2003:299 CIV

LULEÅ UNIVERSITY OF TECHNOLOGY

MASTER'S THESIS

Spin in Quantum Physics

General Theory and Application on "The Proton Spin Crisis"

CHRISTIAN TÜRK

MASTER OF SCIENCE PR OGRAMME

Department of Applied Physics and Mechanical Engineering
Division of Physics

2003:299 CIV • ISSN: 1402 - 1617 • ISRN: LTU - EX - - 03/299 - - SE

LULEÅ UNIVERSITY OF TECHNOLOGY

Master's Thesis

Spin in Quantum Physics

General Theory and Application on 'The Proton Spin Crisis'

Christian Türk

Division of Physics
Luleå University of Technology
SE-97187 Luleå
Sweden

Abstract

The theoretical framework for small scale phenomena is *quantum physics*, where the word "quantum" refers to the smallest possible package of a physical quantity. Especially, in this thesis, we consider the spin of elementary particles, a kind of *"intrinsic" angular momentum.* This property is peculiar to quantum theories and we first discuss how it is connected to non-relativistic (low energy) processes. In this approach the spin is not automatically contained in the theory. It is rather experimental evidence, such as the Stern-Gerlach experiment, which shows that spin must be introduced to fully explain all observations. We also look at the connection to spin in relativistic high energy theories. This requires a knowledge of the Poincaré group, as this group determines the structure of space-time to a large extent.

We also discuss the article by E. Wigner from 1939, where he classified the little group connected to the inhomogeneous Lorentz group, and all its fundamental representations. This allowed Wigner to classify fundamental particles according to their masses and spins. Spin is revealed to result from the symmetries of space-time.

Finally, we try to introduce a phenomenological "hybrid particle"-model, composed of quarks and gluons, of the proton in order to explain the infamous *"Proton Spin Crisis"* problem, the experimental observation that little or nothing of the spin of a proton seems to be carried by the quarks of which it consists. This was first observed by the European Muon Collaboration (EMC) at CERN in 1988.

The conclusions so far are that spin is a property not fully understood and that a better understanding is necessary for obtaining more accurate theories. It is also a powerful and sensitive "tool" to test different theories of nature.

Contents

Introduction

— *Physicist is an atom's way of knowing about atoms.* —
G. Wald

The science of physics for small scale phenomena in the world of atoms is attributed to *quantum physics*, where the word "quantum" refers to the smallest possible "package" of a physical quantity. The Standard Model of today classifies the elementary particles by their invariant mass and spin, where spin is a kind of *"intrinsic" angular momentum* not yet fully understood. This classification is originally due to the work of E. Wigner in 1939 [**1**] where he classified all the unitary irreducible representations of the inhomogeneous Lorentz group and the corresponding little group.

Spin physics plays a very fundamental role in particle physics and it is an important "tool" to gain insight especially into the theory of strong interactions or the color force between gluons and quarks, the building blocks of hadrons. The spin property is peculiar to quantum theories. It is not predicted by the non-relativistic theory, but inserted "by hand" to explain experimental facts. In a relativistic approach it is however an automatic and inherent property.

One of the most infamous problems is the *"Proton Spin Crisis"* [**2**] based on experimental observations by the European Muon Collaboration (EMC) in 1988 [**3**], an experiment at CERN scattering muons off polarized protons. The theoretical models available did not agree with the experimental facts and this is actually a very common feature whenever spin is considered more carefully. This catalyzed several experimental investigations and different theoretical aspects on the problem were proposed. None of the explanations are yet fully comprehensive and verified by experimental observations, see [**4**] for further details. Therefore, experimental observations on spin observables such as polarization, spin correlation and spin asymmetries [**5**] provides important information on the dynamical properties of the interactions between particles, and it hopefully will lead to an agreement between theory and experimental observations of the strong interaction.

In this thesis, we will in Chapter 2 discuss the theories for spin in a non-relativistic regime starting from the classical angular momentum. As it turns out the theory must be modified in order to contain spin. The relativistic regime is discussed in Chapter 3 together with the Poincaré group. This group includes the "little group" of Wigner which leaves the four-momentum of a particle invariant. The results show that spin emerges automatically by considering special relativity and quantum mechanics together. In the last chapter we discuss the "Proton Spin Crisis" and also present the first steps towards an alternative model of the proton in a phenomenological sense, in terms of "hybrid particles" consisting of both quarks and gluons.

CHAPTER 1

Historical Overview of Spin

— *Models are to be used, not believed.* —
H. Teil

The notion and idea of *spin* and its connection to elementary particles such as electrons has its origin in the attempts to describe spectral lines and their multiplicity in atoms. The rotation of an electron about its own axis, i.e. the electron spin, was first proposed by G. E. Uhlenbeck and S. A. Goudsmit in 1925 [**6**] as a real physical characteristic and replaced other *ad hoc* interpretations at that time. The story begins with the discovery of the multiplicity of spectral terms and the anomalous Zeeman effect.

In 1913 N. Bohr published a theory of the spectrum of the hydrogen atom, the simplest atom, with only one electron and one proton, where he proposed three quantum numbers. The principal quantum number $n = 1, 2, 3...$ which described the size of the orbit of the electron, a subordinate quantum number k, $n \geq k$, related to the angular momentum and therefore determining the shape of the orbital and finally a magnetic number m. This final number indicates the component along the vector $\mathbf{k}$, parallel to the magnetic field, and takes the values $-k \leq m \leq k$. However, there are some limits such as the Zeeman effect or in the presence of an external magnetic field a single energy level split into $2k + 1$ levels. The introduction of those quantum numbers made it possible to classify spectral terms of atoms other than hydrogen and especially for hydrogen-like systems, i.e. atoms with a heavier nucleus but still only one electron, by specifying n, k and m. As it turned out, the three numbers were not sufficient to account for all levels and in 1920 A. J. W. Sommerfeld introduced a fourth quantum number j and called it the inner quantum number. In this new convention the previous m now instead specifies the sub-levels which are split for a level specified by n, j and k in the presence of an external magnetic field. This splitting is also of Zeeman-type but its pattern is slightly different from the one before the use of j and experiments showed additional selection rules concerning the multiplicity. All these features were incorporated differently in models by Sommerfeld, A. Landé and W. Pauli who competed with each other to classify multiplicity and its Zeeman effect by introducing their own selection rules. At that time there were different opinions on how to treat an atom and the most favored one was a core built on the nucleus and all the electrons except for one, the outermost electron, sometimes called the radiant electron.

The model by Landé was based on this fact and he called it an *Ersatzmodell.* By using this model many valid results came out when it was compared to experiments but it had some minor problems which Pauli pointed out. Instead Pauli derived a new selection of rules, probably since he never though in terms of models, and it lead to an important conclusion, the fact that the origin of the multiplicity is not the core but the electron itself. The ideas of Pauli concerned the criticism of the assumption that the atomic core is in the K shell, that the multiplicity is not due to the interaction of the core and the radiant electron but a characteristic of the electron itself. He introduced the set of four quantum numbers n, k, j and m and he tried to avoid the use of a model to describe the numbers. But this new adopted idea about the multiplicity related the quantum number j and therefore m to belong to the electron itself and it had a "classically indescribable two-valuedness". With this new set of number he was able to explain how to fill the orbits of an atom with a very clear rule, the well-known *Pauli's exclusion principle.* Pauli published his ideas and, with the paper of Uhlenbeck and Goudsmit [**6**] who introduce the spinning electron, served as a guideline for the understanding of the anomalous Zeemann effect.

The work of L. H. Thomas of the doublet splitting as well as the work of R. de L. Kronig with the account for the electron self-rotating finally confirmed the notion of *spin as a conceptual property* in the description of the electron and it had a great impact on atomic physics. Most of this material is taken from [**7**] and we recommend it for the interested reader.

CHAPTER 2

Spin Physics in Non-Relativistic Processes

— *Nature does nothing without purpose or uselessly.* —
Aristotle

A great achievement of the twentieth-century is the theory about light and matter known as quantum mechanics which in particular describes the law of nature at an atomic level. At this level of fundamental particles and atoms, the classical theory of mechanics is not comprehensive enough to explain all observations and a higher accuracy to a certain limit is achieved with a non-relativistic form of quantum mechanics.

Compared to classical mechanics, where we have a "material" particle with its path determined by the equations of motion, a physical system in quantum mechanics is determined by a "wave" function. As a consequence the formalism in the description of physical events dramatically changes and we will discuss how the angular momentum is affected with spin as one important outcome.

Atomic behavior is very unlike daily life observation and the non-relativistic approach was resolved during the 1920s by people such as N. Bohr, W. Heisenberg and E. Schrödinger. They managed to build a consistent description of phenomena on small scales with new physical concepts such as the wave-particle duality, the uncertainty principle, etc.

The transition between the classical picture and the quantum regime is governed by the following replacements

$$E \mapsto \hat{E} = i\hbar\partial_t, \;\; \mathbf{p} \mapsto \hat{\mathbf{p}} = -i\hbar\nabla, \tag{2.1}$$

or in a covariant notation

$$p^\mu \mapsto \hat{p}^\mu = i\hbar\partial^\mu. \tag{2.2}$$

REMARK 2.0.1. An expression is distinguished between a classical quantity and an operator in quantum mechanics by a caret, such as $\mathbf{p}$ and $\hat{\mathbf{p}}$.

The classical energy momentum relation for a free particle is

$$E = \frac{\mathbf{p}^2}{2m}, \tag{2.3}$$

and by the use of (2.1) we can write

$$i\hbar\partial_t\psi(t,\mathbf{x}) = \frac{1}{2m}(-i\hbar)^2\nabla^2\psi(t,\mathbf{x}) \Leftrightarrow i\hbar\frac{\partial\psi(t,\mathbf{x})}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi(t,\mathbf{x})}{\partial\mathbf{x}^2}, \tag{2.4}$$

where the operators act on a function $\psi(t, \mathbf{x})$ denoted as the wave function of the particle. If we include the interaction of particles by a potential $U(\mathbf{x})$, see for example [**8**], we get the general Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\psi(t,\mathbf{x}) = \hat{\mathbf{H}}\psi(t,\mathbf{x}), \tag{2.5}$$

where $\hat{\mathbf{H}}$ is the Hamiltonian operator

$$\hat{\mathbf{H}} = \frac{\hat{\mathbf{p}}}{2m} + U(\mathbf{x}). \tag{2.6}$$

In this chapter we discuss the equivalent of the classical angular momentum of a particle in quantum mechanics and we also make a review of the Stern-Gerlach experiment in Section 2.2 showing that elementary particles possesses an extra internal degree of freedom similar to the angular momentum. The theory is not complete and must be modified to fit experimental observations and we consider a more general angular momentum in Section 2.3 with its close connection to rotations and finally in Section 2.4 the property of spin for non-relativistic processes .

2.1. A First Glimpse of Angular Momentum

The law of conservation of angular momentum is a consequence of the isotropy of space, all directions are equivalent, with respect to a closed system, both in the classical limit and also in the viewpoint of quantum mechanics. It is only a special case of the more famous and celebrated theorem of E. Noether [**9**] from 1918 which relates a continuous symmetry or invariance, in this case rotational, to a conservation law. This statement has been used extensively by physicists in the effort to understand natural phenomenon and in particle physics the principle is used to a large extent.

The classical notion of the angular momentum of a particle in a spherically potential $U(\mathbf{r})$ is defined as

$$\mathbf{l} = \mathbf{r} \times \mathbf{p}, \tag{2.7}$$

where $\mathbf{r}$ is a distance and $\mathbf{p}$ is the momentum. By the use of (2.1) the angular momentum (2.7) becomes

$$\hat{\mathbf{l}} = -i\hbar\left(\mathbf{r} \times \nabla\right), \tag{2.8}$$

where the cartesian components are

$$\hat{l}_x = -i\hbar\left(y\partial_z - z\partial_y\right) \tag{2.9a}$$
$$\hat{l}_y = -i\hbar\left(z\partial_x - x\partial_z\right) \tag{2.9b}$$
$$\hat{l}_z = -i\hbar\left(x\partial_y - y\partial_x\right). \tag{2.9c}$$

EXAMPLE 2.1.1. The commutator of $\hat{l}_x$ and $\hat{l}_y$ is

$$
\begin{aligned}
[\hat{l}_x, \hat{l}_y]\psi(\mathbf{x}) = & \left(\hat{l}_x\hat{l}_y - \hat{l}_y\hat{l}_x\right)\psi(\mathbf{x}) = \\
& -\hbar^2 \left[(y\partial_z - z\partial_y)(z\partial_x - x\partial_z)\right. \\
& \left. - (z\partial_x - x\partial_z)(y\partial_z - z\partial_y)\right]\psi(\mathbf{x}) = \\
& -\hbar^2 \left[y\psi_x + yz\psi_{xz} - yx\psi_{zz} - z^2\psi_{xy} + zx\psi_{zy} - zy\psi_{zx}\right. \\
& \left. + z^2\psi_{yx} + xy\psi_{zz} - x\psi_y - xz\psi_{yz}\right] = \\
& -\hbar^2 \left[y\psi_x - x\psi_y\right] = i\hbar\hat{l}_z\psi(\mathbf{x}).
\end{aligned}
$$

We find that

$$[\hat{l}_x, \hat{l}_y] = i\hbar\hat{l}_z. \tag{2.10}$$

and in a similar way

$$[\hat{l}_y, \hat{l}_z] = i\hbar\hat{l}_x, \ [\hat{l}_z, \hat{l}_x] = i\hbar\hat{l}_y. \tag{2.11}$$

The commutator relations (2.10) and (2.11) can be summarized as

$$\hat{\mathbf{l}} \times \hat{\mathbf{l}} = i\hbar\hat{\mathbf{l}}, \tag{2.12}$$

where the nature of the operator of the angular momentum is clearly seen. If $\hat{\mathbf{l}}$ would be a vector the cross product would be zero. One of the basic principles in quantum mechanics is that any two operators representing measurable quantities can only be measured simultaneously to arbitrary precision if they commute with each other [**8, 10**]. In the case of the angular momentum any two components do not commute and therefore cannot be measured simultaneously. Since the angular momentum is a constant of motion its operators commute with the Hamiltonian

$$[\hat{\mathbf{H}}, \hat{\mathbf{l}}] = 0, \tag{2.13}$$

of the system

$$\hat{\mathbf{H}}\psi(t, \mathbf{x}) = E\psi(t, \mathbf{x}), \tag{2.14}$$

which determines the energy, E, of the particle determined by $\psi(t, \mathbf{x})$. A commuting set of observables is one of the components of $\hat{\mathbf{l}}$ together with $\hat{\mathbf{H}}$.

The operator representing the square of the magnitude of the angular momentum is defined as

$$\hat{\mathbf{l}}^2 = \hat{l}_x^2 + \hat{l}_y^2 + \hat{l}_z^2. \tag{2.15}$$

It commutes with each of the operators $\hat{l}_x$, $\hat{l}_y$ and $\hat{l}_z$, verified by direct calculations.

EXAMPLE 2.1.2. The commutator of $\hat{\mathbf{l}}^2$ and $\hat{l}_x$ is

$$\begin{aligned}[\hat{\mathbf{l}}^2, \hat{l}_x] = & [\hat{l}_x^2 + \hat{l}_y^2 + \hat{l}_z^2, \hat{l}_x] = [\hat{l}_x^2, \hat{l}_x] + [\hat{l}_y^2, \hat{l}_x] + [\hat{l}_z^2, \hat{l}_x] = \\ & \hat{l}_x[\hat{l}_x, \hat{l}_x] + [\hat{l}_x, \hat{l}_x]\hat{l}_x + \hat{l}_y[\hat{l}_y, \hat{l}_x] + [\hat{l}_y, \hat{l}_x]\hat{l}_y \\ & +\hat{l}_z[\hat{l}_z, \hat{l}_x] + [\hat{l}_z, \hat{l}_x]\hat{l}_z = \\ & -\hat{l}_y i\hbar\hat{l}_z - i\hbar\hat{l}_z\hat{l}_y + \hat{l}_z i\hbar\hat{l}_y + i\hbar\hat{l}_y\hat{l}_z = \\ & i\hbar\left(\hat{l}_z\hat{l}_y + \hat{l}_y\hat{l}_z - \hat{l}_y\hat{l}_z - \hat{l}_z\hat{l}_y\right) = 0,\end{aligned}$$

and both operators commute with each other since

$$[\hat{\mathbf{l}}^2, \hat{l}_x] = 0. \tag{2.16}$$

It also holds for $\hat{l}_x$ and $\hat{l}_y$ by similar calculations

$$[\hat{\mathbf{l}}^2, \hat{l}_y] = [\hat{\mathbf{l}}^2, \hat{l}_z] = 0. \tag{2.17}$$

The commutator relations (2.16) and (2.17) can be summarized as

$$[\hat{\mathbf{l}}^2, \hat{\mathbf{l}}] = 0, \tag{2.18}$$

and we may choose our complete set of commuting observables based on the Hamiltonian operator $\hat{\mathbf{H}}$, a component of the angular momentum $\hat{l}_z$[1] and $\hat{\mathbf{l}}^2$. All the information, or properties, of a particle is determined by its wave function, where the angular momentum classifies the states according to their transformation properties under rotation of the coordinate system. This can be accomplished by the two commuting operators $\hat{l}_z$ and $\hat{\mathbf{l}}^2$, since there exists simultaneous eigenfunctions connected to each operator and also two eigenvalues, usually denoted l and m in the literature. First we change the coordinate system to a spherical polar coordinate system (r, θ, ϕ) and each component of the angular momentum (2.9) can be rewritten as

$$\hat{l}_x = -i\hbar\left(-\sin\phi\partial_\theta - \cot\theta\cos\phi\partial_\phi\right) \tag{2.19a}$$

$$\hat{l}_y = -i\hbar\left(\cos\phi\partial_\theta - \cot\theta\sin\phi\partial_\phi\right) \tag{2.19b}$$

$$\hat{l}_z = -i\hbar\partial_\phi. \tag{2.19c}$$

The use of the $\hat{l}_z$ operator on the time-independent wave function gives

$$\hat{l}_z\psi(r, \theta, \phi) = m\hbar\psi(r, \theta, \phi) \Leftrightarrow -i\hbar\partial_\phi\psi(r, \theta, \phi) = m\hbar\psi(r, \theta, \phi), \tag{2.20}$$

Its solution is

$$\psi(r, \theta, \phi) = F(r, \theta)\Phi_m(\phi), \;\; \Phi_m(\phi) \equiv e^{im\phi}, \tag{2.21}$$

[1]We adopt the usual conventional to consider the z-component of the angular momentum operator and also has the simplest form in spherical polar coordinates.

where $F(r,\theta)$ is an arbitrary function. The wave function $\psi(r,\theta,\phi)$ must be single-valued (we have not yet introduced the possibility of an internal degree of freedom) the spin, that can be double-valued, it must be 2π-periodic in ϕ

$$\Phi_m(2\pi) = \Phi_m(0) \Rightarrow e^{im2\pi} = e^0 \Leftrightarrow \cos m2\pi + i\sin m2\pi = 1, \tag{2.22}$$

a kind of rotational symmetry around the z-axis. Hence, a measurement of the z-component of the angular momentum can only yield the values $0, \pm\hbar, \pm 2\hbar, \ldots$ since the restriction that $\psi(r,\theta,\phi)$ must be single-valued gives

$$m = 0, \pm 1, \pm 2, \ldots\ . \tag{2.23}$$

The quantum number m is usually denoted as the *magnetic quantum number*, as it is close connected to the Zeeman effect. Because the z-axis can be chosen along an arbitrary direction, each component of the angular momentum must be quantized as well. If the eigenfunctions $\Phi_m(\phi)$ of $\hat{l}_z$ satisfy

$$\langle \Phi_m | \Phi_{m'} \rangle = \delta_{mm'} \tag{2.24}$$

they are orthonormal. The left hand side of (2.24) can be expanded as

$$\langle \Phi_m | \Phi_{m'} \rangle = \int_0^{2\pi} \Phi_m^* \Phi_{m'} \mathrm{d}\phi = \int_0^{2\pi} e^{-im'\phi} e^{im\phi} \mathrm{d}\phi = \int_0^{2\pi} e^{i(m'-m)\phi} \mathrm{d}\phi. \tag{2.25}$$

If $m' = m$:

$$\langle \Phi_m | \Phi_{m'} \rangle = \int_0^{2\pi} \mathrm{d}\phi = 2\pi,$$

and $m' \neq m$:

$$\begin{aligned}\langle \Phi_m | \Phi_{m'} \rangle &= \int_0^{2\pi} \cos((m'-m)\phi)\mathrm{d}\phi + i\int_0^{2\pi} \sin((m'-m)\phi)\mathrm{d}\phi = \\ &\frac{1}{m'-m}\left[\underbrace{\sin((m'-m)2\pi)}_{=0} - i\left[\underbrace{\cos((m'-m)2\pi)}_{=1} - 1\right]\right] = 0.\end{aligned}$$

Clearly the eigenfunctions are orthogonal and if we also divide each function with the factor $\sqrt{2\pi}$ they are orthonormal. The orthonormal eigenfunctions of the operator $\hat{l}_z$ are defined as

$$\Phi_m(\phi) \equiv (2\pi)^{-1/2} e^{im\phi},\ m = 0, \pm 1, \pm 2, \ldots\ , \tag{2.26}$$

and we have a stationary wave function as

$$\psi_m(r,\theta,\phi) = R(r)\Theta(\theta)\Phi_m(\phi), \tag{2.27}$$

where the choice of factorization will be proven to be very convenient. The use of (2.19) gives

$$\hat{\mathbf{l}}^2 = -\hbar^2 \left[\frac{1}{\sin\theta}\partial_\theta\left[\sin\theta\partial_\theta\right] + \frac{1}{\sin^2\theta}\partial_{\phi\phi}\right], \tag{2.28}$$

and it only depends on θ and ϕ. As it is a purely angular operator it commutes with any arbitrary function of r, which is also true for $\hat{l}_z$. But, before we try to determine the simultaneous eigenfunctions of $\hat{\mathbf{l}}^2$ and $\hat{l}_z$ we also take a look at the eigenvalues of the square of the angular momentum.

The stationary states which only differs in the value of m actually have the same energy, see for example [**8**], thus the energy levels whose angular momentum is conserved are always degenerate (except for a zero value). Therefore we assume that $\psi_m(r, \theta, \phi)$ is a state with the same value of $\hat{\mathbf{l}}^2$ and belongs to one degenerate energy level distinguished by the value of m. Since the directions of the z-axis are physical equivalent, there exists for each $m = |m|$ a corresponding value $m = -|m|$. Now let $l \geq 0$ denote the greatest possible value of $|m|$ for the given $\hat{\mathbf{l}}^2$. The existence of this upper limit follows from

$$\hat{\mathbf{l}}^2 = \hat{l}_x^2 + \hat{l}_y^2 + \hat{l}_z^2 \Leftrightarrow \hat{\mathbf{l}}^2 - \hat{l}_z^2 = \underbrace{\hat{l}_x^2 + \hat{l}_y^2}_{\geq 0}, \tag{2.29}$$

and the eigenvalues of the operator $\hat{\mathbf{l}}^2 - \hat{l}_z^2$ cannot be negative. Instead of using $\hat{l}_x$ and $\hat{l}_y$ there exists a useful and convenient combination as

$$\hat{l}_\pm = \hat{l}_x \pm i\hat{l}_y, \tag{2.30}$$

with the following relations

$$[\hat{l}_+, \hat{l}_-] = 2\hbar\hat{l}_z, \;\; [\hat{l}_z, \hat{l}_\pm] = \pm\hbar\hat{l}_\pm \tag{2.31a}$$

$$\hat{\mathbf{l}}^2 = \hat{l}_+\hat{l}_- + \hat{l}_z^2 - \hbar\hat{l}_z = \hat{l}_-\hat{l}_+ + \hat{l}_z^2 + \hbar\hat{l}_z \; . \tag{2.31b}$$

Now we can write

$$\begin{aligned} \hat{l}_z\hat{l}_\pm\psi_m &= \{(2.31)\} = \left(\hat{l}_\pm\hat{l}_z \pm \hbar\hat{l}_\pm\right)\psi_m = \\ &\hat{l}_\pm m\hbar\psi_m \pm \hbar\hat{l}_\pm\psi_m = \hbar\,(m \pm 1)\,\hat{l}_\pm\psi_m, \end{aligned} \tag{2.32}$$

and $\hat{l}_\pm\psi_m$ is an eigenfunction to $\hat{l}_z$ with an eigenvalue of $\hbar(m \pm 1)$, apart from some constant connected to the normalization condition. In total we have

$$\psi_{m+1} = c_1\hat{l}_+\psi_m \tag{2.33a}$$

$$\psi_{m-1} = c_2\hat{l}_-\psi_m, \tag{2.33b}$$

where c_i are arbitrary constants. The operators (2.30) are therefore respectively *raising* and *lowering* operators, similar to the ladder operators A and $A^\dagger$ in the treatment of the linear harmonic oscillator, see [**8**]. If $m = l$

$$\hat{l}_+\psi_l = 0,$$

since there exists no states with $m > l$, and if we operate $\hat{l}_-$ on both sides

$$\begin{aligned}\hat{l}_-\hat{l}_+\psi_l = 0 \Rightarrow \hat{l}_-\hat{l}_+\psi_l &= \{(2.31)\} = \left(\hat{\mathbf{l}}^2 - \hat{l}_z^2 - \hbar\hat{l}_z\right)\psi_l = \\ &\hat{\mathbf{l}}^2\psi_l - \hat{l}_z^2\psi_l - \hbar\hat{l}_z\psi_l = \\ &a_{l^2}\psi_l - \hbar^2 l^2\psi_l - \hbar^2 l\psi_l = \\ &\underbrace{\left(a_{l^2} - \hbar^2 l^2 - \hbar^2 l\right)}_{=0}\psi_l = 0,\end{aligned} \tag{2.34}$$

and the corresponding eigenvalues of $\hat{\mathbf{l}}^2$ can be determined as

$$a_{l^2} = \hbar^2 l(l+1). \tag{2.35}$$

The expression (2.35) determines the eigenvalues to the square of the angular momentum $\hat{\mathbf{l}}^2$ and takes all positive integral numbers as well as zero. For a given value of l usually denoted as the *azimuthal quantum number*, the component $\hat{l}_z$ eigenvalues are limited to

$$m = -l, (-l+1), ..., (l-1), l. \tag{2.36}$$

Thus, the energy level specified by an angular momentum l has a $(2l+1)$-fold degeneracy, called degeneracy with respect to the direction of the angular momentum.

We now turn to the problem of obtaining the common eigenfunctions to the operators $\hat{\mathbf{l}}^2$ and $\hat{l}_z$. Since they are purely angular we only need to consider the angular part of the wave function $\psi_m(r, \theta, \phi)$[2]. We have two eigenvalue equations

$$\hat{l}_z Y_{lm}(\theta, \phi) = m\hbar Y_{lm}(\theta, \phi), \tag{2.37}$$

and

$$\hat{\mathbf{l}}^2 Y_{lm}(\theta, \phi) = l(l+1)\hbar^2 Y_{lm}(\theta, \phi), \tag{2.38}$$

together with the results on the eigenvalues. The eigenfunctions are denoted as $Y_{lm}(\theta, \phi)$ and can be factorized as

$$Y_{lm}(\theta, \phi) = \Theta_{lm}(\theta)\Phi_m(\phi), \tag{2.39}$$

where $\Phi_m(\phi)$ is given by (2.26). We also require the normalization condition

$$\langle Y_{lm}|Y_{l'm'}\rangle = \delta_{ll'}\delta_{mm'}. \tag{2.40}$$

One method of obtaining the required functions is by directly solving the problem of finding the eigenfunctions of the operator $\hat{\mathbf{l}}^2$ (2.28) in spherical coordinates, (2.38) becomes

$$-\hbar^2\left[\frac{1}{\sin\theta}\partial_\theta\left(\sin\theta\partial_\theta\right) + \frac{1}{\sin^2\theta}\partial_{\phi\phi}\right]Y_{lm}(\theta, \phi) = \hbar^2 l(l+1)Y_{lm}(\theta, \phi),$$

[2] The wave function of a particle is not completely determined by only considering l and m since the eigenfunctions can contain an arbitrary factor depending on a function of r

or

$$\left[\frac{1}{\sin\theta}\partial_\theta\left(\sin\theta\partial_\theta\Theta_{lm}(\theta)\right) - \frac{m^2}{\sin^2\theta}\Theta_{lm}(\theta) + l(l+1)\Theta_{lm}(\theta)\right]\Phi_m(\phi) = 0. \quad (2.41)$$

From the theory of special functions and especially spherical harmonics, see for example [**11**] for an introduction, we can identify the following differential equation from (2.41)

$$\frac{1}{\sin\theta}\partial_\theta\left(\sin\theta\partial_\theta\Theta_{lm}(\theta)\right) - \frac{m^2}{\sin^2\theta}\Theta_{lm}(\theta) + l(l+1)\Theta_{lm}(\theta) = 0, \quad (2.42)$$

where the corresponding solutions are called associated Legendre polynomials $P_l^m(\cos\theta)$. The solutions satisfy the conditions for finiteness and single-valuedness for positive integral values of $l \geq |m|$ and the condition (2.40) for the solutions of $\Theta_{lm}(\theta)$ reads

$$\int_0^\pi |\Theta_{lm}|^2 \sin\theta \mathrm{d}\theta = 1 \quad (2.43)$$

which gives the final solutions for $m \geq 0$ as

$$\Theta_{lm}(\theta) = (-1)^m \left[\frac{(2l+1)(l-m)!}{2(l+m)!}\right]^{\frac{1}{2}} P_l^m(\cos\theta). \quad (2.44)$$

For negative m values we have

$$\Theta_{l,-|m|} = (-1)^m \Theta_{l|m|}. \quad (2.45)$$

The angular momentum eigenfunctions are just spherical harmonic functions normalized in a particular way and the complete expression for the eigenfunctions $Y_{lm}(\theta,\phi)$ is

$$\begin{aligned} Y_{lm}(\theta,\phi) &= (-1)^m \left[\frac{(2l+1)(l-m)!}{4\pi(l+m)!}\right]^{\frac{1}{2}} P_l^m(\cos\theta)e^{im\phi},\ m \geq 0 \\ &= (-1)^m Y^*_{l,-m}(\theta,\phi),\ m < 0. \end{aligned} \quad (2.46)$$

REMARK 2.1.3. A state determined by l and m is written as $|lm_l\rangle$ in the most general form without any specific representation. In the position representation $|lm_l\rangle$ is equivalent to $|Y_{lm}\rangle$. Another possible representation is a matrix representation discussed in Section 2.3.3.

2.2. Experimental Evidence of A Twofold Ghost

In 1922 [**12**], O. Stern and W. Gerlach conducted an experiment suggested by Stern in 1921 [**13**], where he proposed a way to measure the magnetic moments of atoms. Today this is known as the *Stern-Gerlach experiment* and it measures the deflection of an atomic beam caused by an inhomogeneous magnetic field.

A hot oven is used to create a source of an atomic beam and a system of slits creates a narrow and almost parallel beam which enters between the poles of a magnet. The inhomogeneous magnetic field causes a deflection of the beam and a detector at the end of the setup detects the various trajectories, the apparatus is shown in Figure 2.1. In the beginning, Stern and Gerlach used silver atoms which allowed them to study the effect on a single electron since there is only one outermost electron and it moves in a Coulomb potential caused by the protons and shielded by all the other electrons. If an atom with a magnetic moment $\boldsymbol{\mu}$ is placed in a magnetic field $\mathbf{B}$, a net force $\mathbf{F}$ acts on the atom where each component is

$$F_i = \boldsymbol{\mu}\partial_i \mathbf{B}, \; i = x, y, z. \tag{2.47}$$

However, the magnets shape gives only one non-vanishing force-component. This causes a deflection of the beam in the z-direction. As the z-component of the magnetic moment is proportional to L_z we may expect two different outcomes:

i. *The classical picture.* All orientations of the magnetic moment μ_z are possible since L_z is a continuous parameter, and therefore we expect to see a continuous smear of strikes on the collector.

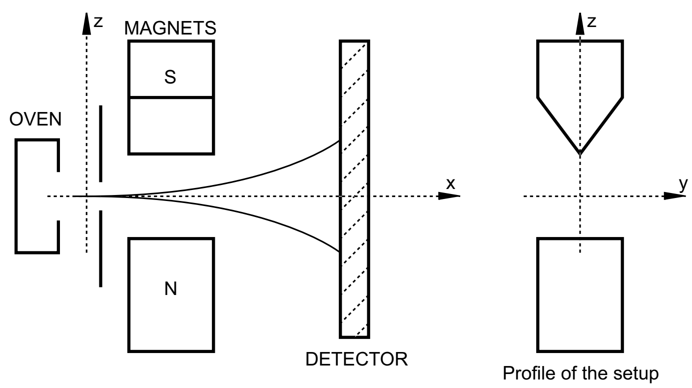

FIGURE 2.1. This is a schematic picture of the experimental setup of the Stern-Gerlach experiment. A beam of atoms is created in an oven and collimated through a system of slits, the beam must be as narrow and parallel as possible, before it enters between the poles of a magnet. A profile of the magnets is also shown and there is a detector at the end of the setup. It detects the deflection of the atomic beam.

ii. *The quantum picture.* We must instead consider the operator $\hat{l}_z$ and we expect a $2l+1$ splitting of the atomic beam trajectories. For $l=0$ the beam should be unaffected and for $l=1$, the collector should show three distinct deflections of the atomic beam, one for $z=0$ (no deflection) and two for $z=\pm z_0$ where z_0 is some constant value.

The surprising result by Stern and Gerlach was that neither of the two different explanations could be correct. First, there was no continuous smear of the strikes and it was one of several different experiments and evidence showing that the classically picture was insufficient. Most textbooks on quantum mechanics discuss and explain these different experiments, see for example [**8, 14**]. Second, indeed there was a kind splitting except with one problem. It was *not* a $2l+1$ splitting. The beam should be unaffected since the considered electron has a zero angular momentum $l=0$ but it was separated into two distinct parts. There was one upper and lower part symmetrically about the point of no deflection. Similar results were also found for other atoms such as gold and copper and later on by others for sodium, potassium, caesium and hydrogen.

If we denote the multiplicity by α we have that

$$\alpha_{theory} \neq \alpha_{experiment}. \tag{2.48}$$

But, an agrement between the theory and experiment could be accomplished if we would also allow non-integral values of l such as $l=1/2$

$$\alpha_{theory} = 2l+1 = 2(\frac{1}{2})+1 = 2 = \alpha_{experiment}. \tag{2.49}$$

The theoretical explanation for the outcome of the Stern-Gerlach experiment came in 1925 by G. E. Uhlenbeck and S. A. Goudsmit [**6**] when they analyzed the Zeeman effect, the splitting of spectral lines from atoms placed in a magnetic field. The electron itself possesses an *"intrinsic" angular momentum* or *spin* independent of its *orbital angular momentum.* The spin is usually described by the *spin quantum number s* and can be both integral and non-integral values where the azimuthal quantum number l can only be integral ones. The discovery of the electron spin had a very fundamental impact on physics in general and it is known today that all particles can be assigned an intrinsic angular momentum with a corresponding quantum number s.

We conclude the Stern-Gerlach experiment by noting that a particle state $|\psi\rangle$ require an additional spin state decoupled from the orbital angular momentum state $|lm_l\rangle$, and this fact is implemented by writing the state for a particle with momentum $\mathbf{p}$ as

$$|\psi\rangle \propto \underbrace{|\mathbf{p}\rangle}_{\text{"Space"}} \otimes \underbrace{|sm_s\rangle}_{\text{"Spin"}}. \tag{2.50}$$

The first factor refers to the usual kinematic degrees of freedom such as the energy and the orbital angular momentum. Our new factor is the spin degree of freedom and we cannot use the eigenstates $|sm_s\rangle \equiv |Y_{lm}\rangle$ as a representation, since the rules must also include half-odd integers. The eigenfunctions $Y_{lm}(\theta, \phi)$ for the orbital angular momentum $l = 1/2$ are

$$Y_{1/2,\pm 1/2} \propto \sqrt{\sin\theta} e^{\pm i \frac{\phi}{2}}, \tag{2.51}$$

and the lowering operator $\hat{l}_-$ gives

$$\hat{l}_- Y_{1/2,1/2} \propto \frac{\cos\theta}{\sqrt{\sin\theta}} e^{-i\frac{\phi}{2}}. \tag{2.52}$$

This, however is not proportional to $Y_{1/2,-1/2}$ if we compare with (2.51), and it indicates the problem to establish the rules to include half-odd integers.

The theory developed in Section 2.1 for the orbital angular momentum is not comprehensive enough to include the spin. One of the reasons for this problem comes from the condition that the state $|\psi\rangle$ of the wave function should have an unique value at each point in space (2.22) for the orbital momentum. Since the spin is not associated with the spatial dependence of the wave function this uniqueness does not apply for the spin. Fortunately, this issue can be solved by the representations considered in Section 2.3.3, namely the state of the particle is considered as a vector and the operators as matrices acting on this vector. The way to build such representations in general is discussed in the next section.

2.3. A More General Angular Momentum

In the classical viewpoint we adopt the following picture of a rotation. Assume that we have a dynamical system defined by all the position and momentum vectors $(\mathbf{x}, \mathbf{p})$. A rotation is possible if we can rotate all the dynamical quantities, in this case $(\mathbf{x}, \mathbf{p})$, while the distance element stays invariant. This new system is a rotated system compared to the original one. The rotation of physical quantities influence nearby all classical physical areas and it is of direct importance in quantum mechanics as well. There is a close connection to a more general angular momentum and influences areas such as atomic, molecular, condensed matter, nuclear and particle physics. The outcome of rotations in the quantum mechanical world is not always simple to understand by intuitive means. The reason lies in that a particle is defined as a wave function. If we assume that a particle is in a state $|\psi\rangle$ the rotated state is defined such that all expectation values of all operators in the rotates state are rotated relative to the original values, there is a similarity to vectors in Euclidian space. But, what does it mean physically to rotate a quantum state?

The answer to this question is based on the conceptual foundations of quantum mechanics and therefore this becomes a very fundamental issue and we leave it at this stage. Instead we try to explain how to understand a rotation of a quantum state by starting with the classical rotation group in three space dimensions, $SO(3)$, and see how these rotations are represented in quantum mechanics. This can be accomplished by finding the unitary operators connected to the rotations and we give the general strategy on how to find these representations by means of the irreducible representations of the general angular momentum. In principle it is not possible to have a complete description of rotations in quantum mechanical systems by only considering the pure group $SO(3)$ and we must instead use the special unitary group $SU(2)$. Both groups have a similar structure but the topological global effects are different and one reason is that the property of spin must be taken into account. For the more interested reader we recommend one of the textbooks [**15, 16, 17, 18**].

2.3.1. Rotations in Ordinary Space. We first assume that we work in a well-defined Euclidian coordinate system with an origin $\mathcal{O}$. This reference system S is spanned by a set of three orthogonal unit vectors $\hat{\mathbf{e}}_{(i)}$. In an abstract sense rotations are operators acting on this three-dimensional space and maps all the points into other points except one which remains fixed. All distances should remain invariant.

REMARK 2.3.1. We denote r as the physical *rotation operator* and follow the notations in [**4**].

The components of a rotated three-dimensional vector $\mathbf{A}^r$ are related to $\mathbf{A}$ by

$$A_i^r = R_{ij}(r)A_j, \tag{2.53}$$

where the elements of the 3×3 matrix $\mathbf{R}$ depend on a given rotation r. In the literature one usually speaks about two different types of rotations, the *active* point of view where the object is rotated compared to the *passive* one, instead the coordinate axis rotate. As the rotation of an axis is accomplished by an active rotation we simply state that all the rotations are of the first type. Meanwhile, $\mathbf{A}$ is visualized by a set of three orthogonal unit vectors $\hat{\mathbf{e}}_{(j)}$ where each can be rotated by the relation (2.53). But, each $\hat{\mathbf{e}}^r_{(i)}$ can also be expressed as a linear superposition of the non-rotated bases $\hat{\mathbf{e}}_{(j)}$, *videlicet*

$$\hat{\mathbf{e}}^r_{(i)} = R(r)_{ji}\hat{\mathbf{e}}_{(j)} = (\mathbf{R}^T)_{ij}\hat{\mathbf{e}}_{(j)}. \tag{2.54}$$

Thus, a fixed vector $\mathbf{A}$ seen in a reference system S by an observer $\mathcal{O}$

$$\mathbf{A} = A_j\hat{\mathbf{e}}_{(j)}, \tag{2.55}$$

is described as

$$\mathbf{A} = (A_l)_{S^r}\hat{\mathbf{e}}^r_{(l)} \tag{2.56}$$

by the observer $\mathcal{O}^r$ in a rotated frame S^r. We get the equality

$$A_j\hat{\mathbf{e}}_{(j)} = (A_l)_{S^r}\hat{\mathbf{e}}^r_{(l)} = (A_l)_{S^r}(\mathbf{R}^T)_{lj}\hat{\mathbf{e}}_{(j)}, \tag{2.57}$$

and

$$A_j = (A_l)_{S^r}(\mathbf{R}^T)_{lj} = R(r)_{jl}(A_l)_{S^r} \Rightarrow (A_l)_{S^r} = [R(r)_{jl}]^{-1}A_j. \tag{2.58}$$

The use of $[R(r)_{jl}]^{-1} \equiv R_{kn}(r^{-1})$ finally gives

$$(A_k)_{S^r} = R_{kn}(r^{-1})A_n. \tag{2.59}$$

In general the rotation matrices $R_{ij}(r)$ form a representation of the geometrical rotation operator r and these matrices belong to the orthogonal group $O(3)$, for the definition see [**16**]. But we need to distinguish between right- and lefthanded coordinate systems, done via

$$\det R_{ij}(r) = \pm 1. \tag{2.60}$$

We only use righthanded systems and therefore we usually work with the special orthogonal group $SO(3)$, we speak of *proper* rotations. One representation of the rotation matrices is based on a *axis-angle parametrization* and we denote $\mathbf{R}_{\hat{\mathbf{n}}}(\theta)$ as a rotation around an arbitrary axis by an angle θ, we use the right hand rule. From [**16**] we have

$$\mathbf{R}_{\hat{\mathbf{e}}_{(1)}}(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}, \tag{2.61}$$

and similar matrices for $\mathbf{R}_{\hat{\mathbf{e}}_{(2)}}$ and $\mathbf{R}_{\hat{\mathbf{e}}_{(3)}}$. From a geometrical sense it is not very difficult to see that rotations around a fix axis commute with each other

$$\mathbf{R}_{\hat{\mathbf{n}}}(\theta_1)\mathbf{R}_{\hat{\mathbf{n}}}(\theta_2) = \mathbf{R}_{\hat{\mathbf{n}}}(\theta_2)\mathbf{R}_{\hat{\mathbf{n}}}(\theta_1). \tag{2.62}$$

From the work initiated by S. Lie and and E. Noether a century ago on continuous groups of geometrical transformations and the connected invariance, see for example [**15, 17**], it is much more useful to consider the rotations as *infinitesimal* or *near-identity* rotations,

$$\mathbf{R} = \mathbf{I} + \epsilon\mathbf{B} + \mathcal{O}(\epsilon^2), \tag{2.63}$$

where $\mathbf{B}$ is a matrix, ϵ a small parameter and $\mathbf{R}$ represents rotations by an infinitesimal angle about some arbitrary axis. The success of Lie is that we can determine the rotation matrices by considering the Lie algebra of the group. The algebra is determined by the generators and the exponential of them contains the essential information to determine the transformations. Whenever the words "Lie algebra" is mentioned one should think of the infinitesimal rotations of some symmetry group and with the commutators we can in principle use them to build the finite symmetry transformations. For the rotation group $SO(3)$ the Lie algebra is given by the commutators

$$[l_x, l_y] = l_z, \ [l_y, l_z] = l_x, \ [l_z, l_x] = l_y, \tag{2.64}$$

where $\mathbf{l}$ is given by (2.7) and the rotation matrix $\mathbf{R}_{\hat{\mathbf{n}}}(\theta)$ can be determined

$$\mathbf{R}_{\hat{\mathbf{n}}}(\theta) = e^{\theta(\hat{\mathbf{n}}\cdot\mathbf{l})}, \tag{2.65}$$

with the results of Lie and Noether. This explicit shows the connection between the classical angular momentum and invariant rotations in ordinary space.

EXAMPLE 2.3.2. The rotation matrix $\mathbf{R}_{\hat{\mathbf{e}}_{(1)}}(\theta)$ is determined by

$$\mathbf{R}_{\hat{\mathbf{e}}_{(1)}}(\theta) = e^{\theta(\hat{\mathbf{e}}_{(1)} \cdot \mathbf{l})} = e^{\theta(l_x)} =$$
$$\mathbf{I} + \theta l_x + \frac{1}{2!}(\theta l_x)^2 + \frac{1}{3!}(\theta l_x)^3 + ... = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}, \quad (2.66)$$

which gives the same result from the axis-angle parametrization and the representation of l_x [**17**] is taken to be

$$l_x \equiv \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}. \quad (2.67)$$

A much more useful parametrization of rotations is the *Euler angles*. This parametrization is based on three angles, usually denoted (α, β, γ). First we consider the rotation of the z-axis and it can be specified by the two angles (α, β), the spherical angles,

$$\hat{\mathbf{e}}'_{(3)} = \mathbf{R}_1(\alpha, \beta)\hat{\mathbf{e}}_{(3)}, \ \mathbf{R}_1(\alpha, \beta) \equiv \mathbf{R}_{\hat{\mathbf{e}}_{(3)}}(\alpha)\mathbf{R}_{\hat{\mathbf{e}}_{(2)}}(\beta). \quad (2.68)$$

However, this rotation cannot specify the most general rotation since we can get the $\hat{\mathbf{e}}'_{(1)}$- and $\hat{\mathbf{e}}'_{(2)}$-axis wrong. We need a third rotation defined by an other angle (γ)

$$\mathbf{R} = \mathbf{R}_{\hat{\mathbf{e}}'_{(3)}}(\gamma)\mathbf{R}_1(\alpha, \beta), \quad (2.69)$$

and this determines the general rotation matrix. Now we have a mixing of both the "old" and "new" axes and this is not very convenient. From [**19**] we have the important identity

$$\mathbf{R}_{\hat{\mathbf{e}}'_{(3)}}(\gamma)\mathbf{R}_1(\alpha, \beta) = \mathbf{R}_1(\alpha, \beta)\mathbf{R}_{\hat{\mathbf{e}}_{(3)}}(\gamma), \quad (2.70)$$

and the usual zyz-convention for the Euler angles is

$$\mathbf{R}(\alpha, \beta, \gamma) = \mathbf{R}_{\hat{\mathbf{e}}_{(3)}}(\alpha)\mathbf{R}_{\hat{\mathbf{e}}_{(2)}}(\beta)\mathbf{R}_{\hat{\mathbf{e}}_{(3)}}(\gamma). \quad (2.71)$$

There exists other conventions as well but this particular one is very well suited for quantum mechanical applications, we usually speak about a well-defined z-axis. The Euler angles have a one-to-one correspondence to the group manifold $SO(3)$ apart from the exceptional points of the angles, the end of their ranges. From a geometrical viewpoint two rotations do not commute in general and in fact the group $SO(3)$ is non-Abelian. If we look closer to the commutator relations of the generators (2.64) they clearly do not commute. The full proper rotation matrix can now be determine by

$$\mathbf{R}(\alpha, \beta, \gamma) = e^{\alpha(l_z)}e^{\beta(l_y)}e^{\gamma(l_z)}. \quad (2.72)$$

So far we have presented the most important concept on how to obtain a representation of rotations in ordinary space. The next step is to implement all of them into quantum mechanical systems.

2.3.2. The Quantum Mechanical Rotations. In quantum mechanics it is common to represent the effects of an operation by an operator acting directly on the states and can in a very general form be written as

$$|\psi\rangle' = U|\psi\rangle. \tag{2.73}$$

We do not give any representations of the states nor the operator U which represents an arbitrary operation of some kind. We must preserve the probability for physical interpretations

$$|\langle\phi|\psi\rangle|^2 = |'\langle\phi|\psi\rangle'|^2 = |\langle\phi|\underbrace{U^\dagger U}_{=1}|\psi\rangle|^2, \tag{2.74}$$

showing that the operator U must be unitary.

PROPOSITION 1 ([**17**]). *We assume that the unitary proper rotation operators are associated to the classical rotations* $\mathbf{R} \in SO(3)$ *by*

$$\mathbf{R} \mapsto U(\mathbf{R}), \tag{2.75}$$

with U *as a function of* $\mathbf{R}$. *If the rotation operators satisfy*

$$U(\mathbf{I}) = 1,\ U(\mathbf{R}_1)U(\mathbf{R}_2) = U(\mathbf{R}_1\mathbf{R}_2), \tag{2.76}$$

and

$$U(\mathbf{R}^{-1}) = U(\mathbf{R})^{-1} = U(\mathbf{R})^\dagger, \tag{2.77}$$

we say that $U(\mathbf{R})$ *forms a* representation of $SO(3)$ by means of unitary operators.

REMARK 2.3.3. The statement of Proposition 1 is actually to strong for rotations of spin systems as they do not have any spatial dependence. More on this matter is discussed later on and in Section 2.4.

The way to find the unitary representations of the rotations can be based on the use of infinitesimal rotations as

$$\mathbf{R}_{\hat{\mathbf{n}}}(\theta) = \mathbf{I} + \theta\,(\hat{\mathbf{n}} \cdot \mathbf{l}) + \mathcal{O}(\theta^2) \mapsto U(\mathbf{R}_{\hat{\mathbf{n}}}(\theta)) = 1 - \frac{i}{\hbar}\theta\left(\hat{\mathbf{n}} \cdot \hat{\mathbf{j}}\right) + \mathcal{O}(\theta^2), \tag{2.78}$$

where $\hat{\mathbf{j}}$ is a vector of Hermitian operators and the factor $i/\hbar$ is conventional. We do not yet give a specific explanation of these operators. From the theory of S. Lie and E. Noether as earlier we can obtain the association between the classical rotations and the corresponding unitary operators by the use of exponents

$$\mathbf{R}_{\hat{\mathbf{n}}}(\theta) = e^{\theta(\hat{\mathbf{n}} \cdot \mathbf{l})} \mapsto U_{\hat{\mathbf{n}}}(\theta) = e^{-i\theta\left(\hat{\mathbf{n}} \cdot \hat{\mathbf{j}}\right)/\hbar}. \tag{2.79}$$

Now we must consider the components of $\hat{\mathbf{j}}$ as generators compatible to the classical case of rotations in ordinary space. It can also be shown that these generators must as well satisfy the familiar commutator relations

$$[\hat{j}_x, \hat{j}_y] = i\hbar\hat{j}_z, \ [\hat{j}_y, \hat{j}_z] = i\hbar\hat{j}_x, \ [\hat{j}_z, \hat{j}_x] = i\hbar\hat{j}_y. \tag{2.80}$$

More specific, the standard commutator relations of the orbital angular momentum operators $\hat{\mathbf{l}}$. We do not interpret $\hat{\mathbf{j}}$ as $\hat{\mathbf{l}}$ since they are not sufficient for the spin property, deduced from experimental facts, and we make the following definition:

DEFINITION 2.3.4 ([**14**]). A vector operator $\hat{\mathbf{j}}$ is a *general angular momentum* operator if its components are Hermitian operators satisfying

$$[\hat{j}_x, \hat{j}_y] = i\hbar\hat{j}_z, \ [\hat{j}_y, \hat{j}_z] = i\hbar\hat{j}_x, \ [\hat{j}_z, \hat{j}_x] = i\hbar\hat{j}_y, \tag{2.81}$$

equivalent to

$$\hat{\mathbf{j}} \times \hat{\mathbf{j}} = i\hbar\hat{\mathbf{j}}. \tag{2.82}$$

Finally we have a general strategy to obtain the unitary operators and their representations for classical rotations with the three basic steps as follows:

(i) Find the representation of the most general $\hat{\mathbf{j}}$ with the given commutators (2.82) in some given basis.
(ii) Exponent the linear combination of these representations and it determines the corresponding unitary rotation operators $U_{\hat{\mathbf{n}}}(\theta)$.
(iii) Explore the physical properties of such rotations.

2.3.3. Representations of $\hat{\mathbf{j}}$. In quantum mechanics there is a close connection between the angular momentum operators and unitary rotation operators. From the previous section the first important step towards this connection requires the knowledge of the representations of the general angular momentum.

We start in analog to Section 2.1 but without adopting any specific representation of the states we are interested to obtain the eigenvalues of the commuting operators $\hat{\mathbf{j}}^2$ and $\hat{j}_z$ together with the simultaneous eigenvectors $|jm_j\rangle$. First, the eigenvalues of $\hat{\mathbf{j}}^2$ and $\hat{j}_z$ are real since both operators are Hermitian and the eigenvalues of $\hat{\mathbf{j}}^2$ must be either positive or zero. Denoting the simultaneous eigenvectors of the two operators by $|jm_j\rangle$ we have a similar situation as with the orbital angular momentum

$$\hat{\mathbf{j}}^2|jm_j\rangle = j(j+1)\hbar^2|jm_j\rangle, \tag{2.83}$$

and

$$\hat{j}_z|jm_j\rangle = m_j\hbar|jm_j\rangle. \tag{2.84}$$

For later convenience we have written the eigenvalues of $\hat{\mathbf{j}}^2$ in the form $j(j+1)\hbar^2$ and for $\hat{j}_z$ as $m_j\hbar$. Second, because the expectation value of the square of a

Hermitian operator must be positive, or zero, we have

$$\langle \hat{\mathbf{j}}^2 \rangle = \langle \hat{j_x}^2 \rangle + \langle \hat{j_y}^2 \rangle + \langle \hat{j_z}^2 \rangle \geq \langle \hat{j_z}^2 \rangle, \tag{2.85}$$

and it follows that

$$j(j+1) \geq m_j. \tag{2.86}$$

We introduce the raising and lowering operators

$$\hat{j_\pm} = \hat{j_x} \pm \hat{j_y}, \tag{2.87}$$

and from Section 2.1 we have that $\hat{j_\pm}|jm_j\rangle$ are simultaneous eigenvectors of $\hat{\mathbf{j}}^2$ and $\hat{j_z}$ belonging to the eigenvalues $j(j+1)\hbar^2$ and $(m_j \pm 1)\hbar$, respectively. If we repeatedly operate $\hat{j_\pm}$ on the state $|jm_j\rangle$, a sequence of eigenvectors can be constructed with the eigenvalues $(m_j \pm 1)\hbar$, $(m_j \pm 2)\hbar$, $(m_j \pm 3)$ and so on, of which each of them being an eigenvector to $\hat{\mathbf{j}}^2$ with the eigenvalue $j(j+1)\hbar^2$. From the condition (2.86) we see that there must exist a maximum and minimum eigenvalue, namely $m_j^{max}\hbar$ and $m_j^{min}\hbar$, since we cannot continue to infinitely in either direction. The difference of the maximum and minimum value must be positive or zero

$$m_j^{max} - m_j^{min} = k, \; k = 0, 1, 2, \ldots \;, \tag{2.88}$$

and the following holds

$$\hat{j_+}|jm_j^{max}\rangle = 0 = \hat{j_-}|jm_j^{min}\rangle. \tag{2.89}$$

This together with the relations (2.31) where we only replace $\hat{\mathbf{l}}$ with $\hat{\mathbf{j}}$ yields

$$\begin{aligned} 0 = \hat{j_\mp}\left[\hat{j_\pm}|jm_j^{\substack{max \\ min}}\rangle\right] &= \left[\hat{\mathbf{j}}^2 - \hat{j_z} \mp \hbar\hat{j_z}\right]|jm_j^{\substack{max \\ min}}\rangle = \\ &\underbrace{\left[j(j+1) - \left(m_j^{\substack{max \\ min}}\right)^2 \mp m_j^{\substack{max \\ min}}\right]}_{=0}\hbar^2|jm_j^{\substack{max \\ min}}\rangle, \end{aligned} \tag{2.90}$$

and we get

$$\left(m_j^{max}\right)^2 + m_j^{max} = j(j+1) = \left(m_j^{min}\right)^2 - m_j^{min}, \tag{2.91}$$

or

$$\left(m_j^{max}\right)^2 + m_j^{max} = \left(m_j^{min}\right)^2 - m_j^{min}. \tag{2.92}$$

The solutions of (2.92) are

$$m_j^{max} = -m_j^{min}, \; m_j^{max} = m_j^{min} - 1, \tag{2.93}$$

and since (2.88) must hold the only acceptable solution is $m_j^{max} = -m_j^{min}$ and with equation (2.91)

$$m_j^{max} = j = -m_j^{min}. \tag{2.94}$$

As the difference (2.88) between m_j^{max} and m_j^{min} must be a positive integer or zero

$$2j = k,\ \ k = 0, 1, 2, ...\ \ \Rightarrow j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, ...\ , \tag{2.95}$$

and the allowed $2j+1$ values for m_j are $(-j, -j+1, ..., j-1, j)$. We have not yet considered any representation of the operators $\hat{\mathbf{j}}$, but the listed values of j are compatible with the commutator relations (2.82). For any specific operators satisfying these relations it may be that in some cases only some values of j are present while other are absent. It depends on how the operators are parameterized or on their representations. The orbital angular momentum do have the same commutator relations, but there are only integer values of j present, denoted as l. The particular multiplicity connected to the spin, obtained from the Stern-Gerlach experiment, should allow half-odd integers and we immediately see how well the vector operator $\hat{\mathbf{j}}$ is suited for spin systems. In principle we may use $\hat{\mathbf{j}}$ as the most general representation of both the orbital angular momentum as well as the spin.

Let us now return to the simultaneous eigenvectors $|jm_j\rangle$ connected to the two operators $\hat{\mathbf{j}}^2$ and $\hat{j}_z$. These eigenvectors can be used as a basis for any space on which the angular momentum commutator relations act, see [**8, 17**]. As the operators are Hermitian, the eigenvectors are orthogonal and with some normalization constant we have the orthonormality condition

$$\langle j'm_j'|jm_j\rangle = \delta_{j'j}\delta_{m_j'm_j}. \tag{2.96}$$

The fact that the two operators commute with each other

$$[\hat{\mathbf{j}}^2, \hat{j}_z] = 0, \tag{2.97}$$

implies, *videlicet*

$$\begin{aligned}0 = [\hat{\mathbf{j}}^2, \hat{j}_z] = \langle j'm_j'|[\hat{\mathbf{j}}^2, \hat{j}_z]|jm_j\rangle = \langle j'm_j'|\hat{\mathbf{j}}^2\hat{j}_z - \hat{j}_z\hat{\mathbf{j}}^2|jm_j\rangle = \\ \langle\hat{\mathbf{j}}^2 j'm_j'|\hat{j}_z|jm_j\rangle - \langle j'm_j'|\hat{j}_z|\hat{\mathbf{j}}^2 jm_j\rangle = \\ j'(j'+1)\hbar^2\langle j'm_j'|\hat{j}_z|jm_j\rangle - j(j+1)\hbar^2\langle j'm_j'|\hat{j}_z|jm_j\rangle = \\ \hbar^2\left(j'(j'+1) - j(j+1)\right)\langle j'm_j'|\hat{j}_z|jm_j\rangle,\end{aligned} \tag{2.98}$$

for all the eigenvectors given by (2.96). The last factor reads

$$\langle j'm_j'|\hat{j}_z|jm_j\rangle = m_j\hbar\langle j'm_j'|jm_j\rangle = m_j\hbar\delta_{j'j}\delta_{m_j'm_j}, \tag{2.99}$$

and it can be represented as the matrix elements of the operator $\hat{j}_z$. But, the matrix elements can only be nonzero if $j' = j$ or only between the states that have the same total general angular momentum quantum numbers

$$\langle jm_j'|\hat{j}_z|jm_j\rangle = m_j\hbar\delta_{m_j'm_j}. \tag{2.100}$$

EXAMPLE 2.3.5. The matrix representation of the $\hat{j}_z$ operator for $j = 1$ is determined by

$$\langle 1m'_j|\hat{j}_z|1m_j\rangle = m_h\hbar\delta_{m'_j m_j} \tag{2.101}$$

where m'_j and m_j take the values 1, 0 and -1. The only nonzero elements are

$$\hat{j}_{z1,0} = -\hat{j}_{z-1,-1} = \hbar, \tag{2.102}$$

and we have a 3×3 matrix as

$$\left[\hat{j}_{zm'_j m_j}\right]_{j=1} = \hbar \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \tag{2.103}$$

which could also be used as a representation for the orbital angular momentum $\hat{l}_z$ with $l = 1$. For half-odd integers of j such as $j = 1/2$

$$\left[\hat{j}_{zm'_j m_j}\right]_{j=1/2} = \frac{\hbar}{2} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \tag{2.104}$$

and this will be proven to be very useful when the spin is taken into account and could serve as a representation for $s = 1/2$.

The matrix elements of $\hat{\mathbf{j}}^2$ can be calculated directly by its definition but we are still missing the matrix representations of $\hat{j}_x$ and $\hat{j}_y$. In order to find these matrices it is convenient to first obtain those of the raising and lowering operators

$$\hat{j}_\pm = \hat{j}_x \pm i\hat{j}_y, \tag{2.105}$$

as we can rearrange the two relations to

$$\hat{j}_x = \frac{1}{2}\left(\hat{j}_+ + \hat{j}_-\right) \tag{2.106a}$$

$$\hat{j}_y = \frac{i}{2}\left(\hat{j}_- - \hat{j}_+\right). \tag{2.106b}$$

Since the raising and lowering operators commute with $\hat{\mathbf{j}}^2$

$$[\hat{\mathbf{j}}^2, \hat{j}_\pm] = 0, \tag{2.107}$$

the same principles and calculations used for $\hat{j}_z$ holds for $\hat{j}_+$ and $\hat{j}_-$. The matrix elements are determined by

$$\langle j'm'_j|\hat{j}_\pm|jm_j\rangle = c_\pm\langle j'm'_j|j(m_j \pm 1)\rangle = c_\pm\delta_{j'j}\delta_{m'_j m_j \pm 1}. \tag{2.108}$$

We used equation (2.33) with $\hat{l}_\pm$ replaced by $\hat{j}_\pm$. The constants $c_\pm$ are

$$\begin{aligned}|c_\pm|^2 &= |c_\pm|^2 \cdot 1 = |c_\pm|^2 \langle j(m_j \pm 1)|j(m_j \pm 1)\rangle = \\ &\langle j(m_j \pm 1)||c_\pm|^2|j(m_j \pm 1)\rangle = \\ &\langle c_\pm j(m_j \pm 1)|c_\pm j(m_j \pm 1)\rangle = \\ &\langle \hat{j}_\pm jm_j|\hat{j}_\pm jm_j\rangle = \langle jm_j|\hat{j}_\mp \hat{j}_\pm|jm_j\rangle = \\ &\langle jm_j|\hat{\mathbf{j}}^2 - \hat{j}_z^{\,2} \mp \hbar \hat{j}_z|jm_j\rangle = \\ &(j(j+1)\hbar^2 - m\hbar^2 \mp m\hbar^2)\langle jm_j|jm_j\rangle = \\ &\hbar^2 \left[j(j+1) - m_j(m_j \pm 1)\right], \end{aligned} \tag{2.109}$$

or

$$c_\pm = \hbar\sqrt{j(j+1) - m_j(m_j \pm 1)}. \tag{2.110}$$

The general expression for the matrix elements now reads

$$\langle j' m_j'|\hat{j}_\pm|jm_j\rangle = \hbar\sqrt{j(j+1) - m_j(m_j \pm 1)}\delta_{j'j}\delta_{m_j' m_j \pm 1} \tag{2.111}$$

and the only nonzero matrix elements are for $j' = j$.

EXAMPLE 2.3.6. The matrix representation of $\hat{j}_z$ for $j = 1$ is given by the matrix (2.103) and for the operators $\hat{j}_\pm$

$$\left[\hat{j}_{\pm m_j' m_j}\right]_{j=1} = \hbar\sqrt{2 - m_j(m_j \pm 1)}\delta_{m_j' m_j \pm 1}, \tag{2.112}$$

where the only nonzero elements are

$$\hat{j}_{+1,0} = \hat{j}_{+0,-1} = \hbar\sqrt{2} \tag{2.113a}$$

$$\hat{j}_{-0,1} = \hat{j}_{--1,0} = \hbar\sqrt{2}. \tag{2.113b}$$

From the relations (2.106) the matrices for $\hat{j}_x$ and $\hat{j}_y$ are

$$\left[\hat{j}_{x m_j' m_j}\right]_{j=1} = \frac{\hbar}{\sqrt{2}}\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad \left[\hat{j}_{y m_j' m_j}\right]_{j=1} = \frac{\hbar}{\sqrt{2}}\begin{bmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{bmatrix}, \tag{2.114}$$

and finally the matrix of $\hat{\mathbf{j}}^2$ becomes

$$\begin{aligned}\left[\hat{\mathbf{j}}^2{}_{m_j' m_j}\right]_{j=1} &= \hbar^2\left(\frac{1}{2}\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right) = \\ &\hbar^2\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right) = \\ &\hbar^2\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = 2\hbar^2\mathbf{I}. \end{aligned} \tag{2.115}$$

We have considered each representation separately for $j' = j$ and each space is spanned by the eigenvectors $|jm\rangle$ with dimension $(2j+1)$. We could take all these representations into a single form, a matrix of infinite rank and block-diagonal form, so each representations by a given value of j is a subspace.

DEFINITION 2.3.7 ([**17**]). A subspace is considered *invariant* under the action of an operator if every vector in this space is mapped into another vector in the same subspace by the operator.

It is a fact, see [**17**], that the considered representation of the operators $\hat{\mathbf{j}}$ posses no smaller subspaces which are invariant under the operator. We speak about invariant subspaces of minimal dimensionality.

DEFINITION 2.3.8 ([**17**]). If a subspace is an invariant subspace of minimal dimensionality, these spaces are called *irreducible invariant subspaces*.

The most convenient way to work with the representations of the general angular momentum is to consider each representation separately since they form invariant subspaces, see Definition 2.3.7. We say that each matrix form an *irreducible representation*, see Definition 2.3.8, of the commutator relations (2.82) and they actually do not depend on the physical interpretation of the vector operator $\hat{\mathbf{j}}$. They are universal matrices that can be used for any problem orbital angular momentum, spin, and isospin.

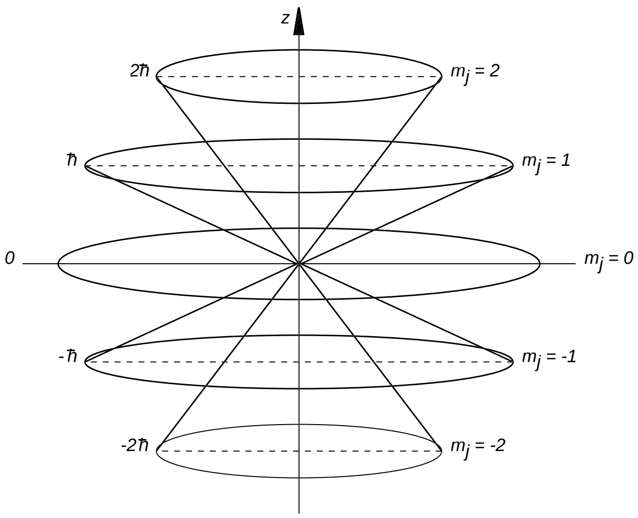

FIGURE 2.2. An illustration of the vector model for the general angular momentum with $j = 2$. The vector $\hat{\mathbf{j}}$ of length $\sqrt{j(j+1)}\hbar$ precesses around the z-axis quantized by the $2j+1$ projections on the axis given by $m_j\hbar$.

We have now a way to obtain the infinite number of matrices representing the different operators connected to the general angular momentum in the basis of the eigenvectors $|jm_j\rangle$. Each matrix is an irreducible representation of the operator and characterized by a given value of j and dimension $(2j+1)\times(2j+1)$. We note that the matrix representation for $l = j = 0, 1, 2, 3, 4, ...$ may be used for the orbital angular momentum and all the properties in Section 2.1 hold as well. For more details see [**8, 14**]. Next we discuss on how to get the rotation representations and explore their physical effects.

We end this section by mentioning a very convenient way to visualize the general angular momentum known as the *vector model*, see [**8, 14**] and Figure 2.2. The model visualize $\hat{\mathbf{j}}$ as a vector of length $\sqrt{j(j+1)}\hbar$ that precesses around the $z-$axis with different positions determined by the $2j+1$ allowed projections of $\hat{\mathbf{j}}$ on the axis given by $m_j\hbar$. Thus, $\hat{\mathbf{j}}$ lies on the surface of a cone with altitude $m_j\hbar$ with the $z-$axis as its symmetry axis.

2.3.4. Unitary Rotation Operators $\mathcal{D}^j(r)$. The components of the general angular momentum vector operator $\hat{\mathbf{j}}$ are considered as generators of rotations in quantum mechanics. In the previous section we considered the representations of those generators and noticed that these generators were more general then $\hat{\mathbf{l}}$ in the sense that they could correspond to different types of angular momentum. We also know that the unitary rotation operators are determined from (2.79)

$$U_{\hat{\mathbf{n}}}(\theta) = e^{-i\theta\left(\hat{\mathbf{n}}\cdot\hat{\mathbf{j}}\right)/\hbar}.$$

If we use the irreducible matrix representations of the general angular momentum, we obtain the irreducible representations of the unitary rotation operators. In analog to rotations in ordinary space we can express them in axis-angle form or in the Euler-angles

$$\mathcal{D}^j_{m_j' m_j}(\hat{\mathbf{n}},\theta) = \langle jm_j'|U_{\hat{\mathbf{n}}}(\theta)|jm_j\rangle \tag{2.116a}$$

$$\mathcal{D}^j_{m_j' m_j}(\alpha,\beta,\gamma) = \langle jm_j'|U(\alpha,\beta,\gamma)|jm_j\rangle, \tag{2.116b}$$

where the representations of $\mathcal{D}^j_{m_j' m_j}(r)$ are the $(2j+1)$-dimensional unitary matrices of rotations r, see [**4**] for more information.[3]

EXAMPLE 2.3.9. The representation of a rotation around the z-axis is a diagonal matrix

$$\begin{aligned}\mathcal{D}^j_{m_j' m_j}(\hat{\mathbf{e}}_{(3)},\theta) &= \langle jm_j'|e^{-i\theta\hat{j}_z/\hbar}|jm_j\rangle = \\ &\langle jm_j'|1 - \frac{i}{\hbar}\theta\hat{j}_z + \frac{1}{2!}(\frac{i}{\hbar}\theta\hat{j}_z)^2 + \mathcal{O}(\theta^3)|jm_j\rangle = \\ &\langle jm_j'|e^{-i\theta m_j}|jm_j\rangle = e^{-i\theta m_j}\delta_{m_j' m_j},\end{aligned} \tag{2.117}$$

[3] $\mathcal{D}$ is a common notation as it can stand for one of the two German words, *Drehung* (rotation) or *Darstellung* (representation).

and the representation of the Euler angles becomes

$$\begin{aligned}\mathcal{D}^j_{m'_j m_j}(\alpha,\beta,\gamma) = \langle jm'_j|e^{-i\alpha\hat{j}_z/\hbar}e^{-i\beta\hat{j}_y/\hbar}e^{-i\gamma\hat{j}_z/\hbar}|jm_j\rangle = \\ \sum_{m^1_j,m^2_j}\langle jm'_j|e^{-\frac{i\alpha}{\hbar}\hat{j}_z}|jm^1_j\rangle\langle jm^1_j|e^{-\frac{i\beta}{\hbar}\hat{j}_y}|jm^2_j\rangle\langle jm^2_j|e^{-\frac{i\gamma}{\hbar}\hat{j}_z}|jm_j\rangle = \\ \sum_{m^1_j,m^2_j} e^{-i\alpha m^1_j}\delta_{m'_j m^1_j}\langle jm^1_j|e^{-i\beta\hat{j}_y/\hbar}|jm^2_j\rangle e^{-i\gamma m_j}\delta_{m^2_j m_j} = \\ e^{-i(\alpha m'_j+\gamma m_j)}\langle jm'_j|e^{-i\beta\hat{j}_y/\hbar}|jm_j\rangle.\end{aligned} \tag{2.118}$$

We notice that the choice of the zyz-convention of the Euler angles is very convenient as we only need to consider one factor multiplied by an exponent.

From the previous example we notice that it is only necessary, besides some exponent, to consider the representation

$$d^j_{m'_j m_j}(\beta) = \langle jm'_j|e^{-i\beta\hat{j}_y/\hbar}|jm_j\rangle \tag{2.119}$$

called the *reduced rotation matrix* for the irreducible representations of the unitary rotation operators by the Euler angles. Since the representations of $\hat{j}_y$ are purely imaginary each element of $d^j_{m'_j m_j}(\beta)$ is purely real and in principle is it sufficient to only consider the reduced rotation matrices for the general rotations. This is due to the zyz-convention which again proves to be very convenient for our purposes.

EXAMPLE 2.3.10. There exists explicit formulas [**4**] for the reduced rotation matrices, and we list the first two non-trivial ones. For $j = 1/2$

$$\left[d^{1/2}_{m'_j m_j}(\beta)\right] = \begin{bmatrix}\cos(\beta/2) & -\sin(\beta/2)\\ \sin(\beta/2) & \cos(\beta/2)\end{bmatrix}, \tag{2.120}$$

and for $j = 1$

$$\left[d^{1}_{m'_j m_j}(\beta)\right] = \begin{bmatrix}\frac{1}{2}(1+\cos\beta) & -\sin\beta/\sqrt{2} & \frac{1}{2}(1-\cos\beta)\\ \sin\beta/\sqrt{2} & \cos\beta & -\sin\beta/\sqrt{2}\\ \frac{1}{2}(1-\cos\beta) & \sin\beta/\sqrt{2} & \frac{1}{2}(1+\cos\beta)\end{bmatrix}. \tag{2.121}$$

We have the representations of the rotations by means of the irreducible representations of the general angular momentum and we now discuss the physical effects. Let ϕ be the infinitesimal angle of a rotation about the z-axis and $\mathcal{D}^j_{m'_j m_j}(\hat{\mathbf{e}}_{(3)},\phi)$ is a diagonal matrix with the phases $e^{-i\theta m_j}$. For a complete 2π-rotation the phases become

$$e^{-i2\pi m_j} = (-1)^{2j}, \tag{2.122}$$

and a rotated state would return to its original value for integral values of j but not for half-odd values. It would require another rotation of 2π. This is actually a nonclassical effect and very important to remember for our treatment of the spin later on. As pointed out earlier we see that the statement of Definition 1 is too strong and we must redefine it.

DEFINITION 2.3.11 ([**17**]). The unitary proper rotation operators are associated to the classical rotations $\mathbf{R} \in SO(3)$ by

$$\mathbf{R} \mapsto \begin{cases} U(\mathbf{R}) \\ U(\mathbf{R} + \mathbf{R}_{2\pi}) = -U(\mathbf{R}), \end{cases} \tag{2.123}$$

with U as a function of $\mathbf{R}$. If the rotation operators satisfy

$$U(\mathbf{I}) = 1,\ U(\mathbf{R}_1)U(\mathbf{R}_2) = U(\mathbf{R}_1\mathbf{R}_2), \tag{2.124}$$

and

$$U(\mathbf{R}^{-1}) = U(\mathbf{R})^{-1} = U(\mathbf{R})^{\dagger}, \tag{2.125}$$

we say that $U(\mathbf{R})$ forms a *representation of* $SO(3)$ *by means of unitary operators.*

From the previous definition we see that the irreducible representations of the unitary rotation operators form a double-valued representation of $SO(3)$ for half-odd integers j, and a single-valued representation for integral j. In general the matrices of $\mathcal{D}^j_{kl}(r)$ form a proper and single valued irreducible representations of the group $SU(2)$, see [**4, 17**]. This group is a standard group and its elements are 2×2 complex unitary matrices with determinant 1, for a formal definition see [**16**]. We end this section with this important conclusion and we try to understand the one-to-two association between $SO(3)$ and $SU(2)$ with the discussion of spin. The spin is closely connected to the general representations of the rotation operators $\mathcal{D}^j_{kl}(r)$ and therefore it can be quite difficult to understand rotations of spin states with this two-valuedness.

2.4. The Physical Property of Spin

The explanation of the Stern-Gerlach experiment demonstrates that a particle must be assigned an "intrinsic" angular momentum, namely the spin with a spin quantum number s. It is very important to remember that an interpretation to the orbital angular momentum is not completely correct and as well in principle no classical interpretation exist. If the electron would be a composite particle, the motion of its constituents parts, the orbital angular momentum would give raise to a $2l + 1$ splitting. Such constituent elements does not exist from what we know today and the electron is regarded as an elementary particle. Thus, the spin of a particle must be ascribed to it whether it is "composite" or "elementary" and completely unconnected to its motion in space. The electron, an elementary particle, is assigned a spin quantum number $s = 1/2$ as well as the proton regarded as a composite particle. This kind of property is peculiar to quantum theories but at this stage a non-relativistic approach does not predict the spin as the theory is modified "by hand" to fit experimental data. In relativistic theories the situation is opposite, spin appears to be a result of the equations and symmetries which forbids it as a decoupled degree of freedom, more on this matter is discussed later on in Chapter 3. One of the conclusions in Section 2.3.3 classifies the orbital angular momentum as a special case of the more general angular momentum but with the same basic structure given by Definition 2.3.4. The spin is also a kind of angular momentum and in non-relativistic quantum mechanics the spin is introduced as an additional degree of freedom by the operator

$$\hat{\mathbf{s}} = (\hat{s}_x, \hat{s}_y, \hat{s}_z), \tag{2.126}$$

and it must satisfy the familiar commutator relations (2.82)

$$\hat{\mathbf{s}} \times \hat{\mathbf{s}} = i\hbar\hat{\mathbf{s}}. \tag{2.127}$$

The irreducible representations obtained in Section 2.3.3 are well suited as the spin quantum number s must take the values $0, 1/2, 1, 3/2, 2, ...$ to agree with experiments. We only replace $\hat{\mathbf{j}}$ with $\hat{\mathbf{s}}$

$$\hat{\mathbf{s}}^2|sm_s\rangle = s(s+1)\hbar^2|sm_s\rangle \tag{2.128a}$$

$$\hat{s}_z|sm_s\rangle = m_s\hbar|sm_s\rangle, \tag{2.128b}$$

and the spin operators $\hat{\mathbf{s}}^2$ and $\hat{s}_z$ can be represented by $(2s+1)\times(2s+1)$ matrices and the simultaneous eigenvectors $|sm_s\rangle$ as column vectors. It is common in literature that the symbol χ_{sm_s} denotes the vector representation of $|sm_s\rangle$. The spin eigenvector has $(2s+1)$-elements, all elements are zero except for one that is equal to unity. The orthonormal condition now reads

$$\chi^\dagger_{sm'_s}\chi_{sm_s} = \delta_{m'_s m_s}, \tag{2.129}$$

where † denotes the Hermitian adjoint. We illustrate the material presented so far with an example.

EXAMPLE 2.4.1. The irreducible representation of the operator $\hat{s}_z$ for a spin one particle is

$$[\hat{s}_z]_{s=1} = \hbar \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \tag{2.130}$$

and the spin eigenvectors are determined by (2.128b). We write

$$\left([\hat{s}_z]_{s=1} - m_s\hbar\mathbf{I}\right)\chi_{1m_s} = \mathbf{0}, \;\; \chi_{1m_s} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}, \tag{2.131}$$

and a, b and c are constants to be calculated. The condition to be satisfied is

$$\det\left([\hat{s}_z]_{s=1} - m_s\hbar\mathbf{I}\right) = 0, \tag{2.132}$$

or

$$(1 - m_s)m_s(1 + m_s) = 0 \Rightarrow m_s = \{1, 0, -1\}. \tag{2.133}$$

For $m_s = 1$ we get $\chi_{1,1}$ from

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \mathbf{0} \Rightarrow \chi_{1,1} = a \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \tag{2.134}$$

and since a is arbitrary it can for simplicity be set to unity. The other eigenvectors are determined by similar calculation and in total we get

$$\chi_{1,1} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \ \chi_{1,0} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \ \chi_{1,-1} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \tag{2.135}$$

The condition (2.129) is also satisfied, for instance

$$\chi_{1,0}^{\dagger}\chi_{1,0} = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = 1, \ \chi_{1,0}^{\dagger}\chi_{1,-1} = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 0. \tag{2.136}$$

The theory of non-relativistic quantum mechanics is sufficient for spinless particles with its state represented by a wave function $\psi(t, \mathbf{x})$, see [**8**]. The function only depends on a position variable $\mathbf{x}$ and time variable t. There is no connection to a spin variable, only experimental facts gave raise to questioning the correctness of the theory, and from the material presented so far we must modify the theory to include the spin. The wave function must as well depend on a spin variable, denoted as σ. This new variable denotes the component of the spin along the z-axis which only can take *discrete* values in contrary to the *continuously* variables $\mathbf{x}$ and t. We assume that the basic postulates in quantum mechanics, discussed in detail in [**8, 10**], still holds if we include σ. The wave function for a particle having spin can be expanded as

$$\psi(t, \mathbf{x}, \sigma) = \sum_{m_s=-s}^{s} \psi_{m_s}(t, \mathbf{x})\chi_{sm_s}, \tag{2.137}$$

whereby the properties of a spin s particle is embedded in a wave function with $2s + 1$ components. Each component corresponds to a particular value of the spin variable.

EXAMPLE 2.4.2. The wave function for a spin one particle is

$$\psi(t,\mathbf{x},\sigma) = \psi_{-1}(t,\mathbf{x})\chi_{1,-1} + \psi_0(t,\mathbf{x})\chi_{1,0} + \psi_1(t,\mathbf{x})\chi_{1,1} = \psi_{-1}(t,\mathbf{x})\begin{bmatrix}0\\0\\1\end{bmatrix} + \psi_0(t,\mathbf{x})\begin{bmatrix}0\\1\\0\end{bmatrix} + \psi_1(t,\mathbf{x})\begin{bmatrix}1\\0\\0\end{bmatrix} = \begin{bmatrix}\psi_1(t,\mathbf{x})\\ \psi_0(t,\mathbf{x})\\ \psi_{-1}(t,\mathbf{x})\end{bmatrix}, \tag{2.138}$$

and each component represents a fixed value of $\sigma = \{\hbar, 0, -\hbar\}$, respectively.

The way to test theories in physics is by experiments and the basic requirement is that a measurement must agree with the predicted one. A measurement of an observable on a particle, described by its wave function, the expectation value of this observable is

$$\langle \hat{A} \rangle = \frac{\langle \psi|\hat{A}|\psi\rangle}{\langle\psi|\psi\rangle}, \tag{2.139}$$

where $\hat{A}$ is the Hermitian operator associated with the dynamical variable. This is one of the basic postulates in quantum mechanics and was deduced from the treatment with a spinless theory. It still holds when the spin is considered by the expansion (2.137). With the condition that the wave function is normalized to unity

$$\langle\psi|\psi\rangle = \int \psi^\dagger(t,\mathbf{x},\sigma)\psi(t,\mathbf{x},\sigma)\mathrm{d}\mathbf{x} = \sum_{m_s=-s}^{s}\int |\psi_{m_s}(t,\mathbf{x})|^2\mathrm{d}\mathbf{x} = 1, \tag{2.140}$$

we interpret $|\psi_{m_s}(t,\mathbf{x})|^2$ as the probability of locating the particle at the time t in the volume $\mathrm{d}\mathbf{x}$ with the component of the spin along the z-axis equal to $m_s\hbar$. The expectation value of an Hermitian operator $\hat{A}$ reads

$$\langle \hat{A} \rangle = \frac{\langle \psi|\hat{A}|\psi\rangle}{\langle\psi|\psi\rangle} = \langle \psi|\hat{A}|\psi\rangle = \int \psi^\dagger(t,\mathbf{x},\sigma)\hat{A}\psi(t,\mathbf{x},\sigma)\mathrm{d}\mathbf{x} = \sum_{m'_s=-s}^{s}\sum_{m_s=-s}^{s}\int \psi^*_{m'_s}(t,\mathbf{x})A_{m'_s m_s}\psi_{m_s}(t,\mathbf{x}), \tag{2.141}$$

with

$$A_{m'_s m_s} \equiv \langle \chi_{sm'_s}|\hat{A}|\chi_{sm_s}\rangle, \tag{2.142}$$

a matrix representing operators $\hat{A}$ in a so called $(2s+1)$-dimensional *spin space*, and note the similarity for having the representations of $\hat{\mathbf{j}}$. We can therefore treat the spin separately in this space and use (2.137) to assemble the final wave function with its first factor unconnected to the spin. In general we can write a spin state of a particle with spin s as a superposition of the spin eigenvectors, the spin wave function becomes

$$\chi_s = \sum_{m_s=-s}^{s} a_{m_s} \hat{\boldsymbol{\eta}}_{(m_s)}, \tag{2.143}$$

where a_{m_s} are constants and $\hat{\boldsymbol{\eta}}_{(m_s)}$ denotes the spin eigenvectors χ_{sm_s}. If $\chi_s^\dagger \chi_s = 1$, the factor $|a_{m_s}|^2$ denotes the probability of finding the particle in the spin state χ_{sm_s} and for orthonormal eigenvectors

$$\sum_{m_s=-s}^{s} |a_{m_s}|^2 = 1. \tag{2.144}$$

So far we have labelled the spin states according to m_s and its reference to a z-axis in a well-defined, fixed coordinate system with an origin $\mathcal{O}$. Although it is not very simple to understand what we mean by physical rotating a spin state in its space, we are interested in the relation between the observations of the same spin state in different reference systems.

2.4.1. Rotations in Spin Space. An observer in our well-defined reference system S describes a spin state according to (2.143). In complete analogy to the rotations of vector in Euclidian space discussed in Section 2.3.1, the components of a spin state χ_s in the rotated frame S^r are related to the non-rotated components by

$$(a_i)_{S^r} = \mathcal{D}^s_{ij}(r^{-1}) a_j \tag{2.145}$$

and the spin eigenvectors

$$\hat{\boldsymbol{\eta}}_{(i)} = \mathcal{D}^s_{ji}(r^{-1}) \hat{\boldsymbol{\eta}}^r_{(j)}, \tag{2.146}$$

where the representation of $\mathcal{D}^s_{ij}(r)$ are the $(2s+1)$-dimensional unitary matrices of rotations r, see Section 2.3.4 and [**14, 4**]. The interpretation in physical terms is that the state $\hat{\boldsymbol{\eta}}_{(i)}$ observed by $\mathcal{O}$ is described by $\mathcal{O}^r$ as a superposition of its own states $\hat{\boldsymbol{\eta}}^r_{(j)}$. This result can be summarized to more general terms. Assume that the observer $\mathcal{O}$ sees a particle with spin s. Its spin state is $|sm_s\rangle$ and the same state becomes

$$|sm_s\rangle_{S^r} = \mathcal{D}^s_{m_s' m_s}(r^{-1}) |sm_s'\rangle, \tag{2.147}$$

in a rotated frame S^r, by the rotation r, in the view of the observer $\mathcal{O}^r$. To extend the ideas on rotations we can also talk about an active rotation of a spin state in the reference system S, but it is not simple to understand physically what happens to the state as the effects with integral values of s are different from those of half-odd integers already mentioned in Section 2.3.4. The phase factor can either be symmetric or antisymmetric for a complete 2π-rotation. We interpret $|sm_s\rangle^r$ as a rotated spin state

$$|sm_s\rangle^r = \mathcal{D}^s_{m_s' m_s}(r) |sm_s'\rangle, \tag{2.148}$$

analogy to represent the effects of a general operator acting on some state vectors. Assume that a particle is in definite state, say χ_s, with well defined coefficients a_{m_s} and $|a_{m_s}|^2$ as the probability to find the particle in a specific state. As χ_s rotates the coefficients changes and as well the probability. The image of a rotation is the change of the probability not compatible with a physical rotation of an object or property in ordinary space. We also note the following very important new behavior of the general wave function (2.137), with respect to rotations about the z-axis. Without the spin factor, a complete rotation of 2π does not alter the wave function as the function is symmetric. This was one of the requirements in the treatment of the angular momentum without the existence of spin. For the spin state χ_s we a complete 2π-rotation becomes

$$\chi_s^r = \mathcal{D}^s_{m_s' m_s}(\hat{\mathbf{e}}_{(3)}, \phi)\chi_s\Big|_{\phi=2\pi} = (-1)^{2s}\chi_s. \tag{2.149}$$

where we have used the representation of $\mathcal{D}^s_{m_s' m_s}(r)$ from Section 2.3.4. Thus, a rotation of 2π about the z-axis returns the general wave function to its original value for particles with integral spin and those with half-odd integers spin change sign.

A wave function can therefore either be symmetric or antisymmetric and this fact is used to separate fundamental particles into two types. Particles having integral values of s are called *bosons* because they obey the Bose-Einstein statistics and those for half-odd integers are denoted as *fermions*, since they instead admit the Fermi-Dirac statistics. The connection between the spin and statistics becomes even clearer in relativistic quantum theories since for the non-relativistic case no such connection exists, by the fact that spin is not predicted.

The physical conclusion is that a spin state behaves differently under rotations depending if the particle has an integral value of s compared to half-odd integers. This supports the use of the mathematical entities known as *spinors*.

DEFINITION 2.4.3 ([**4**]). A mathematical entity S is a *spinor* if it satisfies

$$S(\theta + 2\pi) = -S(\theta), \tag{2.150}$$

where θ is an arbitrary rotation parameter and it is assumed that S can have a sign associated with it.

A geometrical visualization of a spinor is explained with respect to Figure 2.3 and 2.4. The unit circle in the xz-plane with an origin in $\mathcal{O}$ intercepts the z-axis at a positive point P. An arbitrary point on this circle C makes an angle θ with the z-axis, the usual polar angle, and we have a ray $\mathcal{O}C$. If we bisect this angle we get a new point S and this in turn bisects the ray PC because $\mathcal{O}P = \mathcal{O}C$. We have

$$\mathcal{O}S(\theta) = \cos(\frac{\theta}{2}),\ PS(\theta) = \sin(\frac{\theta}{2}), \tag{2.151}$$

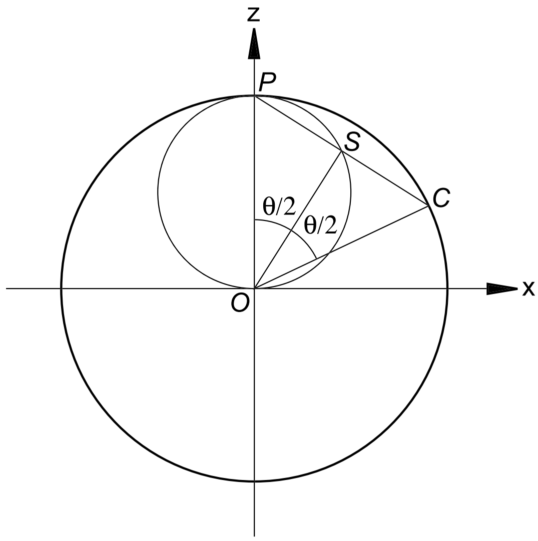

FIGURE 2.3. A geometrical visualization of a spinor for a two-dimensional representation. The two quantities $\mathcal{O}S$ and PS are the components of the spinor, determined by the smaller circle, with respect to the ray $\mathcal{O}C$. For a revolution of the point C, the point S moves from P to $\mathcal{O}$ and back to P by a second revolution of C.

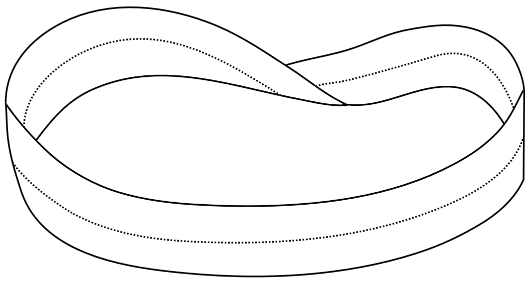

FIGURE 2.4. A popular geometrical visualization of a spinor by the use of the Möbius strip. If we start at one point we end up on the other side with one revolution. A second revolution is required to return to the starting point.

and these quantities are taken to be signed, $\mathcal{O}S$ is negative for $\pi < \theta < 2\pi$ and PS for $2\pi < \theta < 4\pi$. From the first revolution of C, the point S goes from P to $\mathcal{O}$. A second revolution of C returns S to P. The path of S is a new circle and it is traced out in a period of 4π. The quantities $\mathcal{O}S$ and PS fulfill the definition of a spinor as they are signed and change their sign when $\theta \to \theta + \pi$. The pair of $(\mathcal{O}S, PS)$ form a spinor associated with C. In a two-dimensional representation, the spin one half case, one may view the ray $\mathcal{O}C$ as a vector associated to the point C by the polar angle θ. Then $\mathcal{O}S$ and PS are the components of the spinor with respect to the vector $\mathcal{O}C$. The particular nature of a spinor can as well be visualized by the path on the Möbius strip. It takes two revolutions to return to the starting point.

The next section is devoted to the treatment of spin one half particles as it appears to be the most important case. These particles can therefore be represented by spinors since they are fermions and all the results so far must be taken into account. The reason why this particular case is so important has its origin in the building blocks of atoms. All electrons and nucleons appear to be spin one half particles as well as the more elementary constituents of nucleons, the quarks.

2.4.2. Spin-$1/2$ Building Blocks. The theory of spin one half particles in non-relativistic processes was first developed by W. Pauli in 1927[4], to explain the properties of the electron and to achieve an agreement with experimental facts. Several of the important fundamental particles such as the *electron* and *nucleons*, the building blocks of atoms, have spin one half. Indeed, it is also believed that all strongly interacting particles the *hadrons* are shaped by smaller constituents, *the quarks*. These so-called quarks share the same spin properties as the electron and it is the reason why the spin one half case is the most important one.

For any particle having spin one half, the spin variable σ takes the values $\hbar/2$ and $-\hbar/2$. Its wave function is determined by the expansion (2.137) with the spin considered separately. The spin wave function $\chi_{1/2}$ is a superposition of two eigenspinors

$$\chi_{1/2} = a_{\uparrow}\hat{\boldsymbol{\eta}}_{(\uparrow)} + a_{\downarrow}\hat{\boldsymbol{\eta}}_{(\downarrow)}, \tag{2.152}$$

where $\uparrow\downarrow$ are equivalent to $m_s = \pm 1/2$, respectively. It is also required that

$$|a_{\uparrow}|^2 + |a_{\downarrow}|^2 = 1, \tag{2.153}$$

where each factor is referred to the probability of finding a particle in a specific state. Since we work in a well-defined reference system S, each value of m_s is equivalent to the projection of the spin on the z-axis and it is common to speak about the 'spin up" ($\uparrow$) and "spin down" ($\downarrow$) states. The reason for this convention is closely related to the vector model discussed earlier in the treatment of the general angular momentum. The eigenspinors span a two-dimensional spin space and their representations are

$$\hat{\boldsymbol{\eta}}_{(\uparrow)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \; \hat{\boldsymbol{\eta}}_{(\downarrow)} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \tag{2.154}$$

[4]W. Pauli seldom published his papers, he instead preferred to have a correspondence with his colleagues about his research. Still he is credited for a lot of his work on spin.

and we immediately see that the general wave function has two components. This particular function is sometimes called the *Pauli wave function*. Each component $\psi_\uparrow(t,\mathbf{x})$ and $\psi_\downarrow(t,\mathbf{x})$ represents the wave function for a fix value of σ. The two eigenspinors are used as a basis for the spin space, thus the irreducible representations of the spin operators are represented by 2×2 Hermitian matrices according to the results from Section 2.3.3 with $j = s = 1/2$. We have

$$\hat{s}_x = \frac{\hbar}{2}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},\ \hat{s}_y = \frac{\hbar}{2}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix},\ \hat{s}_z = \frac{\hbar}{2}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{2.155a}$$

$$\hat{s}_+ = \hbar\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},\ \hat{s}_- = \hbar\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \tag{2.155b}$$

$$\hat{\mathbf{s}}^2 = \frac{3\hbar^2}{4}\mathbf{I}. \tag{2.155c}$$

Another aspect of the spin one half case is its uniqueness in the sense that the spin operator $\hat{\mathbf{s}}$ can be represented as

$$\hat{\mathbf{s}} = \frac{\hbar}{2}\boldsymbol{\sigma}, \tag{2.156}$$

where $\boldsymbol{\sigma}$ is the set of three 2×2 matrices

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},\ \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix},\ \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{2.157}$$

known as the *Pauli matrices*, see [**8, 14**]. The Pauli spin matrices play a dual role since they form the irreducible representations of the spin operators and together with the identity matrix they form as well a basis for any general 2×2 Hermitian matrix, see [**17**]. Such a matrix can always be written as

$$M = \frac{\hbar}{2}\left(a\mathbf{I} + \mathbf{b}\cdot\boldsymbol{\sigma}\right) = \begin{bmatrix} a + b_3 & b_1 - ib_2 \\ b_1 + ib_2 & a - b_3 \end{bmatrix}, \tag{2.158}$$

with four arbitrary constants and it is a representation of an element in the special unitary group $SU(2)$. The Pauli matrices are the generators and span the Lie algebra $\mathfrak{su}(2)$. This two-dimensional representation is sometimes called the fundamental representation as all other representations of $s = 1, 3/2, 2, 5/2, \ldots$ can be built on these σ_i-matrices, for some examples see [**20**], sometimes leading to confusions. The σ_i-matrices are used for different purposes, it becomes even more obvious in relativistic processes.

Let us now consider the unitary rotation operators and they are determined from (2.79)

$$U_{\hat{\mathbf{n}}}(\theta) = e^{-i\theta\,(\hat{\mathbf{n}}\cdot\hat{\mathbf{s}})\,/\hbar} = \left\{\hat{\mathbf{s}} = \frac{\hbar}{2}\boldsymbol{\sigma}\right\} = e^{-i\theta\,(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})\,/2} =$$
$$\cos(\theta/2)\mathbf{I} - i\,(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})\sin(\theta/2). \tag{2.159}$$

The validity of the last equality of (2.159) can be checked by direct calculations. For example if we expand one factor as

$$
\begin{aligned}
e^{-i\theta\hat{s}_z/\hbar} = e^{-i\theta\sigma_z/2} = \{\epsilon \equiv -\theta/2\} = \\
\mathbf{I} + i\epsilon\sigma_z + \frac{(i\epsilon)^2}{2!}\sigma_z^2 + \frac{(i\epsilon)^3}{3!}\sigma_z^3 + \frac{(i\epsilon)^4}{4!}\sigma_z^4 + ... = \\
\mathbf{I} + i\epsilon\sigma_z + \frac{(i\epsilon)^2}{2!}\mathbf{I} + \frac{(i\epsilon)^3}{3!}\sigma_z + \frac{(i\epsilon)^4}{4!}\mathbf{I} + ... = \\
\left(1 - \frac{(\epsilon)^2}{2!} + \frac{(\epsilon)^4}{4!} - ...\right)\mathbf{I} + i\left(\epsilon - \frac{(\epsilon)^3}{3!} + \frac{(\epsilon)^5}{5!} - ...\right)\sigma_z = \\
\mathbf{I}\cos\epsilon + i\sigma_z\sin\epsilon = \mathbf{I}\cos(\theta/2) - i\sigma_z\sin(\gamma/2),
\end{aligned}
\tag{2.160}
$$

and similar results hold for $\hat{s}_x$ and $\hat{s}_y$. The irreducible representations of the spinor rotations about each axis are

$$\left[\mathcal{D}^{1/2}_{m_s'm_s}(\hat{\mathbf{e}}_{(1)},\theta)\right] = \begin{bmatrix} \cos(\theta/2) & -i\sin(\theta/2) \\ -i\sin(\theta/2) & \cos(\theta/2) \end{bmatrix} \tag{2.161a}$$

$$\left[\mathcal{D}^{1/2}_{m_s'm_s}(\hat{\mathbf{e}}_{(2)},\theta)\right] = \begin{bmatrix} \cos(\theta/2) & -\sin(\theta/2) \\ \sin(\theta/2) & \cos(\theta/2) \end{bmatrix} \tag{2.161b}$$

$$\left[\mathcal{D}^{1/2}_{m_s'm_s}(\hat{\mathbf{e}}_{(3)},\theta)\right] = \begin{bmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{bmatrix}, \tag{2.161c}$$

and the Euler angles

$$
\begin{aligned}
\left[\mathcal{D}^{1/2}_{m_s'm_s}(\alpha,\beta,\gamma)\right] = \begin{bmatrix} e^{-i\alpha/2} & 0 \\ 0 & e^{i\alpha/2} \end{bmatrix}\begin{bmatrix} \cos(\beta/2) & -\sin(\beta/2) \\ \sin(\beta/2) & \cos(\beta/2) \end{bmatrix}\begin{bmatrix} e^{-i\gamma/2} & 0 \\ 0 & e^{i\gamma/2} \end{bmatrix} = \\
e^{-i(\alpha m_j' + \gamma m_j)}\begin{bmatrix} \cos(\beta/2) & -\sin(\beta/2) \\ \sin(\beta/2) & \cos(\beta/2) \end{bmatrix} = \\
e^{-i(\alpha m_j' + \gamma m_j)} d^{1/2}_{m_s'm_s}(\beta).
\end{aligned}
\tag{2.162}
$$

Since we rotate spinors, the effect of the Euler angles are not exactly compatible with the classical case, instead we have

$$0 \le \alpha \le 2\pi,\ 0 \le \beta \le \pi,\ 0 \le \gamma \le 4\pi, \tag{2.163}$$

where the angle (γ) is allowed to be in the range to 4π. It is required by the definition of a spinor, see Definition 2.4.3, as it has a signed associated with it. We note the following property of the $\mathcal{D}^{1/2}$-matrices

$$\mathcal{D}^{1/2}_{m_s'm_s}(r^{-1}) = [\mathcal{D}^{1/2}(r)^{-1}]_{m_s'm_s} = [\mathcal{D}^{1/2}(r)^{\dagger}]_{m_s'm_s} = \mathcal{D}^{1/2*}_{m_sm_s'}(r), \tag{2.164}$$

which also is true for other values of s [**4**].

Now let $\hat{\mathbf{n}}$ be a unit vector in a three-dimensional reference system defined by the unit vectors $\hat{\mathbf{e}}_{(i)}$. We are interested of the component of the spinor along $\hat{\mathbf{n}}$, namely $\hat{\mathbf{n}} \cdot \hat{\mathbf{s}}$, as it is not necessary to keep $\hat{s}_z$ diagonal. First, we consider a spin up state defined by the eigenspinor $\hat{\boldsymbol{\eta}}_{(\uparrow)}$. The spinor rotation that maps this state into the $\hat{\mathbf{n}}$ direction is

$$\left[\mathcal{D}^{1/2}_{m_s' m_s}(\alpha,\beta,0)\right] = \begin{bmatrix} e^{-i\alpha/2} & 0 \\ 0 & e^{i\alpha/2} \end{bmatrix} \begin{bmatrix} \cos(\beta/2) & -\sin(\beta/2) \\ \sin(\beta/2) & \cos(\beta/2) \end{bmatrix} \mathbf{I}. \tag{2.165}$$

which is not an unique one as we could allow any value of the angle (γ). But we do not care about this overall phase as it is totally irrelevant for now. We get

$$\hat{\boldsymbol{\eta}}^{\hat{\mathbf{n}}}_{(\uparrow)} = \begin{bmatrix} e^{-i\alpha/2}\cos(\beta/2) \\ e^{i\alpha/2}\sin(\beta/2) \end{bmatrix}. \tag{2.166}$$

The Cartesian components of $\hat{\mathbf{n}}$ are

$$\hat{\mathbf{n}} = \sin\beta\cos\alpha\hat{\mathbf{e}}_{(x)} + \sin\beta\sin\alpha\hat{\mathbf{e}}_{(y)} + \cos\beta\hat{\mathbf{e}}_{(z)}, \tag{2.167}$$

and

$$\begin{aligned} [\hat{s}_n] &= \hat{\mathbf{n}}\cdot\hat{\mathbf{s}} = \sin\beta\cos\alpha\,[\hat{s}_x] + \sin\beta\sin\alpha\,[\hat{s}_y] + \cos\beta\,[\hat{s}_z] = \\ & \frac{\hbar}{2}\left(\sin\beta\cos\alpha\sigma_x + \sin\beta\sin\alpha\sigma_y + \cos\beta\sigma_z\right) = \\ & \frac{\hbar}{2}\begin{bmatrix} \cos\beta & \sin\beta e^{-i\alpha} \\ \sin\beta e^{i\alpha} & -\cos\beta \end{bmatrix}. \end{aligned} \tag{2.168}$$

Let us investigate the expression $\hat{s}_n\hat{\boldsymbol{\eta}}^{\hat{\mathbf{n}}}_{(\uparrow)}$

$$\begin{aligned} [\hat{s}_n]\,\hat{\boldsymbol{\eta}}^{\hat{\mathbf{n}}}_{(\uparrow)} &= \frac{\hbar}{2}\begin{bmatrix} \cos\beta & \sin\beta e^{-i\alpha} \\ \sin\beta e^{i\alpha} & -\cos\beta \end{bmatrix}\begin{bmatrix} e^{-i\alpha/2}\cos(\beta/2) \\ e^{i\alpha/2}\sin(\beta/2) \end{bmatrix} = \\ & \frac{\hbar}{2}\begin{bmatrix} e^{-i\alpha/2}(\cos\beta\cos(\beta/2) + \sin\beta\sin(\beta/2)) \\ e^{i\alpha/2}(\sin\beta\cos(\beta/2) - \cos\beta\sin(\beta/2)) \end{bmatrix} = \\ & \frac{\hbar}{2}\begin{bmatrix} e^{-i\alpha/2}\cos(\beta-\beta/2) \\ e^{i\alpha/2}\sin(\beta-\beta/2) \end{bmatrix} = \frac{\hbar}{2}\hat{\boldsymbol{\eta}}^{\hat{\mathbf{n}}}_{(\uparrow)}. \end{aligned} \tag{2.169}$$

In fact $\hat{\boldsymbol{\eta}}^{\hat{\mathbf{n}}}_{(\uparrow)}$ is an eigenspinor to the operator $\hat{s}_n$ with the eigenvalue $\hbar/2$, i.e. the component of the spin in the direction of $\hat{\mathbf{n}}$. In this sense we shall say that the spin is "up" in the direction of $\hat{\mathbf{n}}$. Second, for the spin down state $\hat{\boldsymbol{\eta}}_{(\downarrow)}$ the spinor rotation (2.165) maps it to

$$\hat{\boldsymbol{\eta}}^{\hat{\mathbf{n}}}_{(\downarrow)} = \begin{bmatrix} -e^{-i\alpha/2}\sin(\beta/2) \\ e^{i\alpha/2}\cos(\beta/2) \end{bmatrix}. \tag{2.170}$$

and in accordance to (2.169) the following holds

$$\hat{s}_n\hat{\boldsymbol{\eta}}^{\hat{\mathbf{n}}}_{(\downarrow)} = -\frac{\hbar}{2}\hat{\boldsymbol{\eta}}^{\hat{\mathbf{n}}}_{(\downarrow)}. \tag{2.171}$$

We say that the spin is "down" in the direction of $\hat{\mathbf{n}}$.

The spin one half case is unique in the sense that every spinor "points" in some direction and that this spinor is an eigenspinor of $\hat{\mathbf{n}}\cdot\hat{\mathbf{s}}$ in some direction

$\hat{\mathbf{n}}$. It does not hold for higher values of the spin. In the literature one usually speaks about a spinor pointing in that direction or another one and this property can even be used to define helicity states, where the direction is defined by the momentum. This is very useful in relativistic processes and will be discussed more in detail in Chapter 3.

The spin-polarization vector $\mathcal{P}_{\chi_s}$ is extensively used since it can specify a spin state.

DEFINITION 2.4.4 ([**4**]). The *spin-polarization vector* $\mathcal{P}_{\chi_s}$ is defined as

$$\mathcal{P}_{\chi_s} \equiv \langle \hat{\mathbf{s}} \rangle_{\chi_s} / s. \tag{2.172}$$

The knowledge of $\mathcal{P}_{\chi_{1/2}}$ is very important for spin one half since it completely specifies the spin state. We can write an arbitrary spin one half state as

$$\chi_{1/2} = \begin{bmatrix} e^{-i\alpha/2} \cos(\beta/2) \\ e^{i\alpha/2} \sin(\beta/2) \end{bmatrix}, \tag{2.173}$$

without any loss of generality from (2.166) and (2.170). We find that

$$\mathcal{P}_{\chi_{1/2}} = \langle \boldsymbol{\sigma} \rangle_{\chi_{1/2}} = \chi_{1/2}^{\dagger} \boldsymbol{\sigma} \chi_{1/2} = [\sin\beta\cos\alpha, \sin\beta\sin\alpha, \cos\beta]\,. \tag{2.174}$$

which demonstrates the usefulness of the spin-polarization vector $\mathcal{P}_{\chi_{1/2}}$. Since the spin one half case is governed by an $SU(2)$ symmetry it allows us to use all the results on other physical objects that have a similar symmetry. We close this section with an example on isospin of the nucleon.

EXAMPLE 2.4.5. In nuclear physics the proton and neutron may be considered as identical particles if we would neglect the minimal mass difference and the Coulomb interaction of the proton. They are two different states of the nucleon having an internal degree of freedom compatible to the $SU(2)$ symmetry group. The isospin generators are a copy of the spin one half generators

$$\hat{\mathbf{I}} \times \hat{\mathbf{I}} = i\hbar\hat{\mathbf{I}}, \tag{2.175}$$

In the fundamental representation the isospin generators are

$$\hat{\mathbf{I}} = \frac{\hbar}{2}\boldsymbol{\sigma}, \tag{2.176}$$

and they act on the states

$$p = \begin{bmatrix} 1 \\ 0 \end{bmatrix},\ n = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \tag{2.177}$$

considered as the proton and neutron, respectively.

CHAPTER 3

Spin Physics in Relativistic Processes

— There are no solved problems; there are only problems that are more or less solved. —
H. Poincaré

The work of P. A. M Dirac [**21**] was motivated by the fact that the only relativistic generalization of the one-particle Schrödinger equation at that time, the Klein-Gordon equation, was facing serious problems. The Schrödinger equation was obtained from the classical, non-relativistic, energy relation together with the replacement (2.2). The Klein-Gordon equation was instead relativistic, we set $c \equiv 1$. The energy relation now reads

$$E^2 = \mathbf{p}^2 + m^2 \Leftrightarrow p^\mu p_\mu = m^2, \tag{3.1}$$

and in a quantum mechanical approach the free particle equation becomes

$$(\Box + m^2)\psi(x) = 0, \tag{3.2}$$

with $\Box$ being the D'Alembertian operator. The relation (3.1) admits both positive and negative energies and the major problem is that for $E < 0$ the probability density is not positive. How should one interpret a negative probability? At first these solutions were rejected, but they cannot be simply ignored as we then would not have a complete set of states. Dirac's major goal was to overcome these difficulties. He required that the solutions must have a positive definite probability density and that the equation should as well be first order in time. The most general equation that could satisfy these requirements and as well be covariant is

$$H\psi(x) = (\boldsymbol{\alpha} \cdot \mathbf{p} + \beta m)\psi(x), \tag{3.3}$$

and the unknown coefficients β and α_i are determined by the energy relation given by (3.1). It can be rewritten as

$$H^2\psi(x) = (\mathbf{p}^2 + m^2)\psi(x). \tag{3.4}$$

Dirac's suggestions succeeded in overcoming the problems and he also showed that the spin could no longer be treated as a decoupled degree of freedom. It emerges automatically and the predictions were well suited for spin one half particles. Still it is not trivial to see exactly how the spin is to be described relativistically, nor how it is to be interpreted physically.

Of course physicist still had problems with these one-particle theories since they do not preserve causality in general and thus violates a basic postulate in the special theory of relativity [**22**]. The most successful approach today for overcoming this problem is with multiparticle-theories, or *quantum field theories*, sometimes called the "second" quantization of both the Klein-Gordon and Dirac equation. The first being suitable for spinless particles and the second for spin one half particles, for more information see [**23**].

The most successful and verified quantum field theory at present time is quantum electrodynamics, QED, which is denoted by R. P. Feynman as "The Strange Theory of Light and Matter" [**24**]. The theory of the interaction between lepton and quarks is denoted QCD, or quantum chromodynamics, which is still only partially understood. As an example we have the "Proton Spin Crisis" discussed in Chapter 4.

With field theories such as QED and QCD one usually is forced to use (approximate) perturbation analysis, which has both its advantages and disadvantages. It is a very successful approach when one wants to describe scattering processes. But, with bound-state problems the perturbation approach breaks down and we need some other alternatives in order to test the current theories. The proton with its constituent quarks or partons, depending on the energy regime, serves as a excellent example, see Chapter 4. In the physics community there are different opinions how to best overcome these difficulties, where one is to continue to develop mathematical techniques which enables us to formulate a bound-state problem within current field theories or the idea that all particles could be brought to free-particle asymptotic states, because perturbative quantum field theories only handle scattering processes.

A very fruitful and promising approach is a covariant formulation of the theories which brings the behavior of quantum mechanics and special relativity together. If a bound-state model is developed from the two approaches, mentioned earlier, it should have the same space-time symmetry and this incorporates the transformations of the *inhomogeneous Lorentz group*, also know as the *Poincaré group*.

In this chapter we discuss the space-time symmetry group and how to build the representations of the transformations. They can be used to classify particles with the work by E. P. Wigner in 1939 [**1**] where he classified all the unitary irreducible representations of the inhomogeneous Lorentz group. The classification is actually based on the invariant mass and spin of the particle. We also give the theory on how to consider spin effects in relativistic processes and it comes out as a consequence when special relativity and quantum mechanics are merged together.

REMARK 3.0.6. In this chapter we set both $c = 1$ and as well $\hbar = 1$ for simplicity if not stated explicitly otherwise.

3.1. Theory of the Poincaré Group

DEFINITION 3.1.1 ([**18**]). The *Poincaré group* (PG) describes the fundamental space-time symmetries in the four-dimensional Minkowski-space-time and hence is the invariance group of all (special) relativistic theories and thereby determines the general structure to a large extent.

The Poincaré group is a combination of the homogeneous Lorentz transformations and space-time translations on the four-dimensional Minkowski-space-time in the form

$$x^\mu \mapsto \tilde{x}^\mu = \Lambda^\mu{}_\nu x^\mu + a^\mu, \tag{3.5}$$

where $\Lambda^\mu{}_\nu$ is the Lorentz transformations and a^μ the space-time translations.

To achieve a complete understanding of the PG we first study the Lorentz group and then the little group of Wigner, which leaves the four-momentum of a particle invariant. The little group together with the space-time translations a^μ give a complete description of the PG. This allowed Wigner to construct a classification of particles based on their invariant masses and spin. He also made the important conclusion that in a strict sense spin could *not* be treated as a decoupled degree of freedom.

3.1.1. Group of Lorentz Transformations. A basic postulate in special relativity tells us that all inertial frames must be equivalent or all physical processes should be invariant under linear transformations from one to another inertial frame [**22**]. Each frame is a four-dimensional space, with the space-time coordinates as four-vectors

$$x^\mu = (x_0, \mathbf{x}) = (t, x_i) = (t, x, y, z), \tag{3.6}$$

which is the contravariant form. The covariant one takes the form

$$x_\mu = (x_0, -\mathbf{x}) = (t, -x_i) = (t, -x, -y, -z), \tag{3.7}$$

and is related to x^μ by

$$x_\mu = \eta_{\mu\nu} x^\nu. \tag{3.8}$$

The factor $\eta_{\mu\nu}$ is the metric tensor usually called the *Minkowski metric* and its inverse properties are defined as

$$\eta_{\mu\nu}\eta^{\nu\mu} = \delta^\nu_\mu. \tag{3.9}$$

This geometry differs from the common Euclidian geometry in the sense that an *interval* between two events is given by the means of

$$\mathrm{d}s^2 \equiv (x_\mu - y_\mu)(x^\mu - y^\mu) \ = (x_0 - y_0)^2 - (\mathbf{x} - \mathbf{y})^2, \tag{3.10}$$

where the Minkowskian inner product is

$$x^2 = x_\mu x^\mu = x_\mu \eta^{\mu\nu} x_\nu = x_0^2 - \mathbf{x}^2. \tag{3.11}$$

In general a *linear transformation* which operates on this four-dimensional space can be written as

$$x^\mu \mapsto \tilde{x}^\mu = \Lambda^\mu{}_\nu x^\nu = \Lambda^\mu{}_0 x^0 + \Lambda^\mu{}_i x^i, \tag{3.12}$$

and relates two different positions in space-time. This is the required linear transformation between inertial frames in special relativity and we make a formal definition.

DEFINITION 3.1.2 ([**23**]). The transformations (3.12) that leave the length of four-vectors, the *interval* (3.10), invariant are called *Lorentz transformations* (LT's).

LEMMA 3.1.3 ([**23**]). *Any transformation (3.12) that satisfies the relation*

$$\eta_{\lambda\xi} = \eta_{\rho\nu}\Lambda^\rho{}_\lambda \Lambda^\nu{}_\xi \tag{3.13}$$

is a Lorentz transformation.

PROOF. The interval (3.10) must be invariant under a LT

$$\mathrm{d}s^2 = x^2 = x_\mu x^\mu = x_0^2 - \mathbf{x}^2 = y_0^2 - \mathbf{y}^2 = y_\mu y^\mu = \mathrm{d}\tilde{s}^2, \tag{3.14}$$

and we can write

$$\begin{aligned} x^\xi x_\xi &= \tilde{x}^\rho \tilde{x}_\rho \Rightarrow \\ \tilde{x}^\rho \tilde{x}_\rho &= \eta_{\rho\nu}\tilde{x}^\rho \tilde{x}^\nu = \eta_{\rho\nu}\Lambda^\rho{}_\lambda x^\lambda \Lambda^\nu{}_\xi x^\xi = \eta_{\rho\nu}\Lambda^\rho{}_\lambda \Lambda^\nu{}_\xi x^\lambda x^\xi = \\ &\qquad \eta_{\lambda\xi} x^\lambda x^\xi = x^\xi x_\xi, \end{aligned} \tag{3.15}$$

which proves the statement. □

A four-dimensional space-time continuum in which (3.10) is invariant and gives the condition (3.13) is usually called *the Minkowski-space-time.*

REMARK 3.1.4. We denote x_μ as *time-like* if $x^2 > 0$, *light-like* if $x^2 = 0$ and *space-like* for $x^2 < 0$.

A convenient way to understand the LT's is by the use of a matrix notation, where x^μ is a four-dimensional column vector $\mathbf{X}$ and $\eta_{\mu\nu}$ as a 4×4 matrix $\boldsymbol{\eta}$

$$\mathbf{X} = \left[\begin{array}{c} x_0 \\ \hline \mathbf{x} \end{array}\right], \quad \boldsymbol{\eta} = \left[\begin{array}{c|c} 1 & \mathbf{0} \\ \hline \mathbf{0} & -\mathbf{I} \end{array}\right], \tag{3.16}$$

written in block form where the indices run from 0 to 3. A LT is denoted by a 4×4 matrix $\boldsymbol{\Lambda}$

$$\boldsymbol{\Lambda} = \left[\begin{array}{c|c} \Lambda^0{}_0 & \Lambda^0{}_i \\ \hline \Lambda^i{}_0 & \Lambda^i{}_j \end{array}\right], \tag{3.17}$$

and with these notations we can write (3.10) and (3.12) as

$$\mathrm{d}s^2 = \mathbf{X}^\top \boldsymbol{\eta} \mathbf{X}, \qquad \tilde{\mathbf{X}} = \boldsymbol{\Lambda}\mathbf{X}. \tag{3.18}$$

The condition (3.13) given in tensor notation now reads

$$\boldsymbol{\eta} = \boldsymbol{\Lambda}^\top \boldsymbol{\eta} \boldsymbol{\Lambda}. \tag{3.19}$$

DEFINITION 3.1.5 ([**17**]). A *group* is a set G with a binary operator, called the group composition $(\cdot)$, i.e. a map from $G \times G$ to G which is associative and has both a unit element and an inverse.

We now show that the LT's form a group in the mathematical sense, by considering the four group properties:

(i) Assume that $\boldsymbol{\Lambda}_1$ and $\boldsymbol{\Lambda}_2$ are LT's

$$(\boldsymbol{\Lambda}_1\boldsymbol{\Lambda}_2)^\top \boldsymbol{\eta} (\boldsymbol{\Lambda}_1\boldsymbol{\Lambda}_2) = \boldsymbol{\Lambda}_2^\top \boldsymbol{\Lambda}_1^\top \boldsymbol{\eta} \boldsymbol{\Lambda}_1 \boldsymbol{\Lambda}_2 = \boldsymbol{\Lambda}_2^\top \boldsymbol{\eta} \boldsymbol{\Lambda}_2 = \boldsymbol{\eta}. \tag{3.20}$$

which shows that closure is satisfied.

(ii) The associativity is governed by the matrix multiplication

$$(\boldsymbol{\Lambda}_1\boldsymbol{\Lambda}_2)\boldsymbol{\Lambda}_3 = \boldsymbol{\Lambda}_1(\boldsymbol{\Lambda}_2\boldsymbol{\Lambda}_3). \tag{3.21}$$

(iii) The identity exists $\boldsymbol{\Lambda} = \mathbf{I}$ and it is a LT since

$$\mathbf{I}^\top \boldsymbol{\eta} \mathbf{I} = \boldsymbol{\eta}. \tag{3.22}$$

(iv) For any LT there exists an inverse $\boldsymbol{\Lambda}^{-1}$ by

$$\boldsymbol{\eta} = \boldsymbol{\Lambda}^\top \boldsymbol{\eta} \boldsymbol{\Lambda} \Rightarrow \det \boldsymbol{\Lambda} \neq 0, \tag{3.23}$$

and $\boldsymbol{\Lambda}$ has an inverse. The inverse is also a LT since

$$\left(\boldsymbol{\Lambda}^{-1}\right)^\top \boldsymbol{\eta} \boldsymbol{\Lambda}^{-1} = \boldsymbol{\eta}. \tag{3.24}$$

With the condition (3.13) in matrix notation, the considered LT's must satisfy ten conditions on sixteen unknown elements which gives six unknown parameters. These unknown parameters correspond to six different generators of its group structure. The corresponding generators and their representations are of interest and we will return to them later on.

In a more formal manner by the use of mathematical group theory the LT's belong to one of the classical Lie groups, a pseudo-orthogonal group. These different classical groups are based on the work and results of S. Lie on continuous groups of transformations initiated already a century ago, see for example [**15, 16**].

DEFINITION 3.1.6 ([**16**]). The *pseudo-orthogonal group* is defined as

$$O(m-P, P) \equiv \{\mathbf{L} \in GL(m, \mathbb{C}) \mid \mathbf{L} \text{ preserves the Lorentizian metric}\}, \tag{3.25}$$

where $GL(m,\mathbb{C})$ is the *general linear group* and the *Lorentzian metric* on $\mathbb{R}^m$ with $\mathbf{X},\mathbf{Y}\in\mathbb{R}^m$ is given as

$$\mathrm{d}s^2 = \mathbf{X}^\top \boldsymbol{\eta} \mathbf{Y} = \mathbf{X}^\top \begin{bmatrix} \mathbf{I}_P & 0 \\ 0 & -\mathbf{I}_{m-P} \end{bmatrix} \mathbf{Y} = = \sum_{i=1}^{P} x_i y_i - \sum_{i=P+1}^{m} x_i y_i,$$

and $\mathbf{I}_P$ is the $P\times P$ unit matrix. With the condition det $\mathbf{L}=1$ the *special pseudo-special orthogonal group* is defined as

$$SO(m-P,P) \equiv \{\mathbf{L}\in O(m-P,P) \mid \det \mathbf{L}=1\} < O(m-P,P). \tag{3.26}$$

From the previous definition we get the metric of the Minkowski-space-time by setting $P=1$ and $m=4$ and the LT's belong to the pseudo-orthogonal group denoted as $O(3,1)$, called the *homogeneous Lorentz group.* The LT's can be both complex and real. We only focus in the real ones since we map a real space into another real space. But the complex LT's are still very important in the proof of the PCT theorem in quantum field theory, see [**23**] for the proof.

REMARK 3.1.7. Whenever we speak about LT's they are presumed to be real and not complex if not stated specifically.

In the literature, there is a common scheme on how to classify the real LT's to be *proper*, *improper*, *orthochronous* or *non-orthochronous.* We first consider the determinant of the condition (3.13) in matrix notation

$$\det(\boldsymbol{\eta}) = \det\left(\boldsymbol{\Lambda}^\top\boldsymbol{\eta}\boldsymbol{\Lambda}\right) = \det\left(\boldsymbol{\Lambda}^\top\right)\det(\boldsymbol{\eta})\det(\boldsymbol{\Lambda}), \tag{3.27}$$

and since $\det(\boldsymbol{\eta})=1$ we get

$$1 = \det\left(\boldsymbol{\Lambda}^\top\right)\det(\boldsymbol{\Lambda}) = [\det(\boldsymbol{\Lambda})]^2 \Rightarrow \det\boldsymbol{\Lambda} = \pm 1. \tag{3.28}$$

The two different values of the determinant of $\boldsymbol{\Lambda}$ account to the proper/improper $(+/-)$ LT's. Second, by setting $\lambda=\xi=0$ in (3.13) we get

$$1 = \eta_{\rho\nu}\Lambda^\rho{}_0\Lambda^\nu{}_0 = \left[\Lambda^0{}_0\right]^2 - \left[\Lambda^i{}_0\right]^2 \Leftrightarrow \left[\Lambda^0{}_0\right]^2 = 1 + \left[\Lambda^i{}_0\right]^2, \tag{3.29}$$

and thus

$$\left[\Lambda^0{}_0\right]^2 \geq 1 \text{ or } \left|\Lambda^0{}_0\right| \geq 1. \tag{3.30}$$

If $\Lambda^0{}_0 \geq 1$ the transformations are said to be orthochronous, while $\Lambda^0{}_0 \leq -1$ account for non-orthochronous transformations.

DEFINITION 3.1.8 ([**23**]). The real LT's of the group $O(3,1)$ are said to be:

(i) *proper orthochronous* $(\mathrm{L}_+^\uparrow)$ if det $\mathbf{L}=1$, $\Lambda^0{}_0\geq 1$.
(ii) *proper non-orthochronous* $(\mathrm{L}_+^\downarrow)$ if det $\mathbf{L}=1$, $\Lambda^0{}_0\leq -1$.
(iii) *improper orthochronous* $(\mathrm{L}_-^\uparrow)$ if det $\mathbf{L}=-1$, $\Lambda^0{}_0\geq 1$.
(iv) *improper non-orthochronous* $(\mathrm{L}_-^\downarrow)$ if det $\mathbf{L}=-1$, $\Lambda^0{}_0\leq -1$.

The four different classes of the LT's

$$\mathrm{L}_+^\uparrow,\ \mathrm{L}_+^\downarrow,\ \mathrm{L}_-^\uparrow,\ \mathrm{L}_-^\downarrow,$$

are disconnected. The reason is that $\Lambda^0_{\ 0}$ cannot change continuously from 1 to -1 and as well the determinant. But a LT can be continuously deformed into any other LT in the same class and each class can therefore be characterized with one of four basic transformations. These basic matrix transformations are **I**, **T**, **P** and **TP**, where we have the identity, the time inversion, parity and a combination. The time inversion only affects the time ($x_0 \to -x_0$) while parity only the space ($x_i \to -x_i$). However, we can establish a connection between the different classes since the discrete transformations **T**, **P** and **TP** can also be used to map LT's between different classes

$$\mathbf{T}L_+^\uparrow = L_-^\downarrow \tag{3.31a}$$

$$\mathbf{P}L_+^\uparrow = L_-^\uparrow \tag{3.31b}$$

$$\mathbf{TP}L_+^\uparrow = L_+^\downarrow. \tag{3.31c}$$

The discontinuity and the connection between the different classes of real LT's can also be illustrate with Figure 3.1 and we can make an important conclusion, see Summary 3.1.9.

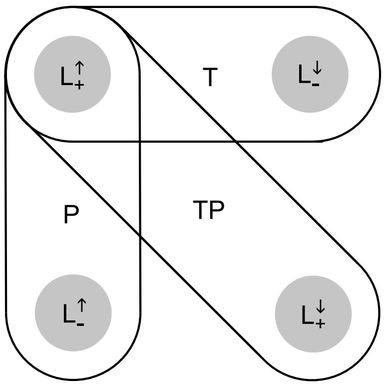

FIGURE 3.1. A schematic picture on the four different classes of real LT's of the homogeneous Lorentz group and the connection between them by the discrete transformations **T**, **P** and **TP**

SUMMARY 3.1.9. To study the properties of the homogeneous Lorentz group $O(3,1)$ of real LT's, it is *sufficient* to study the class $L_+^\uparrow$ and the discrete transformations **T** (time inversion) and **P** (parity).

There is only one of the classes which contains the identity, namely $L_+^\uparrow$, which is required to form a subgroup of the homogeneous Lorentz group $O(3,1)$.

In fact $L_+^\uparrow$ forms an invariant subgroup of $O(3,1)$, called the *proper Lorentz group* [**18**]. From the Summary 3.1.9 we see the importance of the proper Lorentz group and this group was also used by E. P. Wigner in his study of the little group of the Poincaré group which leaves the four-momentum invariant [**1**].

The most convenient way to work with groups and already noticed in the chapter about spin in non-relativistic processes is the use of the generators with their close connection to the group structure and transformations. If we have the structure of the generators we can speak about the group in a very elementary way, in an abstract sense, then try to find representations suitable for different physical processes.

REMARK 3.1.10. We use the word "like" to specify a local homomorphism between different groups. They may have the same Lie algebra structure but different topological global effects as with the case of $SO(3)$ and $SU(2)$. For example, the first being the rotational group of a particle in ordinary space and the second governs the rotation of spin one half particles in the spin space. We speak about that $SU(2)$ is an $O(3)$-like subgroup of $SL(2,\mathbb{C})$, the special linear group of 2×2 complex matrices.

We now proceed to build the corresponding generators and from Chapter 2 together with the results of S. Lie, the real LT's of the proper Lorentz group acting on the space-time coordinates can be written as

$$\mathbf{\Lambda} = e^{-i\epsilon\hat{\mathbf{\Gamma}}/\hbar}, \tag{3.32}$$

where ϵ is a small parameter, $\hat{\mathbf{\Gamma}}$ as a Hermitian generator and $i/\hbar$ conventional. For simplicity we now set $\hbar = 1$. The LT's can be expanded as

$$\mathbf{\Lambda} = \mathbf{I} - i\epsilon\hat{\mathbf{\Gamma}} + \mathcal{O}(\epsilon^2), \tag{3.33}$$

and we insert it into the condition (3.13)

$$\left(\mathbf{I} - i\epsilon\hat{\mathbf{\Gamma}} + \mathcal{O}(\epsilon^2)\right)^\top \boldsymbol{\eta} \left(\mathbf{I} - i\epsilon\hat{\mathbf{\Gamma}} + \mathcal{O}(\epsilon^2)\right) = \boldsymbol{\eta}, \tag{3.34}$$

or

$$\boldsymbol{\eta} - i\epsilon\left(\hat{\mathbf{\Gamma}}^\top \boldsymbol{\eta} + \boldsymbol{\eta}\hat{\mathbf{\Gamma}}\right) + \mathcal{O}(\epsilon^2) = \boldsymbol{\eta}. \tag{3.35}$$

To first order of ϵ we get

$$\hat{\mathbf{\Gamma}}^\top \boldsymbol{\eta} + \boldsymbol{\eta}\hat{\mathbf{\Gamma}} = 0 \Rightarrow \hat{\mathbf{\Gamma}}^\top = -\boldsymbol{\eta}\hat{\mathbf{\Gamma}}\boldsymbol{\eta}, \tag{3.36}$$

and from this equation we can make the following conclusions about the structure of $\hat{\mathbf{\Gamma}}$:

(i) All the diagonal elements of $\hat{\mathbf{\Gamma}}$ are zero since we get the equality

$$\hat{\Gamma}^\mu{}_\mu = -\hat{\Gamma}^\mu{}_\mu \Rightarrow \hat{\Gamma}^\mu{}_\mu = 0.$$

(ii) Three independent elements which are *antisymmetric* and only have *space-space* components.

$$\hat{\Gamma}^i{}_j = -\hat{\Gamma}^j{}_i.$$

(iii) Three independent elements which are *symmetric* and have *space-time* components

$$\hat{\Gamma}^0{}_i = \hat{\Gamma}^i{}_0.$$

We have six unknown parameters and each of them correspond to one generator. These generators span a six-dimensional vector space where the generators with only space-space components are the generators of rotations and the others are the generators of boosts [**23**]. Since we apply the LT's on the four-vector x^μ or the four-dimensional column vector $\mathbf{X}$ we can take the generators to be basis vectors for this space, the generators take the matrix form [**18**]

$$\hat{L}_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{bmatrix}, \quad \hat{K}_1 = \begin{bmatrix} 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{3.37a}$$

$$\hat{L}_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{bmatrix}, \quad \hat{K}_2 = \begin{bmatrix} 0 & 0 & i & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{3.37b}$$

$$\hat{L}_3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \hat{K}_3 = \begin{bmatrix} 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 \end{bmatrix}. \tag{3.37c}$$

The transformations operators for the proper Lorentz group can be exponentiated by (3.32) now as

$$\mathbf{\Lambda} = e^{-i\sum_j(\epsilon_j \hat{L}_j + \varepsilon_j \hat{K}_j)}, \tag{3.38}$$

where ϵ_j and ε_j a small parameters with $\hat{L}_j$ and $\hat{K}_j$ as the generators of rotations and boosts, respectively. These six independent generators form the proper Lorentz group and they satisfy the following commutation relations

$$[\hat{L}_i, \hat{L}_j] = i\epsilon_{ijk}\hat{L}_k \tag{3.39a}$$

$$[\hat{L}_i, \hat{K}_j] = i\epsilon_{ijk}\hat{K}_k \tag{3.39b}$$

$$[\hat{K}_i, \hat{K}_j] = i\epsilon_{ijk}\hat{L}_k, \tag{3.39c}$$

where ϵ_{ijk} is the three-dimensional antisymmetric symbol normalized to unity, called *the Levi-Cività symbol.* These relations are very important to remember as they must always be satisfied whatever representation we may choose in the discussion of the proper Lorentz group. We see from (3.39) that $\hat{L}_i$ form a subgroup of rotations while $\hat{K}_i$ does not form a subgroup at all, and that a multiplication of two boosts will result in a multiplication of a boost and a rotation.

In general, the generators applicable to an arbitrary function of the coordinate variables are [**18**]

$$\hat{L}_i = -i\epsilon_{ijk} x_j \partial^k \tag{3.40a}$$

$$\hat{K}_i = -i\left(x_0 \partial^i + x_i \partial^0\right) \tag{3.40b}$$

and these representations still satisfy (3.39). In literature it is common to combine all six generators in a single form, *videlicet*

$$\hat{L}_{\mu\nu} = -i\left(x_\mu \partial_\nu - x_\nu \partial_\mu\right), \tag{3.41}$$

where $\hat{L}_i = \epsilon_{ijk}\hat{L}_{kj}$ and $\hat{K}_i = L_{0i}$. The Lie algebra, or commutator relation, of $\hat{L}_{\mu\nu}$ is

$$[\hat{L}_{\mu\nu}, \hat{L}_{\rho\sigma}] = i\left(\eta_{\mu\rho}\hat{L}_{\nu\sigma} - \eta_{\nu\rho}\hat{L}_{\mu\sigma} + \eta_{\mu\sigma}\hat{L}_{\rho\nu} - \eta_{\nu\sigma}\hat{L}_{\rho\nu}\right), \tag{3.42}$$

to be identified with the Lie algebra of the proper Lorentz group. It must *always* be satisfied whatever representation one may choose of $\hat{L}_{\mu\nu}$.

In Chapter 2 we noticed that we had different types of generators, one only acting on the coordinates in ordinary space and another in the spin space. We resolved this problem by considering a more general generator and we have the familiar expression $\hat{\mathbf{j}} = \hat{\mathbf{l}} + \hat{\mathbf{s}}$, with $\hat{\mathbf{l}}$ acting on ordinary space and $\hat{\mathbf{s}}$ on the spin space. A *direct product* of the $SO(3)$ and $SU(2)$ groups. This must as well be taken into account in the representations of $\hat{L}_{\mu\nu}$ since (3.41) only transforms the coordinate variables in the Minkowski space. In principle we must add a factor analog to $\hat{\mathbf{s}}$ denoted as $\hat{S}_{\mu\nu}$. This will become more obvious in Section 3.1.3 and after the discussion of the little group connected to the proper Lorentz group.

3.1.2. Wigner's Little Group. In 1939 [**1**] E. Wigner constructed the maximal subgroups of the Lorentz group whose transformations leave the four-momentum p^μ of a given particle invariant. These subgroups are in literature denoted as *Wigner's little group.* One purpose for his work was to construct the representations of the Lorentz group which are relevant for physics. For example in the previous section we used two different representations.

In principle the little group changes the internal space-time coordinates while leaving the four-momentum invariant as we shall see, and one important result of Wigner is that his work placed the spin of physics into the relativistic world. He noticed that the little group for massive particles is locally isomorphic to the three-dimensional rotation group while massless particles are dictated by the two-dimensional Euclidian group. These results together with the group of translations can be combined into the inhomogeneous Lorentz group. By the existence of two Casimir operators, Wigner managed to classify particles in a systematic way. We will return to this matter in Section 3.1.3.

Now let us now first determine the little group and we do not follow Wigners original work as the mathematical techniques available today are much richer and diverse [**18**].

The little group of the proper Lorentz group is determined from the relation

$$\mathbf{\Lambda}^{(p)}\mathbf{P} = \mathbf{P}, \tag{3.43}$$

where $\mathbf{\Lambda}^{(p)}$ is a proper LT and $\mathbf{P}$ is a four-dimensional column vector representing p^μ. According to [**18**] the orbit of p^μ is a surface in the four-dimensional space on which

$$p^2 = p_\mu p^\mu = m^2, \tag{3.44}$$

is constant and for the proper LT's there exists six different orbits, see Figure 3.2 for an illustration of the orbits. The first two orbits are the time-like surfaces for positive m^2 and the third and fourth are the forward and backward light cones. The fifth orbit is a space-like surface for negative m^2 and finally the last orbit is the origin.

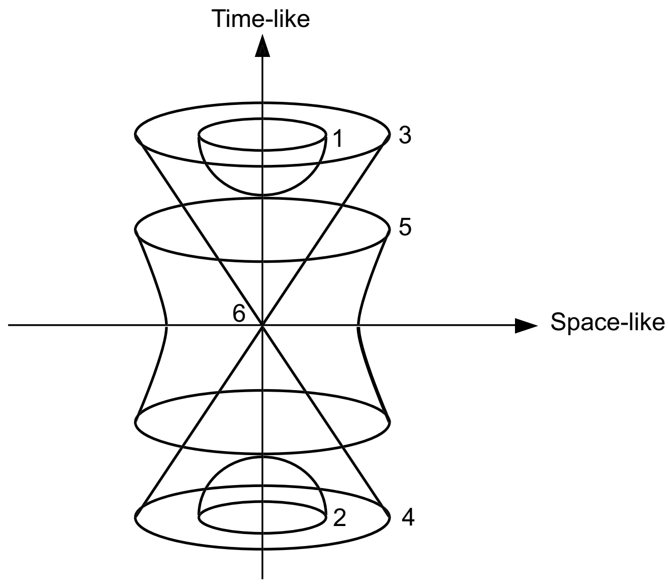

FIGURE 3.2. The orbits of all possible four-momenta of the proper Lorentz group. There are six different orbits and they can be time-like, space-like, light-like or zero. The little groups are given by the different orbits and this schematic view has been adapted from [**18**].

The work on finding all the separately little groups by (3.43) can now be accomplished by considering each orbit separately, as each orbit is specified by a particular four-momentum. The little group for the first and second orbit is governed by the vector $[\pm m, 0, 0, 0]^\top$ which should be invariant. For the third and fourth orbit we need to consider the vector $[\pm w, 0, 0, w]^\top$ and the fifth orbit by $[0, 0, 0, q]^\top$, where both w and q are arbitrary parameters. The last little group leaves the null vector invariant. The relation (3.43) can be written as

$$\underbrace{\begin{bmatrix} l_0^0 & l_1^0 & l_2^0 & l_3^0 \\ l_0^1 & l_1^1 & l_2^1 & l_3^1 \\ l_0^2 & l_1^2 & l_2^2 & l_3^2 \\ l_0^3 & l_1^3 & l_2^3 & l_3^3 \end{bmatrix}}_{\mathbf{\Lambda}^{(p)}} \begin{bmatrix} p_0 \\ p_1 \\ p_2 \\ p_2 \end{bmatrix} = \underbrace{\begin{bmatrix} p_0 \\ p_1 \\ p_2 \\ p_2 \end{bmatrix}}_{\mathbf{P}}, \tag{3.45}$$

and we have to solve linear equations for each p^μ connected to an orbit and as well the condition (3.13) must be taken into account. Before we study any little group in detail we notice that for any non-zero p^μ we get three additional conditions for the parameters l_σ^σ and we know that a proper LT is specified by six independent parameters, so the first to fifth little groups have only three independent parameters. The sixth orbit does not give any additional conditions.

(i) For the first and second orbit, the vector $[\pm m, 0, 0, 0]^\top$ should be invariant under a proper LT. Compared to (3.45) we have $p_0 = \pm m$ and $p_i = 0$. In fact this gives that $l_0^i = 0$ and $l_0^0 = 1$. From condition (3.13) we have that $l_i^0 = 0$ as well. These conclusion can be written explicitly as

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & l_1^1 & l_2^1 & l_3^1 \\ 0 & l_1^2 & l_2^2 & l_3^2 \\ 0 & l_1^3 & l_2^3 & l_3^3 \end{bmatrix} \begin{bmatrix} \pm m \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \pm m \\ 0 \\ 0 \\ 0 \end{bmatrix}, \tag{3.46}$$

and this transformation matrix is a reduced from of a proper LT. It is identified with the rotation subgroup of the proper Lorentz group spanned by the Lie algebra (3.39a).

SUMMARY 3.1.11. The little group for the first and second orbit, time-like surfaces in the four-vector spaces, is generated by the Lie algebra (3.39a), *videlicet*

$$[\hat{L}_i, \hat{L}_j] = i\epsilon_{ijk}\hat{L}_k.$$

(ii) For the third and fourth orbit, the vector $[\pm w, 0, 0, w]^\top$ must be invariant under a proper LT. In this case we must consider each vector separately. We first study the vector $[w, 0, 0, w]^\top$ and by similar conclusions from the treatment of the first and second orbits, we get

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & l_1^1 & l_2^1 & 0 \\ 0 & l_1^2 & l_2^2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} w \\ 0 \\ 0 \\ w \end{bmatrix} = \begin{bmatrix} w \\ 0 \\ 0 \\ w \end{bmatrix}, \tag{3.47}$$

which leaves the vector invariant. This transformation matrix is identified with a rotation around the z-axis and its generator is $\hat{L}_3$. However, in addition there exists another matrix as well which leaves $[w, 0, 0, w]^\top$ invariant [**1**]

$$\underbrace{\begin{bmatrix} 1 + (u^2 + v^2)/2 & u & v & -(u^2 + v^2)/2 \\ u & 1 & 0 & -u \\ v & 0 & 1 & -v \\ (u^2 + v^2)/2 & u & v & 1 - (u^2 + v^2)/2 \end{bmatrix}}_{\equiv \mathbf{T}(u,v)} \begin{bmatrix} w \\ 0 \\ 0 \\ w \end{bmatrix} = \begin{bmatrix} w \\ 0 \\ 0 \\ w \end{bmatrix}, \tag{3.48}$$

and it has the following algebraic properties

$$\mathbf{T}(u, v) = \mathbf{T}(u, 0)\mathbf{T}(0, v) = \mathbf{T}(0, v)\mathbf{T}(u, 0) \tag{3.49a}$$

$$\mathbf{T}(u_1, v_1)\mathbf{T}(u_2, v_2) = \mathbf{T}(u_1 + u_2, v_1 + v_2). \tag{3.49b}$$

These properties are identical with the structure of translations in the Euclidian xy-plane, more precisely it is a subgroup of the special Euclidian group $SE(2)$, for a formal definition see [**16**]. From [**18**] we have the infinitesimal limits of $T(u, 0)$ (translation in x) and $T(0, v)$ (translation in y) as

$$\hat{N}_1 = \begin{bmatrix} 0 & i & 0 & 0 \\ i & 0 & 0 & -i \\ 0 & 0 & 0 & 0 \\ 0 & i & 0 & 0 \end{bmatrix}, \quad \hat{N}_2 = \begin{bmatrix} 0 & 0 & i & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{bmatrix}. \tag{3.50}$$

The $SE(2)$ group contains as well rotations in the xy-plane and this can actually be accomplished with the matrix connected to the generator $\hat{L}_3$. Indeed, the little group connected to the orbit of the vector $[w, 0, 0, w]^\top$ is isomorphic to $SE(2)$. The calculations for the fourth orbit and vector $[-w, 0, 0, w]^\top$ are similar, since we get the same rotation matrix around the z-axis but the $\mathbf{T}(u, v)$ in this case is the Hermitian conjugate of the matrix $\mathbf{T}(u, v)$ in (3.48). The little group for this orbit or vector $[-w, 0, 0, w]^\top$ is as well isomorphic to $SE(2)$.

SUMMARY 3.1.12. The little group of the third orbit, the forward light cone, is generated by $\hat{L}_3$, $\hat{N}_1$ and $\hat{N}_2$ with the commutator relations

$$[\hat{N}_1, \hat{N}_2] = 0 \tag{3.51a}$$
$$[\hat{N}_1, \hat{L}_3] = -i\hat{N}_2 \tag{3.51b}$$
$$[\hat{N}_2, \hat{N}_3] = i\hat{N}_1, \tag{3.51c}$$

which are identical with the structure of the $SE(2)$ group. The little group of the fourth orbit, the backward light cone, is essential identical to the third orbit but $\hat{N}_1$ and $\hat{N}_2$ are replaced by their Hermitian conjugates. The little group of the third and fourth orbit is an $E(2)$-like group.

(iii) The vector $[0, 0, 0, q]^\top$ must be invariant for the fifth orbit. From the calculations of the third orbit we have that a rotation around the z-axis satisfied the invariant requirement. One of the generators of this little group is therefore $\hat{L}_3$. The matrices of boosts along the x and y axes leaves the vector $[0, 0, 0, q]^\top$ as well invariant and the corresponding generators are $\hat{K}_1$ and $\hat{K}_2$. We have three independent parameters and they operate on two spatial dimensions and one time-like dimension. We have the special pseudo-orthogonal group $SO(2, 1)$.

SUMMARY 3.1.13. The little group of the fifth orbit, space-like surfaces, is generated by the Lie algebra

$$[\hat{K}_1, \hat{K}_2] = -i\hat{L}_3 \tag{3.52a}$$
$$[\hat{K}_1, \hat{L}_3] = -i\hat{K}_2 \tag{3.52b}$$
$$[\hat{K}_2, \hat{L}_3] = i\hat{K}_1 \tag{3.52c}$$

to be identified with the group $SO(2, 1)$.

(iv) The sixth orbit is only the origin and therefore all the transformation matrices in the proper Lorentz group leave this point invariant.

SUMMARY 3.1.14. The little group of the sixth orbit, the origin, is generated by the Lie algebra (3.42), *videlicet*

$$[\hat{L}_{\mu\nu}, \hat{L}_{\rho\sigma}] = i\left(\eta_{\mu\rho}\hat{L}_{\nu\sigma} - \eta_{\nu\rho}\hat{L}_{\mu\sigma} + \eta_{\mu\sigma}\hat{L}_{\rho\nu} - \eta_{\nu\sigma}\hat{L}_{\rho\nu}\right),$$

which is the proper Lorentz group.

In total, the little group of Wigner is of four different types. The first type is a rotation subgroup of the proper Lorentz group, an $O(3)$-like group. The second type is an $E(2)$-like group, it consists of translations and rotations on a flat surface. The third type is a the $SO(2, 1)$ group an finally the complete proper Lorentz group sometimes denoted as $SO(3, 1)$.

We have found the little group of Wigner which leaves the four-momentum of a given free-particle invariant. In the next section we incorporate the little group into the PG and from the representations we discuss the physical effects. The most important conclusion is that a massive particle must in general have spin. The spin is predicted form the theory not by experiments as with the case of non-relativistic processes in Chapter 2.

3.1.3. Representations and Classifications. The transformation symmetries of the PG are determined by the space-time translations (3.5)

$$x^\mu \mapsto \tilde{x}^\mu = \Lambda^\mu{}_\nu x^\mu + a^\mu.$$

In Section 3.1.1 we discussed the properties of the homogeneous Lorentz transformations $\Lambda^\mu{}_\nu$ and its group structure. However, we need to consider the translations a^μ as well and from [**18**] we have that the translations separately form an Abelian group. The Hermitian generators are

$$\hat{P}_\mu = -i\partial_\mu, \tag{3.53}$$

applicable to arbitrary functions.

DEFINITION 3.1.15 ([**16**]). A group is said to be *affine* on $\mathbb{R}^m$ if it is a combination of (pseudo-) orthogonal and translations transformations on $\mathbb{R}^m$ with a Lorentzian or Euclidian metric.

With the previous definition, the PG is an affine group and a semidirect product of the Lorentz and translation group in the four-dimensional Minkowski space. Indeed, two PG transformations with the parameters (Λ_1, a_1) and (Λ_2, a_2) maps a space-time point as

$$x^\mu \mapsto \Lambda^\mu_{1\nu} x^\nu + a^\mu_1 \mapsto \Lambda^\mu_{2\rho}\Lambda^\rho_{1\nu} x^\nu + \Lambda^\mu_{2\rho} a^\rho_1 + a^\mu_2, \tag{3.54}$$

which demonstrates the composition law of the PG. Notable is that the translation parameters get rotated.

SUMMARY 3.1.16. The complete group structure of the PG is governed by the Lie algebra

$$[\hat{P}_\mu, \hat{P}_\nu] = 0 \tag{3.55a}$$

$$[\hat{M}_{\mu\nu}, \hat{P}_\rho] = i\left(\eta_{\mu\rho}\hat{P}_\nu - \eta_{\nu\rho}\hat{P}_\mu\right) \tag{3.55b}$$

$$[\hat{M}_{\mu\nu}, \hat{M}_{\rho\sigma}] = i\left(\eta_{\mu\rho}\hat{M}_{\nu\sigma} - \eta_{\nu\rho}\hat{M}_{\mu\sigma} + \eta_{\mu\sigma}\hat{M}_{\rho\nu} - \eta_{\nu\sigma}\hat{M}_{\rho\nu}\right) \tag{3.55c}$$

where $\hat{M}_{\mu\nu}$ is the most general representation of $\hat{L}_{\mu\nu}$. In total there are 10 generators. Three are the momentum operators $\hat{P}_i$ and a fourth is the Hamiltonian operator $\hat{P}_0$, all of them connected to the translations is space-time. The six others are the generators $\hat{M}_{\mu\nu}$ of the homogeneous Lorentz group.

REMARK 3.1.17. The reason why we use the term "most general representation of $\hat{L}_{\mu\nu}$" depend on that we can write $\hat{M}_{\mu\nu}$ as a direct product of $\hat{L}_{\mu\nu}$ and $\hat{S}_{\mu\nu}$ without any loss of generality of the Lie algebra (3.42). $\hat{L}_{\mu\nu}$ is only applicable to the coordinate variables, while $\hat{S}_{\mu\nu}$ to all other variables not affected by $\hat{L}_{\mu\nu}$.

Finally we have a way to give Wigners classifications of all the unitary irreducible representations of the PG, worked out by him in 1939 [**1**] and as stated earlier we do not follow his method exactly.

Instead, we find the set of operators whose eigenvalues label these irreducible representations. The operators needed are the Casimir operators as they commute with all the generators of the PG and we immediately see the analog to Chapter 2, where we used the operator $\hat{\mathbf{j}}^2$ to label the irreducible representations for non-relativistic processes. The Casimir operators of the PG [**18**] are

$$\hat{P}_\sigma \hat{P}^\sigma \tag{3.56a}$$

$$\hat{W}_\sigma \hat{W}^\sigma, \tag{3.56b}$$

where $\hat{W}_\sigma$ is the set of Pauli-Lubanski operators

$$\hat{W}_\sigma = -\frac{1}{2}\epsilon_{\mu\nu\rho\sigma}\hat{M}^{\mu\nu}\hat{P}^\rho. \tag{3.57}$$

$\hat{W}_\sigma$ transforms like a four-vector but with the constraint

$$\hat{P}_\sigma \hat{W}^\sigma = 0, \tag{3.58}$$

and therefore it only has three independent parameters. If a system is in a free-particle state, the operator $\hat{P}^2$ determines the squared invariant mass of the particle. Since $\hat{W}_\sigma$ commutes with $\hat{P}_\sigma$ we can replace $\hat{P}^\sigma$ with the eigenvalue p^σ in (3.56a) and (3.57). As we shall see, the three independent components of $\hat{W}_\sigma$ are actually the generators of the little groups at p^σ.

(i) The first and second orbit was connected to $p^\mu = (\pm m, \mathbf{0})$ and this can represent a a massive particle in its rest frame. The generators $\hat{W}_\sigma$ take the form

$$\hat{W}_0 = 0 \tag{3.59a}$$

$$\hat{W}_i = \pm(m/2)\epsilon_{ijk}\hat{M}_{jk}, \tag{3.59b}$$

and the space part of $\hat{W}_\sigma$ has the commutator relation

$$[\hat{W}_k, \hat{W}_l] = -iM\epsilon_{kln}\hat{W}_n, \quad M \equiv \pm m, \tag{3.60}$$

If $M \neq 0$ we set

$$\hat{s}_i = \frac{1}{M}\hat{W}^i, \tag{3.61}$$

and get the Lie algebra

$$[\hat{s}_j, \hat{s}_k] = i\epsilon_{jkl}\hat{s}_l, \tag{3.62}$$

which clearly shows that the three $\hat{W}_i$ generate the $SO(3)$-like little group. The $\hat{L}_{ij}$ part of $\hat{M}_{ij}$ is dropped as we have a particle at rest and therefore it has no effect on the system. We have a similar situation as with the non-relativistic approach, a kind of internal angular momentum represented by the $\hat{S}_{ij}$ part. *Thus, this little group changes the internal space-time variables of a particle at rest and it is the spin of the particle. Wigner managed to place the spin of physics into the relativistic world.*

(ii) The third and fourth orbit are connected to $p^\mu = (\pm w, 0, 0, w)$. However, it is not possible to bring a massless particle to its rest frame. For such a particle one can fix the coordinate system in which the particle momentum is along the z-axis. Indeed p^μ can therefore describe such a massless particle. The generators $\hat{W}_\sigma$ take the form

$$\hat{W}_0 = w\hat{S}^{12} \tag{3.63a}$$

$$\hat{W}_1 = w\left(\hat{S}^{02} - \hat{S}^{23}\right) \tag{3.63b}$$

$$\hat{W}_2 = w\left(\hat{S}^{01} - \hat{S}^{31}\right) \tag{3.63c}$$

$$\hat{W}_3 = \hat{W}_0, \tag{3.63d}$$

and we see that $\hat{W}_0$ is redundant. The $\hat{L}_{ij}$ parts of $\hat{M}_{ij}$ are dropped out similar to the previous case. These generators are proportional to the generators of the $SE(2)$ group. *The little group is a $SE(2)$-like group and is essential for the structure of massless particles.* i.e. the photon serves as an excellent example. Wigner did not give a complete physical interpretation of this group and it was not until 1971 [**25**] that it was observed that they were actually gauge transformations.

(iii) The fifth orbit has $p^\mu = (0, 0, 0, q)$ as the four-momentum vector and for this case the $\hat{W}_\sigma$ take the form

$$\hat{W}_0 = q\hat{S}_{12} \tag{3.64a}$$

$$\hat{W}_1 = q\hat{S}_{02} \tag{3.64b}$$

$$\hat{W}_2 = q\hat{s}_{10} \tag{3.64c}$$

$$\hat{W}_3 = 0. \tag{3.64d}$$

They are proportional ($\hat{S}_3 \sim S_{12}$, $\hat{K}_1 \sim S_{02}$, $\hat{K}_2 \sim S_{02}$) to the structure of the $SO(2,1)$ group and the the nonzero $\hat{W}_\sigma$ form the little group of this orbit. This little group is not yet relevant to physical particles as it corresponds to the structure of particles moving faster than the speed of light which have not been experimental observed at present time.

$\hat{P}^2$	Subgroup of $O(3,1)$	Subgroup of $SL(2,\mathbb{C})$
Massive: $m^2 > 0$	$O(3)$-like subgroup of $O(3,1)$ (*Hadrons*)	$SU(2)$-like subgroup of $SL(2,\mathbb{C})$ (*Electrons*)
Massless: $m^2 = 0$	$E(2)$-like subgroup of $O(3,1)$ (*Photons*)	$E(2)$-like subgroup of $SL(2,\mathbb{C})$ (*Neutrinos?*)
Spacelike: $m^2 < 0$	$O(2,1)$-like subgroup of $O(3,1)$	$Sp(2)$-like subgroup of $SL(2,\mathbb{C})$
$p_\mu = 0$	$O(3,1)$	$SL(2,\mathbb{C})$

TABLE 1. This is a table of the little groups of Wigner. He showed that massive particles have a rotation structure of an $O(3)$-like group and for massless an $E(2)$-like group. The internal space-time symmetries of massive particles is a $SU(2)$-like group while massless have the structure of an $E(2)$-like group. These conclusions can be used to describe particles such as the hadrons, electrons and photons. The neutrino was first believed to be massless but today there are experimental evidence showing the opposite. The $O(2,1)$- and $Sp(2)$-like subgroups are not physically relevant as they correspond to particles moving faster than light, but they can still play an important role. This table has been adapted from [**18**].

(iv) The sixth orbit is trivial as there is no particle at all since we have $p^\mu = 0$ and there exists no generators of $\hat{W}_\sigma$, in principle we already know this little group as it is the complete group of the homogeneous Lorentz group.

The conclusion so far is that the little group can be used to classify particles and their structures in a schematic way. It is based on the squared invariant mass and the internal structure of the particles, known as the spin. We make a summary with Table 1.

In principle, Wigner's little group can serve as the starting point for the study and construction of the representations of the PG. The PG is a ten parameter Lie group and it is a semi-direct product of Lorentz and space-time translations. We only focus in the physical relevant representations which was one of Wigner's idea. It led him to draw the important conclusion that the little group which leaves the four-momentum of a particle invariant describes the internal space-time symmetries and that they are different for massive and massless particles.

The method of constructing the representations of the PG first starts with the little group and its representations for the four-momentum of a given particle. We need to study in detail these representations as the four-momentum traces its orbit, called an orbit completion. Indeed, the little group is a subgroup of the Lorentz group and we have constructed the representations of the Lorentz group by starting from its subgroups. This is called the *induced representations*, and the same approach is used for the PG. In a mathematical point of view we can study the PG by the construction of induced representations [**18**]. We leave it at this stage as it is not the main theme in this thesis to study the representations in detail, we recommend [**18**] for further discussions. We have the most important conclusion that the spin is an inherent property of special relativity together with quantum mechanics. We proceed in the next section on how to consider the spin for relativistic processes.

3.2. The Basic Spin Formalism

The properties of the PG and Wigner's little group show that a covariant theory already contains the spin by the existence of a covariant spin operator known as the set of Pauli-Lubanski operators (3.57)

$$\hat{W}_\mu = -\frac{1}{2}\epsilon_{\mu\nu\rho\sigma}\hat{M}^{\mu\nu}\hat{P}^\rho.$$

It is not the most convenient operator as with the case of spin in the non-relativistic case which has a much simpler structure at first sight

$$[\hat{s}_j, \hat{s}_k] = i\epsilon_{jkl}\hat{s}_l, \tag{3.65}$$

In the discussion of the little group for a massive particle at rest, $p^\mu = (m, \mathbf{0})$, the only nonzero components of $\hat{W}_\mu$ were the space parts and they satisfied the commutator relations (3.60)

$$[\hat{W}_k, \hat{W}_l] = -im\epsilon_{kln}\hat{W}_n.$$

If we make a substitution as $\hat{s}_i = \hat{W}^i/m$ we may actually identify the Lie algebra (3.65). The physical meaning of the $\hat{W}_\mu$ operator can be even more clarified as follows. We know that $\hat{W}_\mu\hat{W}^\mu$ is a Casimir operator of the PG and if we operate it on a particle state with zero momentum, the rest frame, we should be able to identify the spin quantum number compatible to the non-relativistic case as $\hat{W}_i$ are proportional to the structure (3.65). We show it in the following example.

EXAMPLE 3.2.1. Assume we have a massive particle with mass m in its rest frame. Then we apply the operator $\hat{W}^2$ on this state as follows

$$\hat{W}_\mu\hat{W}^\mu|\Omega\rangle = \eta^{\mu\nu}\hat{W}_\mu\hat{W}_\nu|\Omega\rangle = \left(\hat{W}_0^2 - \hat{W}_i^2\right)|\Omega\rangle, \tag{3.66}$$

where the state $|\Omega\rangle$ is equivalent to $|m, \mathring{\mathbf{p}}\rangle$ for simplicity and $\mathring{\mathbf{p}} \equiv (m, 0, 0, 0)$. From Section 3.1.3 we have

$$\hat{W}_0 = 0, \quad \hat{W}_i = (m/2)\epsilon_{ijk}\hat{M}_{jk},$$

and that the $\hat{L}_{jk}$ parts can be dropped. Now let us investigate one of the components of $\hat{W}_i$

$$\hat{W}_1 = (m/2)\epsilon_{1jk}\hat{M}_{jk} = (m/2)\left[\epsilon_{123}\hat{M}_{23} + \epsilon_{132}\hat{M}_{32}\right] = (m/2)\left[\hat{M}_{23} - \hat{M}_{32}\right] = m\hat{M}_{23} = m\hat{S}_{23} = m\hat{S}_1, \tag{3.67}$$

and by similar calculations we also have

$$\hat{W}_2 = m\hat{M}_{31} = m\hat{S}_2, \quad \hat{W}_3 = m\hat{M}_{12} = m\hat{S}_3. \tag{3.68}$$

If we put our results into (3.66) we have

$$\hat{W}_\mu\hat{W}^\mu|\Omega\rangle = m^2(\hat{S}_1^2 + \hat{S}_2^2 + \hat{S}_3^2)|\Omega\rangle = m^2\hat{\mathbf{S}}^2|\Omega\rangle = m^2 s(s+1)|\Omega\rangle,$$

or explicitly

$$\frac{1}{m^2}\hat{W}^2|m, \mathring{\mathbf{p}}\rangle = s(s+1)|m, \mathring{\mathbf{p}}\rangle. \tag{3.69}$$

The operator $\hat{W}_\mu\hat{W}^\mu$ is invariant, as it is a Casimir operator of the PG, and its eigenvalues are in the form of $m^2 s(s+1)$ where s is the same spin quantum number obtained in Chapter 2.

This example shows that in a relativistic theory a particle is assigned an invariant spin quantum number s and it emerges as an inherent property of the theory. We have an opposite situation as with the non-relativistic case where we modified the theory by considering the spin as a decoupled degree of freedom. On a more fundamental basis the spin has the same fundamental importance as the invariant mass. These two properties are used to get the hierarchy of known particles, denoted the Standard Model, and again we recognize the influence of Wigner's work.

However, we may *only* identify the spin quantum number for a particle at rest and the labelling of a particle with arbitrary momentum is not so clear cut. For the rest state we can invoke the usual formalism of non-relativistic quantum mechanics discussed in Chapter 2, i.e. we have the eigenstates $|sm_s\rangle$ of the operators $\hat{\mathbf{s}}^2$ and $\hat{s}_z$. The eigenvalues are $s(s+1)$ and m_s, respectively.

The most convenient approach to deal with particles of arbitrary momentum is by the use of suitable LT's. These transformations can generate states of arbitrary momentum upon the rest states. We adopt a similar approach as with rotations in Chapter 2, but now we instead have LT's.

REMARK 3.2.2. For simplicity we use the notations in [**4**]. We denote l as an *arbitrary LT* and l_i as *physical pure Lorentz boosts* along an axis. We still continue to denote r as a *physical rotation operator* but of the proper Lorentz group.

EXAMPLE 3.2.3. For example, the rotation sequence of a reference system S by a rotation around the y-axis through some angle (β), a boost along the new z'-axis with speed v and finally a rotation around the new y''-axis by the angle $(-\beta)$ is given as

$$S \mapsto S''' = r_{\hat{\mathbf{e}}''_{(2)}}(-\beta) l_{\hat{\mathbf{e}}'_{(3)}}(v) r_{\hat{\mathbf{e}}_{(2)}}(\beta) S. \tag{3.70}$$

As usually with the LT's the sequence of different operators is important and this particular sequence is referred in literature as a *canonical boost.* We illustrate this sequence of transformations with Figure 3.3.

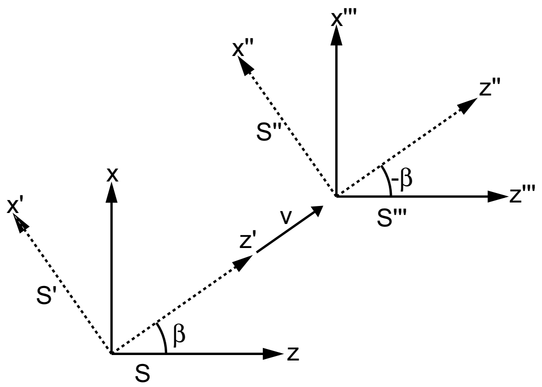

FIGURE 3.3. This is an illustration of a canonical boost along the vector $\mathbf{v}$, which maps a frame S into a new frame S'''. It is a sequence of three LT's, two of them are rotations around the y-axes and one as a boost along the vector $\mathbf{v}$.

An arbitrary LT on a point in space-time can now explicitly be written as

$$x^\mu \mapsto \tilde{x}^\mu = \Lambda^\mu_{\ \nu}(l) x^\nu. \tag{3.71}$$

compatible to Definition 3.1.2. We are particular interested in the connections between different frames. If we have a reference frame S, an arbitrary LT would map all points to a new frame denoted as S^l. In analog to Section 2.3.1, a given four-vector x^μ in S will appear in S^l as having the components

$$(x^\mu)_{S^l} = \Lambda^\mu_{\ \nu}(l^{-1}) x^\nu. \tag{3.72}$$

From Section 2.3.2 and 2.3.4, we were occupied with finding the unitary representations of different Hermitian operators. Especially, the operator of rotations was used to understand the physical properties of spin states. With the LT's we have a similar situation where the generators of rotations and boost are the key elements. However, we do not discuss these representations in detail as it is not really necessary, for further information on this matter see for example [**4**]. We instead use a more abstract approach and in the following sections we discuss the physical outcomes when having either massive or massless particles with their spin properties.

3.2.1. Massive Particles. For a massive particle at rest we saw how we were able to reproduce a notion of spin as an intrinsic property of the particle and could label the spin states similar to the non-relativistic case of angular momentum. The general approach on particles in motion was also deduced in the previous section.

Assume that we have a reference frame S° in which a particle with mass m is at a rest, the spin state is described as $|m, \mathring{\mathbf{p}}, sm_s\rangle$ by the observer O°. Now let O be another observer in a frame S, moving with a velocity $-\mathbf{v}$ with respect to the observer O° in S°. We can in general choose

$$\mathbf{v} = \frac{\mathbf{p}}{\sqrt{\mathbf{p} + m^2}}, \tag{3.73}$$

where $\mathbf{p}$ is some arbitrary momentum, and according to the observer O the state $|m, \mathring{\mathbf{p}}, sm_s\rangle$ becomes explicitly depended on $\mathbf{p}$. As we have the possibility to have arbitrary velocity there are infinitely many frames S attached to the rest frame S°. Each such frame still labels the rest state by the momentum $\mathbf{p}$ but since they are rotated from each other they will report different spin states. In summary we have that the observer O reports *a spin state dependent on which reference frame is used.* We must therefore have a kind of standard convention and in the literature there are two different main choices.

DEFINITION 3.2.4 ([**4**]). The *canonical choice* is given as

$$S^\circ \mapsto S = r^{-1}(-\mathbf{v}) l_{\hat{\mathbf{e}}'_{(3)}}(-\|\mathbf{v}\|) r(-\mathbf{v}) S^\circ, \tag{3.74}$$

where $r(-\mathbf{v})$ is a rotation to $\mathbf{v}$ and this vector is given by (3.73). In this frame, the rest state of a massive particle $|m, \mathring{\mathbf{p}}, sm_s\rangle$ becomes $|m, \mathbf{p}, m'_s\rangle$.

DEFINITION 3.2.5 ([**4**]). The *helicity choice* is given as

$$S^\circ \mapsto S = r_{\hat{\mathbf{e}}''_{(3)}}(-\beta) r_{\hat{\mathbf{e}}'_{(2)}}(-\alpha) l^{-1}_{\hat{\mathbf{e}}_{(3)}}(\|\mathbf{v}\|) S^\circ, \tag{3.75}$$

where we let $\mathbf{p}$ have polar angles (α, β) and $\mathbf{v}$ is given by (3.73). This rotation makes the momentum appear as $\mathbf{p} = (p, \alpha, \beta)$ and if we use the Euler angles the transformation (3.75) becomes

$$S^\circ \mapsto S = r^{-1}(\alpha, \beta, 0) l^{-1}_{\hat{\mathbf{e}}_{(3)}}(\|\mathbf{v}\|) S^\circ. \tag{3.76}$$

In this frame, the rest state of a massive particle $|m, \mathring{\mathbf{p}}, s, \lambda\rangle$, with $m_s = \lambda$, becomes $|m, \mathbf{p}, \lambda\rangle$ denoted as the *helicity state.*

The most convenient choice is the helicity one and we seldom use the canonical one. One reason for this is that the formalism with the helicity is much simpler and it can be used for both massive and massless particles as well shall see.

The helicity state is defined by the Definition 3.2.5 as

$$|m, \mathbf{p}, \lambda\rangle \equiv |m, \overset{\circ}{\mathbf{p}}, s, \lambda\rangle_S,$$

where now S is the helicity frame of the rest frame S°. The relation is governed by the transformation (3.76). If we have a particle in a frame S with helicity λ, the same helicity state would be observed in S°, i.e. the relation between the frames would be

$$S^\circ = l_{\hat{\mathbf{e}}'_{(3)}}(\|\mathbf{v}\|) r(\alpha, \beta, 0) S. \tag{3.77}$$

DEFINITION 3.2.6 ([**4**]). The particular rest frame given by (3.77) is denoted *the helicity rest frame* and the transformation in (3.77) is often given the notation

$$h(\mathbf{p}) \equiv l_{\hat{\mathbf{e}}'_{(3)}}(\|\mathbf{v}\|) r(\alpha, \beta, 0). \tag{3.78}$$

As with the case of the Euler angles of rotations in ordinary space discussed in Section 2.3.1 it is not very convenient to have a mixing of both "new" and "old" axes in $h(\mathbf{p})$. It can be shown from [**26**] that

$$h(\mathbf{p}) = r(\alpha, \beta, 0) l_{\hat{\mathbf{e}}_{(3)}}(\|\mathbf{v}\|), \tag{3.79}$$

and it only involves one reference frame.

EXAMPLE 3.2.7. The matrix form of the LT $\Lambda^\mu{}_\nu(h(\mathbf{p}))$ [**4**] is

$$\mathbf{\Lambda}(h(\mathbf{p})) = \begin{bmatrix} \gamma & 0 & 0 & \beta\gamma \\ \gamma\beta_x & \cos\theta\cos\phi & -\sin\phi & \gamma\beta_x/\beta \\ \gamma\beta_y & \cos\theta\sin\phi & \cos\phi & \gamma\beta_y/\beta \\ \gamma\beta_z & -\sin\theta & 0 & \gamma\beta_z/\beta \end{bmatrix}, \tag{3.80}$$

with polar angles of $\boldsymbol{\beta} = \mathbf{p}/E$ and $\gamma = E/m$. If we operate this transformation on a particle in a rest frame we have

$$\begin{aligned} \mathbf{\Lambda}(h(\mathbf{p}))[m, \overset{\circ}{\mathbf{p}}]^\top &= \begin{bmatrix} \gamma & 0 & 0 & \beta\gamma \\ \gamma\beta_x & \cos\theta\cos\phi & -\sin\phi & \gamma\beta_x/\beta \\ \gamma\beta_y & \cos\theta\sin\phi & \cos\phi & \gamma\beta_y/\beta \\ \gamma\beta_z & -\sin\theta & 0 & \gamma\beta_z/\beta \end{bmatrix} \begin{bmatrix} m \\ 0 \\ 0 \\ 0 \end{bmatrix} = \\ & \begin{bmatrix} \gamma m \\ \gamma\beta_x m \\ \gamma\beta_y m \\ \gamma\beta_z m \end{bmatrix} = \begin{bmatrix} E \\ \mathbf{p} \end{bmatrix} = [E, \mathbf{p}]^\top, \end{aligned} \tag{3.81}$$

and thus, $h(\mathbf{p})$ turns the four-vector $\overset{\circ}{p}{}^\mu$ into p^μ as expected.

The relation (3.77) can now be written as

$$S = h^{-1}(\mathbf{p})S^{\circ}, \tag{3.82}$$

and in a complete analogy to the rotation of spin states in Section 2.4.1 and relation (2.147) we have

$$|m, \mathbf{p}, \lambda\rangle \equiv |m, \overset{\circ}{\mathbf{p}}, s, \lambda\rangle_S = U\left(h(\mathbf{p})\right)||m, \overset{\circ}{\mathbf{p}}, s, \lambda\rangle, \tag{3.83}$$

where U is the corresponding unitary operator. We have finally a way to label the spin states of a massive particle between different observers and before we proceed to the massless case we discuss in more detail the physical meaning of helicity with respect to a massive particle.

The quantity helicity has actually a very simple meaning in an arbitrary frame as we shall see.

DEFINITION 3.2.8 ([**4**]). The helicity operator is defined as

$$\frac{\hat{\mathbf{J}} \cdot \hat{\mathbf{P}}}{\|\hat{\mathbf{P}}\|}, \tag{3.84}$$

where $\hat{J}_i \equiv -(1/2)\epsilon_{ijk}\hat{M}^{jk}$ to be identified with the total angular momentum operators and $\hat{P}_i$ are the momentum operators. It is a *projection operator* of $\hat{\mathbf{J}}$ onto the direction of $\hat{\mathbf{P}}$.

Under this definition the helicity states are the eigenstates of the helicity operators [**27**], i.e.

$$\frac{\hat{\mathbf{J}} \cdot \hat{\mathbf{P}}}{\|\hat{\mathbf{P}}\|}|m, \mathbf{p}, \lambda\rangle = \lambda|m, \mathbf{p}, \lambda\rangle. \tag{3.85}$$

Thus, the helicity λ, the eigenvalues of the helicity operators, is physically *the projections of the total angular momentum onto the direction of the linear momentum of a massive particle.*

REMARK 3.2.9. The total angular momentum refers to both $\hat{L}_i$ and $\hat{S}_i$ as discussed in Section 3.1.2.

By the usual formalism of the non-relativistic approach to spin we can identify the helicity for a massive particle at rest by the eigenvalues of the spin operator $\hat{s}_z$. The helicity is given as $\lambda = m_s$ and the quantization is chosen along a well-defined z-axis where m_s takes all the values $(-s, -s+1, ..., s-1, s)$. Therefore, the eigenvalues λ of the helicity operator are in the same interval but with the reference axis of the momentum. We have a similarity to the discussion of spin one half states in Section 2.4.2, where we discussed the projection of $\hat{\mathbf{s}}$ along an arbitrary vector $\hat{\mathbf{n}}$. The results showed that we could still have the same spin up and down states along the arbitrary vector as with the well-defined z-axis. In principle we could just replace $\hat{\mathbf{n}}$ with $\mathbf{p}$.

So far we have only spoken about massive particles and we turn to the massless ones and how to incorporate them in the next section.

3.2.2. Massless Particles. A generalization to massless particles is now possible with the introduction of the helicity states in the previous section. One reason for having problem with the massless states is that we are not able to define some kind of spin as we did with the massive states in rest. This depends on the definition of the operator $\hat{W}_\mu \hat{W}^\mu$, it is identically to zero for $m = 0$.

Indeed, the unification of the treatment of spin for particles of any spin and mass starts with the fact that the helicity states are the eigenstates of the helicity operator (3.84). There is no reference to the mass neither a rest frame for the helicity and we adopt the equation (3.85) as the definition of the helicity state for a massless particle

$$\frac{\hat{\mathbf{J}} \cdot \hat{\mathbf{P}}}{\|\hat{\mathbf{P}}\|} |\mathbf{p}, \lambda\rangle = \lambda |\mathbf{p}, \lambda\rangle, \tag{3.86}$$

where the eigenstates now explicitly only depends on the momentum. We only have the helicity λ and we define the spin s by the equation

$$s \equiv |\lambda|. \tag{3.87}$$

To begin with we only have one value at the moment. If the parity $\mathcal{P}$ is conserved, the interaction of the particle is invariant under a space reflection, the helicity state $|\mathbf{p}, \lambda\rangle$ must be mapped into a new physical state.

EXAMPLE 3.2.10. Let us determine the commutator of $\hat{\mathbf{J}} \cdot \hat{\mathbf{P}}$ and the rotation operators $\hat{J}_k$.

$$\begin{aligned}[\hat{\mathbf{J}} \cdot \hat{\mathbf{P}}, \hat{J}_k] = {} & [\hat{J}_1 \hat{P}_1 + \hat{J}_2 \hat{P}_2 + \hat{J}_3 \hat{P}_3, \hat{J}_k] = \\ & [\hat{J}_1 \hat{P}_1, \hat{J}_k] + [\hat{J}_2 \hat{P}_2, \hat{J}_k] + [\hat{J}_3 \hat{P}_3, \hat{J}_k]. \end{aligned} \tag{3.88}$$

Now we only consider one of the factors separately

$$\begin{aligned}[\hat{J}_1 \hat{P}_1, \hat{J}_k] = {} & \hat{J}_1 \hat{P}_1 \hat{J}_k - \hat{J}_k \hat{J}_1 \hat{P}_1 = \hat{J}_1 \hat{P}_1 \hat{J}_k - (\hat{J}_1 \hat{J}_k + i\epsilon_{k1l} \hat{J}_l) \hat{P}_1 = \\ & \hat{J}_1 \hat{P}_1 \hat{J}_k - \hat{J}_1 \hat{J}_k \hat{P}_1 - i\epsilon_{k1l} \hat{J}_l \hat{P}_1 = \\ & \hat{J}_1 [\hat{P}_1, \hat{J}_k] - i\epsilon_{k1l} \hat{J}_l \hat{P}_1, \end{aligned} \tag{3.89}$$

and by similar calculations we get

$$[\hat{J}_2 \hat{P}_2, \hat{J}_k] = \hat{J}_2 [\hat{P}_2, \hat{J}_k] - i\epsilon_{k2l} \hat{J}_l \hat{P}_2 \tag{3.90a}$$

$$[\hat{J}_3 \hat{P}_3, \hat{J}_k] = \hat{J}_3 [\hat{P}_3, \hat{J}_k] - i\epsilon_{k3l} \hat{J}_l \hat{P}_3. \tag{3.90b}$$

The commutator (3.88) for $k = 1$ is

$$\begin{aligned}[\hat{\mathbf{J}} \cdot \hat{\mathbf{P}}, \hat{J}_1] = {} & \hat{J}_1 [\hat{P}_1, \hat{J}_1] - i\epsilon_{11l} \hat{J}_l \hat{P}_1 + \hat{J}_2 [\hat{P}_2, \hat{J}_1] - i\epsilon_{12l} \hat{J}_l \hat{P}_2 \\ & + \hat{J}_3 [\hat{P}_3, \hat{J}_1] - i\epsilon_{13l} \hat{J}_l \hat{P}_3 = \\ & \hat{J}_1 [\hat{P}_1, \hat{J}_1] + \hat{J}_2 [\hat{P}_2, \hat{J}_1] + \hat{J}_3 [\hat{P}_3, \hat{J}_1] - i\hat{J}_3 \hat{P}_2 + i\hat{J}_2 \hat{P}_3 = \\ & -i\hat{J}_2 \hat{P}_3 + i\hat{J}_3 \hat{P}_2 - i\hat{J}_3 \hat{P}_2 + i\hat{J}_2 \hat{P}_3 = 0, \end{aligned} \tag{3.91}$$

and it is also zero for $k = 2$ and $k = 3$. We have

$$[\hat{\mathbf{J}} \cdot \hat{\mathbf{P}}, \hat{J}_k] = 0, \tag{3.92}$$

and it demonstrates that $\hat{\mathbf{J}} \cdot \hat{\mathbf{P}}$ is invariant under rotations.

Since $\hat{\mathbf{J}} \cdot \hat{\mathbf{P}}$ is invariant under rotations and a pseudoscalar under parity [**4**], a state with momentum only along the z-component

$$\mathcal{Y}|\mathbf{p}_z, \lambda\rangle \equiv e^{-i\pi \hat{J}_2}\mathcal{P}|\mathbf{p}_z, \lambda\rangle \tag{3.93}$$

has momentum $\mathbf{p}_z$ and it is an eigenstate of the helicity operator with the eigenvalue $-\lambda$. It is usual in literature written as

$$\mathcal{Y}|\mathbf{p}_z, \lambda\rangle = \eta|\mathbf{p}_z, -\lambda\rangle, \tag{3.94}$$

where η is the *intrinsic parity factor*.

A massless particle of spin s whose interactions preserves the parity has two independent helicity states ($\pm\lambda$) for a given momentum and the relation between them are given by (3.94). For example, a photon has two helicity states whereby a neutrino only has one, this particle is is believed to violate parity if it is massless.

For a massless particle it is not possible to identify the spin s directly from the eigenvalues of the operator $\hat{W}_\mu \hat{W}^\mu$ as mentioned earlier. Instead we have the helicity for our purpose but it does not specify the state $|\mathbf{p}, \lambda\rangle$ uniquely. In the case of the massive particle we used the rest frame and now we define a similar standard reference frame S^-. It is chosen in such a way that the massless particle is moving in the z-direction with a definite momentum $\bar{\mathbf{p}} \equiv (\bar{p}, 0, 0, \bar{p})$. The state is labelled as $|\bar{\mathbf{p}}, \lambda\rangle$ and in analog to (3.83)

$$|\mathbf{p}, \lambda\rangle \equiv |\bar{\mathbf{p}}, \lambda\rangle_S = U(h(\mathbf{p}, \bar{\mathbf{p}}))|\bar{\mathbf{p}}, \lambda\rangle, \tag{3.95}$$

where now $h(\mathbf{p}, \bar{\mathbf{p}})$ is the LT such that $h^{-1}(\mathbf{p}, \bar{\mathbf{p}})$ changes the frame S^- into the frame S. The momentum becomes $p^\mu = (p, \mathbf{p})$ with $\mathbf{p} = (p, \alpha, \beta)$. Indeed we have that

$$[\hat{W}_\mu, \hat{W}_\nu]|\bar{\mathbf{p}}, \lambda\rangle = i\epsilon_{\mu\nu\rho\sigma}\hat{W}^\rho \hat{P}^\sigma|\bar{\mathbf{p}}, \lambda\rangle = i\bar{p}(\epsilon_{\mu\nu\rho 0} + \epsilon_{\mu\nu\rho 3})\hat{W}^\rho|\bar{\mathbf{p}}, \lambda\rangle, \tag{3.96}$$

and in specific the commutator relations are

$$[\hat{W}^1, \hat{W}^2] = 0 \tag{3.97a}$$
$$[\hat{W}^1, \hat{W}^3] = -i\bar{p}\hat{W}^2 \tag{3.97b}$$
$$[\hat{W}^2, \hat{W}^3] = i\bar{p}\hat{W}^1, \tag{3.97c}$$

for the $\hat{W}^\mu$ acting on the helicity state $|\bar{\mathbf{p}}, \lambda\rangle$. For a massive particle at rest in Section 3.2.1, the structure of the $\hat{W}^i$ were proportional to one of Wigner's little groups. It was an $O(3)$-like group. As expected we see that the commutator relations (3.97) for the massless particle are instead proportional to the $SE(2)$

group. Indeed, one of Wigner's little groups and being an $E(2)$-like group with its structure given by (3.51). It can be written as

$$[\hat{P}^1, \hat{P}^2] = 0 \tag{3.98a}$$

$$[\hat{P}^1, \hat{J}^3] = -i\hat{P}^2 \tag{3.98b}$$

$$[\hat{P}^2, \hat{J}^3] = i\hat{P}^1, \tag{3.98c}$$

and the set of operators $(\hat{W}^3/\bar{p}, \hat{W}^1, \hat{W}^2)$ acting on the state $|\bar{\mathbf{p}}, \lambda\rangle$ obeys an algebra isomorphic to the set $(\hat{J}^3, \hat{P}^1, \hat{P}^2)$. We note that the eigenvalues of both $\hat{W}^1$ and $\hat{W}^2$ are not quantized and it is *postulated* that the physical massless particles in nature have an eigenvalue zero for these operators [**4**]

$$\hat{W}^{1,2}|\bar{\mathbf{p}}, \lambda\rangle = 0. \tag{3.99}$$

Therefore, the relation (3.96) becomes only

$$\hat{W}^\mu|\bar{\mathbf{p}}, \lambda\rangle = (\bar{p}\hat{J}_3, 0, 0, \bar{p}\hat{J}_3)|\bar{\mathbf{p}}, \lambda\rangle, \tag{3.100}$$

and together with

$$\hat{W}_\mu\hat{W}^\mu|\bar{\mathbf{p}}, \lambda\rangle = 0, \tag{3.101}$$

we get

$$\frac{1}{\bar{p}}\hat{W}^3|\bar{\mathbf{p}}, \lambda\rangle = \hat{J}^3|\bar{\mathbf{p}}, \lambda\rangle = \frac{\hat{\mathbf{J}} \cdot \hat{\mathbf{P}}}{\|\hat{\mathbf{P}}\|}|\bar{\mathbf{p}}, \lambda\rangle = \lambda|\bar{\mathbf{p}}, \lambda\rangle. \tag{3.102}$$

The *physical helicity states* are the eigenstates of $(\hat{W}^3/\bar{p}, \hat{W}^1, \hat{W}^2)$ with eigenvalues $(\lambda, 0, 0)$ and the spin s is now given by the usual rules, i.e. the eigenvalues $s(s+1)$ of the square of the spin operator $\hat{\mathbf{s}}$. We end this section with the conclusion of having a way to treat the spin of both massive and massless particles in relativistic processes.

CHAPTER 4

A Phenomenological Model of the Proton

— Is it, in permanently interacting field theories, still possible to define some kind of mass and spin with current mathematical methods? —
J. Hansson

The structure of the nucleon is not yet completely understood regarding its spin, where does it come from and how is it related to the structure of its constituent parts? How can the knowledge of the nucleon spin structure be extended to also include more general particles such as all hadrons in some manner?

At present time there are mainly two accepted pictures of the nucleon and both are dependent on the more elementary particles of *quarks* and *gluons*. The quarks have an additional charge associated with it, denoted the color charge, and it has three possible values. The interaction between quarks is known as the strong interaction described by the quantum field theory quantum chromodynamics (QCD), or the "color" force and is mediated by bosons known as gluons, analogs of the photons of the electromagnetic interactions.

In a model proposed by M. Gell-Mann and G. Zweig in 1964 [**28, 29**] where they independently suggested that hadrons were composed by smaller constituents, known as quarks (nowadays referred to as the *constituent quark model* or *naive quark model*) aimed to explain the mass, charge and spin of hadrons. The quarks were believed to have a fractional electric charge, a new degree of freedom called flavor (quark type) and to have spin one half. This model can quite well explain the static, low energy properties of the lighter baryons and predicting spectroscopy data. In this picture the properties are calculated by the use of a non-relativistic Scrödinger equation, but the essential point is that an unexcited baryon corresponds to the ground state of a three-particle system. All the quarks are in relative s-state with zero angular momentum and there is no contribution by any gluons. Therefore, the spin of the baryon is equal to the sum of the spin of its constituent quarks, which we shall see is not compatible with experimental facts.

If we would see the nucleon as a non-relativistic bound system, the naive quark model could serve as a basic model for understanding the structures of the proton and neutron. The proton can be visualized as as having two u-quarks and one d-quark (uud), while the neutron has two d-quarks and one u-quark (udd).

However, high energy experiments such as deep inelastic scattering (DIS) has played a vital role in the development of our understanding of the structure among elementary particles. In this high energy regime the results from the experiments first visualized a baryon as made up of pointlike constituents. Initially, these constituents were called *partons*, later identified as quark-antiquarks and gluons by R. P. Feynman [**30**] and also subsumed in the theory of strong interaction. These partons seem to have a lot in common with the quarks of Gell-Mann and Zweig but they have, for instance, different effective masses. Since the effective masses of the partons connected to the proton [**4**] are small compared with the total mass, it is much more favored to see the nucleon as a relativistic bound system. In principle the orbital angular momentum, polarized quark-antiquark pairs and the gluons can contribute to the total angular momentum and therefore the spin.

Consequently we can speak about two different kinds of quarks as *current quarks* (Parton model) and *constituent quarks* (Gell-Mann and Zweig model). The precise relation between these two kind of quarks is not known and especially experiments on polarized DIS raise questions about how the spin of the two models are related to each other.

4.1. Spin of the Nucleon

If we would take a baryon, according to the naive quark model, moving along the z-axis with a helicity $\lambda = 1/2$ we would expect to have

$$\sum_{\text{flavour}} s_z^{\text{flavour}} = \frac{1}{2}, \tag{4.1}$$

where the sum is over the flavour of the present quarks of the given baryon. In the case of a proton we should have

$$\sum_{\text{flavour}} s_z^{\text{flavour}} = s_z^d + s_z^d + s_z^u = \frac{1}{2} + \frac{1}{2} - \frac{1}{2} = \frac{1}{2}, \tag{4.2}$$

and it is consistent with experimental observations at a static or low energy regime. At the other end of the energy scale we have high energy interactions with large momentum transfer between the target and probe which is the case in DIS experiments. The interaction is governed by the strong interaction and DIS experiments of leptons and hadrons have played an important role for the understanding of elementary particles. For nucleons the reaction is

$$\ell + N \rightarrow \ell + X, \tag{4.3}$$

where ℓ is a lepton, N a nucleon and after the collision we have some produced particles X. This kind of DIS is visualized in Figure 4.1 and we consider the

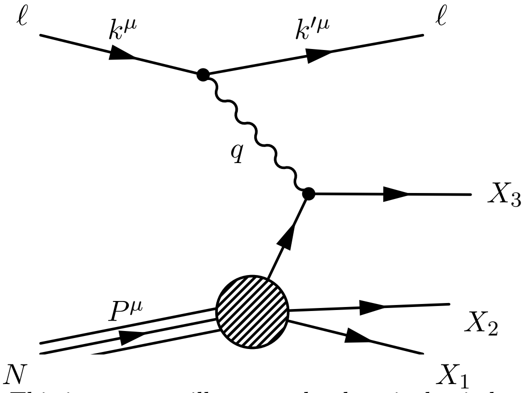

FIGURE 4.1. This is a way to illustrate the deep inelastic lepton-hadron scattering experiments in terms of Feynman diagrams. The lepton ℓ probes the constituent parts of the hadron with a virtual photon and the scattered particles are used to gain information about the inner structure.

reaction in the Lab frame, where the nucleon or proton is at rest. The initial and final four momentum of the lepton is given as

$$k^\mu = (E, \mathbf{k}), \quad k'^\mu = (E', \mathbf{k}'), \tag{4.4}$$

and the initial proton momentum is

$$P^\mu = (M, \mathbf{0}). \tag{4.5}$$

The differential cross-section to find the lepton in a solid angle $\mathrm{d}\Omega$ within some energy range $(E', E'+E)$ is the basic tool in experiments to extract information about the structure of the nucleon. The cross section is given from [**4**] as

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}\Omega\mathrm{d}E'} = \frac{\alpha^2}{2Mq^4}\frac{E'}{E}L_{\mu\nu}W^{\mu\nu}, \tag{4.6}$$

where q is the difference between k^μ and k'^μ, α is the fine structure constant, $L_{\mu\nu}$ the leptonic tensor and $W^{\mu\nu}$ the hadronic tensor. The information of the nucleon structure is given in terms of two spin-independent form factors $W_{1,2}$ and two spin-dependent ones $G_{1,2}$. They depend on the $(P \cdot q, Q^2)$ where $Q^2 \equiv -q^2$ and these form factors can be measured experimentally. In the literature one usually speaks about the scaling functions [**4**] defined as

$$F_1 \equiv MW_1(P \cdot q, Q^2), \quad F_2 \equiv \nu W_2(P \cdot q, Q^2) \tag{4.7a}$$

$$g_1 \equiv \frac{(P \cdot q)^2}{\nu}G_1(P \cdot q, Q^2), \quad g_2 \equiv \nu(P \cdot q)G_2(P \cdot q, Q^2) \tag{4.7b}$$

where ν is given as the difference between the energies of the virtual photon in the Lab. One of the theoretical results was the introduction of the Bjorken scaling variable [**4**]

$$x \equiv \frac{Q^2}{2P \cdot q} = \frac{Q^2}{2M\nu}, \tag{4.8}$$

and this made it possible to have the scaling functions $F_{1,2}$ dependent on x. From experimental observations it came to a conclusion that at fixed values of x it only depends weakly on Q^2 or in the Bjorken limit ($Q^2 \to \infty$, x fixed). It hints that the lepton is scattered off a point charge [**20**]. Since it was believed that nucleons were extended objects, it instead followed that the nucleons in DIS have a sub-structure of point-like constituents "partons". Later these partons were gradually identified as quarks and gluons [**30**]. In the parton model it is possible to treat the nucleons as made up on very fast moving, essentially free, partons or current quarks similar to the constituent quarks, but with very different effective masses. We do not go any further into this and we stay with the spin-dependent scaling functions $g_{1,2}$. For a detailed treatment on the subject of DIS and the corresponding gauge theories see for example [**31, 32**].

The most interesting part is the spin-dependent functions $g_{1,2}$ as it could give answer to much more sophisticated questions about the dynamical effects associated to the spin. But, they are harder to measure experimentally compared to the $F_{1,2}$ functions where the ordinary DIS experiments used unpolarized beams. It was mainly due to technical limits in the beginning since for $g_{1,2}$ it is required to have polarized beams and targets. But technical improvements have made it possible to have such kind of requirements and for more details see [**4**]. A direct way to measure $g_{1,2}$ is to have a longitudinally polarized beam and a target, the nucleon, polarized either along the momentum of the lepton beam or transversely to it. One has

$$\frac{\mathrm{d}^2\sigma^{\rightarrow\Rightarrow}}{\mathrm{d}\Omega\mathrm{d}E'} - \frac{\mathrm{d}^2\sigma^{\rightarrow\Leftarrow}}{\mathrm{d}\Omega\mathrm{d}E'} = -\frac{4\alpha^2 E'}{Q^2 E M\nu}\big((E + E'\cos\theta)g_1 - 2xMg_2\big) \tag{4.9a}$$

$$\frac{\mathrm{d}^2\sigma^{\rightarrow\Uparrow}}{\mathrm{d}\Omega\mathrm{d}E'} - \frac{\mathrm{d}^2\sigma^{\rightarrow\Downarrow}}{\mathrm{d}\Omega\mathrm{d}E'} = -\frac{8\alpha^2 E'^2}{Q^2 M\nu^2}\Big(\frac{\nu}{2E}g_1 + g_2\Big)\sin\theta\cos\phi \tag{4.9b}$$

with $\rightarrow$ and $\Rightarrow$ as spin directions of the lepton and nucleon, θ is the Lab scattering angle and ϕ the azimuthal angle[1].

The interplay between theory and experiments once again showed to be very important. The scaling function g_1 remained for a long time at the level of the naive quark model. It was argued that the expectation value of the helicity of the nucleon, was carried by the current quarks compatible to the constituent quarks and that it should have a value of one half (4.1).

Surprisingly, by the remarkable results of the European Muon Collaboration (EMC) at CERN in 1988 [**3**], an experiment which scattered muons off polarized protons, it is now know that (4.1) is significantly violated. It was not consistent with one half in the constituent quark model, rather with zero. This catalyzed a major programme of experimental investigations and theoretical papers of the problem, also dubbed the *"Proton Spin Crisis"* [**2**] and lead to an intensive theoretical calculations of g_1, see [**4**] for the theoretical situation.

[1] Defined on page 346 in [**4**]

In a modified theoretical picture it is common to introduce the quantity

$$\Delta\Sigma \equiv 2\langle s_z^{\text{flavour}} \rangle, \tag{4.10}$$

and it should be consistent with one if we expect that the spin of the nucleon is entirely due to the spin of the quarks as in the constituent quark model. However, as a consequence of the EMC-results it must be consistent with zero. Recent results show that only about 30 % of the spin can be carried by quarks. An explanation for the “missing” spin could be that it is provided by the gluon spin and by the orbital angular momentum of the quarks and gluons.

In high energy physics it is common to speak about different sum-rules of different phenomena which are used to get a validity between theory and experimental observations. In the literature it is stated that the total helicity of a nucleon satisfies the spin sum rule

$$\frac{1}{2} = \frac{1}{2}\Delta\Sigma(1, Q^2) + \Delta g(1, Q^2) + L_q(Q^2) + L_g(Q^2), \tag{4.11}$$

with

$$\Delta\Sigma(1, Q^2) = \sum_q (\Delta q(1, Q^2) + \Delta\bar{q}(1, Q^2)),$$

is the twice the total helicity of the quarks, $\Delta g(1, Q^2)$ is the total helicity of gluons and $L_{q,g}(Q^2)$ is the contribution of the orbital angular momentum of quarks and gluons along the motion direction. The explanation of the missing spin of the nucleon with the factors of L_q and L_q is at present time the most favored one. But, these quantities have not yet been measured and we do not know if it possible to achieve that. An alternative to solve this puzzle could be our proposed model of an ”hybrid particle” discussed in Section 4.2.

So far we have only dealt with the helicity distribution but also the transverse spin distribution of the quarks is another missing piece in the proton spin puzzle. The scaling function g_2 has no simple interpretation in the perturbative regime of the strong interaction as g_1, and it opens the way to new theoretical and experimental suggestions of spin observables [**5**].

4.2. Our Model

The strong interaction environment is controlled by the nonlinear quantum field theory known as QCD and the structure of the baryon in terms of constituent quarks involves the non-perturbative regime which cannot yet be studied exactly by any theoretical methods. Consequently, we must rely on indirect information from experiments at other regimes such as DIS. However, from the discussion of the previous section it points out that we do not yet have a consistent theory of the hadrons, where the proton with the ”disappearance” of its spin serves as an excellent example. In principle we have the naive quark

	Massive (Slow)	Covariance	Massless (Fast)
Energy and Momentum	$E = \mathbf{p}^2/2m$	$E^2 = \mathbf{p}^2 + m^2$	$E = \mathbf{p}$
Internal space-time symmetry	$\hat{S}_3$ $\hat{S}_1,\ \hat{S}_2$	Wigner's little group	$\hat{S}_3$ Gauge transformations
Relativistic extended particles	Naive quark model	? (Proposition 2)	Parton model

TABLE 1. Einstein's energy and momentum relation unifies the massive and massless particles. Whereby Wigner's little group unifies the internal space time symmetries and the question is whether there exists a connection between the constituents and current quarks? This table has been adapted from [**33**].

model which works well for the proton at rest or at low energy, but at high energies the proton seems to be made of pointlike "partons" not really compatible with the naive model.

What if these two different models could be linked to each other in some manner as Wigner's little group do for the internal structure of massive (slow) and massless (fast) particles. A further motivation is the linking of the energy and momentum relation between fast and slow particles. This kind of question has been addressed earlier such as whether the parton model could be a Lorentz-boosted naive model [**34**], a unification of the constituent and current quarks. A summary is presented in Table 1.

The spin disappearance of the proton motivates our alternative approach where we force the proton to be described always in terms of "clothed" partons. The quarks are coupled to the interaction fields of the gluons and it implies that the quark and gluon fields become irrevocably admixed in an undisturbed proton. Since the coupling is dependent on the effective strong coupling constant α_s which has an asymptotically free behavior as $Q^2 \to \infty$, it weakens at high energies, meaning that we would have a strong contribution of the mixing at low energies "constituents quarks" and less on the opposite range "current quarks". This could serve as a basis for the explanation of the disappearance of the spin due to quarks in the proton. In principle we have two different measurements of the same quasi-particle at given energy regimes, similar in idea to the form factors that are dependent on Q^2.

We make the following proposition.

PROPOSITION 2. *The hadron is treated as being a composition of three quasi-particles of four-momentum*

$$p^\mu \mapsto \tilde{p}^\mu = p^\mu + \alpha_i T^a A_a^\mu, \tag{4.12}$$

where p^μ is the four-momentum of a "free" hypothetical quark, α_i is the effective coupling constant and $T^a A_a^\mu$ the combined contribution of the self-interaction gluon fields.

Equation (4.12) in Proposition 2 is just the normal "minimal coupling" taken literally. The first step towards this phenomenological model is to see how this quasi-particle behaves in current mathematical methods of particle physics and what kind of additional requirements must hold. We discuss this particle in a relativistic approach and we adopt all the methods obtained in the previous chapters. At this stage the most important issue is if we still can define a kind of mass and spin connected to the Casimir operators of the Poincaré group.

In principle, Proposition 2 could be made even more general to include particles such as leptons and other kind if interactions. For a more simple start, this gives us the opportunity to study the quasi-particle with the interaction of the electromagnetic field, which is governed by the symmetry group of $U(1)$ [**23**]. With the electromagnetic interaction there is only *one* interaction field present and the particle field interacts with itself by the photon. The theory which describes such phenomena is known as QED and the strength of the interaction is small, a parameter that changes very little under normal circumstances (diverges logarithmically as $Q^2 \to \infty$). Therefore QED works well in whatever energy regime we may have. However, it also means that the QED quasi-particle always, to a very good approximation, will act like a "normal" charged fermion. This should be contrasted to the quasi-particle in QCD, where the gluon (boson) part dominates at low energies and the quark (fermion) part dominates at high energies.

The four-momentum of a QED quasi-particle becomes

$$\tilde{p}^\mu = p^\mu + \alpha_1 A^\mu. \tag{4.13}$$

with $\alpha_1 \approx 1/137$, A^μ the photon field and p^μ the four-momentum of a "free" charged fermion.

In analog to Section 3.1 the invariant mass $\tilde{m}$ of the quasi-particle is determined by the operator $\hat{P}_\mu \hat{P}^\mu$ and it follows

$$\begin{aligned}\tilde{p}_\mu \tilde{p}^\mu &= \eta^{\mu\nu}\tilde{p}_\mu \tilde{p}_\nu = \tilde{p}_0^2 - \tilde{\mathbf{p}}^2 = \\ &(\tilde{p}_0 + \alpha_1 A_0)^2 - (\tilde{p}_1 + \alpha_1 A_1)^2 - (\tilde{p}_2 + \alpha_1 A_2)^2 - (\tilde{p}_3 + \alpha_1 A_3)^2 = \\ &\tilde{p}_0^2 - \tilde{p}_1^2 - \tilde{p}_2^2 - \tilde{p}_3^2 + \alpha_1^2 \left[A_0^2 - A_1^2 - A_2^2 - A_3^2\right] \\ &+2\alpha_1 \left[\tilde{p}_0 A_0 - \tilde{p}_1 A_1 - \tilde{p}_2 A_2 - \tilde{p}_3 A_3\right] = \\ &\tilde{p}_\mu \tilde{p}^\mu + 2\alpha_1 \tilde{p}_\nu A^\nu + \alpha_1 A_\rho A^\rho, \end{aligned} \tag{4.14}$$

where the term $\tilde{p}_\mu \tilde{p}^\mu$ is the squared invariant mass of the lepton, $2\alpha_1 \tilde{p}_\nu A^\nu$ a new mixing term and $\alpha_1 A_\rho A^\rho$ is identically to zero since the fields obey $A_\rho A^\rho = 0$. The squared invariant mass of the quasi-particle is

$$\tilde{p}_\mu \tilde{p}^\mu = \tilde{m}^2 \tag{4.15a}$$

$$\tilde{m}^2 \equiv \tilde{p}_\mu \tilde{p}^\mu + 2\alpha_1 \tilde{p}_\nu A^\nu. \tag{4.15b}$$

Notice that the mixing term is dependent on the coupling constant α_1. Since it is a small parameter we only have a small contribution to the mass in QED. For example, if we have positronium, a state of an electron and positron with the interaction of photons, we should be able to measure a small difference in the mass compared to the sum of each particle by themselves. At rest such a particle should as well posses some kind of intrinsic angular momentum analog to Section 3.1 in the discussions of Wigner's little group. If we operate the set of Pauli-Lubanski operators (3.57)

$$\hat{W}_\mu = -\frac{1}{2}\epsilon_{\mu\nu\rho\sigma}\hat{M}^{\mu\nu}\hat{P}^\rho.$$

on the state $|\tilde{m}, \mathring{\mathbf{p}}\rangle$, namely $\hat{W}_\sigma \hat{W}^\sigma$ we get

$$\hat{W}_\sigma \hat{W}^\sigma |\tilde{m}, \mathring{\mathbf{p}}\rangle = \tilde{m}^2 \hat{\mathbf{M}}^2 |\tilde{m}, \mathring{\mathbf{p}}\rangle, \tag{4.16}$$

where $\hat{\mathbf{M}} = [\hat{M}_{23}, \hat{M}_{31}, \hat{M}_{12}]$ and it could be identified with the general angular momentum operator $\hat{\mathbf{J}}$ of the helicity operator (3.84). The next step is to determine if there is an eigenvalue similar to $J(J+1)$ present, where J is the quantum number of the general angular momentum. This has not yet been accomplished. One reason for this is due to the mixing part that couples, as it is a combination of a fermion and boson which gives that the quasi-particle is not a pure fermion or boson. However, since the coupling is weak it is a reasonable approximation in QED to speak of only having a pure fermion as the pure lepton and we could invoke all the known material of the mass and spin presented in the previous chapters.

The more interesting part is the approach of quasi-particles in the proton where we use Proposition 2 and write the four-momentum as

$$\tilde{p}^\mu + \alpha_s T^a A_a^\mu, \; a = 1, 2, ..., 8. \tag{4.17}$$

where α_s is the effective strong coupling constant, a running coupling constant $\alpha_s(Q^2)$ with an asymptotic behavior in characteristic energy scales of the process. We have a contribution from eight different interaction fields. This is due to the fact that the $U(1)$ gauge group is replaced by the $SU(3)$ group of gauge transformations. In this case the T^a are dependent on the structure constants f_{abc} of the $SU(3)$ group and the only nonzero elements values of permutations are

$$f_{123} = 1 \tag{4.18a}$$

$$f_{458} = f_{678} = \sqrt{3}/2 \tag{4.18b}$$

$$f_{147} = f_{165} = f_{246} = f_{257} = f_{345} = f_{376} = 1/2, \tag{4.18c}$$

which gives a much more complicated mixing then the $U(1)$ case. So far we have not managed to obtain any satisfying results on either the mass or spin. The more elementary group of $SU(2)$ could allow us to study a non-Abelian case where there still are boson self-interactions, but handling fewer mixing elements as a result. We then have a four-momentum as

$$\tilde{p}^\mu + \alpha_s T^a A_a^\mu, \ a = 1, 2, 3, \tag{4.19}$$

with α_s as before but only handling three different fields. Now, the non-zero structure constants are only the values of permutations as

$$f_{123} = 1, \tag{4.20}$$

and even in this case we have not yet managed to find any satisfying results for mass or spin. The main problem so far is to represent the four-momentum in a reasonable way since to our knowledge the only fundamental representations of T^a are the Gell-Mann matrices for the $SU(3)$ group and the Pauli matrices of the $SU(2)$ group [**20**]. This forces the four-momentum to be expressed in a matrix form which give mathematical results which cannot yet be understood in a clear cut way as with the $U(1)$ case.

We have not yet achieved any conclusive results and there is still a lot of work do be done. Two important questions are:

(i) Is it possible to apply (a generalization of) Wigner's classification in terms of mass and spin to fields with permanent interactions, like bound states?

(ii) Is it possible to deduce a "quantum-statistic" for hybrid particles that continuously go between Fermi-Dirac ($Q^2 \to \infty$) and Bose-Einstein ($Q^2 \to 0$) , as α_s is shifted?

We hope that this can give a clearer insight into spin problems in bound states versus free states generally, and into the "Proton Spin Crisis" particularly, relating (quasi-free) partons to bound constituents quarks.

Summary

In this thesis we have discussed the role of spin in particle physics. It is a purely quantum mechanical phenomenon and there is no valid classical interpretation. In the non-relativistic approach we need to modify the theory by hand, whereas it is beautiful incorporated in the relativistic approach, where it is due to the symmetries of the Poincaré group. The spin of a particle is, besides the invariant mass, used as a fundamental property to classify different types of elementary particles. This is due to the important work of Wigner.

Still, the spin itself is not fully understood physically reflected, for example, in the "Proton Spin Crisis". Is it the spin that we do not understand or is it due to a wrongly built theory? However, the spin property is a very powerful "tool" to get a deeper understanding of nature, as well to verify or falsify different theories.

The approach of using quasi-particles to explain the structure of the proton and in particular its spin has yet only been started. However, it has raised several issues that must be answered in order to get a consistent theory.

The idea of a contribution dependent on the running coupling constant is very promising in our opinion. It could serve as the key element in order to link the constituent and current quarks and also an explanation of the "missing spin" of the proton.

The mixing elements consists of factors that are usually defined with different statistics such as being either fermions or bosons. How should we treat such an element if we would go one step further to build a wave function?

A first step towards the quasi-particle picture could be to do a deeper investigation of the electromagnetic interaction and continue with the simplest non-Abelian example, $SU(2)$. This could give us valuable information on how to proceed with the full $SU(3)$ QCD model.

Acknowledgements

I would like to thank Dr. Johan Hansson, my thesis advisor, for many useful discussions and his critical reading of the manuscript. I am also thankful to Prof. Sverker Fredriksson for his valuable support during my time at Luleå University of Technology.

I also had the pleasure to meet Prof. Mauro Anselmino at the Department of Theoretical Physics at the University of Torino, and I thank him for his kind hospitality and discussions regarding spin in particle physics during my visit.

Finally, a word of gratitude goes to my family for their support and I would also like to thank everybody at the Division of Physics at Luleå University of Technology.

Luleå
October 30, 2003

Christian Türk

Bibliography

1. E. P. Wigner, *On Unitary Representations of the Inhomogeneous Lorentz Group*, Ann. Math. **40** (1939), 149.
2. E. Leader and M. Anselmino, *A Crisis in the Parton Model: where, oh where is the proton's spin?*, Z. Phys. **C41** (1988), 239.
3. J. Ashman *et al.*, *A Measurment of the Spin Asymmetry and Determination of the Structure Function g_1 in Deep Inelastic Muon-Proton Scattering*, Phys. Lett. **B206** (1988), 364.
4. E. Leader, *Spin in Particle Physics*, Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology, Cambridge University Press, Cambridge, 2001.
5. M. Anselmino and G. van der Steenhoven, *Summary of the Working Group on Spin Physics*, arXiv e-Print Archive: `hep-ph/0309126`.
6. G. E. Uhlenbeck and S. Goudsmit, Naturwiss. **13** (1925), 953.
7. Sin itiro Tomonaga, *The Story of Spin*, The University of Chicago Press, 1997.
8. S. Gasiorowicz, *Quantum Physics*, 2 ed., John Wiley & Sons, New York, 1996.
9. E. Noether, *Invariante Variationsprobleme*, Könige Gesellschaft der Wissenschaften Göttingen Nachrichten Mathematik-physik Klasse **2** (1918), 235.
10. J.Neumann, *Mathematical Foundations of Quantum Mechanics*, Princeton University Press, Princeton, 1955.
11. G. Sparr & A. Sparr, *Kontinuerliga System*, 2 ed., Studentlitteratur, Lund, 2000.
12. W. Gerlach and O. Stern, *Der Experimentelle Nachweis der Richtungsquantelung*, Z. Phys. **9** (1922), 349.
13. O. Stern, *Ein Weg zur Experimentellen Prüfung der Richtungsquantelung im Magnetfeld*, Z. Phys. **7** (1921), 249.
14. B.H. Bransden & C.J. Joachain, *Quantum Mechanics*, 2 ed., Pearson Education, Harlow, 2000.
15. G. W. Bluman & S. C. Anco, *Symmetry and Integration Methods for Differential Equations*, 2 ed., Applied Mathematical Sciences, Springer-Verlag New York, 2002.
16. N. Euler, *A First Course in Nonlinear Mathematical Physics*, Luleå University of Technology, 2000, Lecture Notes.
17. J. Fuchs and C. Schweigert, *Symmetries, Lie Algebras and Representations, A Graduate Course for Physicists*, Cambridge Monographs on Mathematical Physics, Cambridge University Press, Cambridge, 1997.
18. Y. S. Kim and M. E. Noz, *Theory and Applications of the Poincaré Group*, Fundamental Theories of Physics, D. Reidel Publishing, 1986.
19. L.D. Landau and E.M. Lifshitz, *Mechanics*, 3 ed., Course of Theoretical Physics, vol. 1, Butterworth-Heinenann, Oxford, 1998.
20. F. Halzen and A.D. Martin, *Quarks & Leptons an Introductory Course in Modern Particle Physics*, John Wiley & Sons, Canada, 1984.
21. P. A. M. Dirac, Proc. Roy. Soc. **A 117** (1927), 610.
22. R. A. d'Inverno, *Introducing Einsteins's Relativity*, Oxford University Press Inc., 2000.

23. F. Gross, *Relativistic Quantum Mechanics and Field Theory*, John Wiley & Sons, New York, 1999.
24. R. P. Feynman, *QED - the Strange Theory of Light and Matter*, Princeton University Press, 1985.
25. A. Janner and T. Janssen, Physica **60** (1972), 292.
26. M. Hamermesh, *Group Theory and Its Application to Physical Problems*, Addison-Wesley, Massachusetts, 1964.
27. S. M. Troshin and N. E. Tyurin, *Spin Phenomena in Particle Interactions*, World Scientific Publishing, 1994.
28. M. Gell-Mann, *A Schematic Model of Bayrons and Mesons*, Phys. Lett. **8** (1964), 214.
29. G. Zweig, *An SU(3) Model for Strong Interaction Symmetry and Its Breaking*, CERN preprints **TH-401, TH-412** (1964).
30. R. P. Feynman, *Very High-Energy Collisions of Hadrons*, Phys. Rev. Lett. **23** (1969), 1415.
31. E. Leader and E. Predazzi, *An Introduction to Gauge Theories and Modern Particle Physics*, Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology, vol. 1, Cambridge University Press, Cambridge, 1996.
32. ———, *An Introduction to Gauge Theories and Modern Particle Physics*, Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology, vol. 2, Cambridge University Press, Cambridge, 1996.
33. Y. S. Kim, *Wigner's Spins, Feynman's Partons, and their Common Ground*, arXiv e-Print Archive: `quant-ph/0205018`.
34. Y. S. Kim and M. E. Noz, Phys. Rev. D **15** (1977), 335.