Notes on Quantum Mechanics

K. Schulten

Department of Physics and Beckman Institute
University of Illinois at Urbana–Champaign
405 N. Mathews Street, Urbana, IL 61801 USA

(April 18, 2000)

Preface

The following notes introduce *Quantum Mechanics* at an advanced level addressing students of Physics, Mathematics, Chemistry and Electrical Engineering. The aim is to put mathematical concepts and techniques like the path integral, algebraic techniques, Lie algebras and representation theory at the readers disposal. For this purpose we attempt to motivate the various physical and mathematical concepts as well as provide detailed derivations and complete sample calculations. We have made every effort to include in the derivations all assumptions and all mathematical steps implied, avoiding omission of supposedly 'trivial' information. Much of the author's writing effort went into a web of cross references accompanying the mathematical derivations such that the intelligent and diligent reader should be able to follow the text with relative ease, in particular, also when mathematically difficult material is presented. In fact, the author's driving force has been his desire to pave the reader's way into territories unchartered previously in most introductory textbooks, since few practitioners feel obliged to ease access to their field. Also the author embraced enthusiastically the potential of the TEX typesetting language to enhance the presentation of equations as to make the logical pattern behind the mathematics as transparent as possible. Any suggestion to improve the text in the respects mentioned are most welcome. It is obvious, that even though these notes attempt to serve the reader as much as was possible for the author, the main effort to follow the text and to master the material is left to the reader.

The notes start out in Section 1 with a brief review of *Classical Mechanics* in the Lagrange formulation and build on this to introduce in Section 2 *Quantum Mechanics* in the closely related *path integral formulation.* In Section 3 the *Schrödinger equation* is derived and used as an alternative description of continuous quantum systems. Section 4 is devoted to a *detailed presentation of the harmonic oscillator*, introducing algebraic techniques and comparing their use with more conventional mathematical procedures. In Section 5 we introduce the *presentation theory of the 3-dimensional rotation group and the group SU(2)* presenting Lie algebra and Lie group techniques and applying the methods to the theory of angular momentum, of the spin of single particles and of angular momenta and spins of composite systems. In Section 6 we present the *theory of many–boson and many–fermion systems* in a formulation exploiting the algebra of the associated creation and annihilation operators. Section 7 provides an introduction to *Relativistic Quantum Mechanics* which builds on the representation theory of the Lorentz group and its complex relative $Sl(2,\mathbb{C})$. This section makes a strong effort to introduce Lorentz–invariant field equations systematically, rather than relying mainly on a heuristic amalgam of Classical Special Relativity and Quantum Mechanics.

The notes are in a stage of continuing development, various sections, e.g., on the semiclassical approximation, on the Hilbert space structure of Quantum Mechanics, on scattering theory, on perturbation theory, on Stochastic Quantum Mechanics, and on the group theory of elementary particles will be added as well as the existing sections expanded. However, at the present stage the notes, for the topics covered, should be complete enough to serve the reader.

The author would like to thank Markus van Almsick and Heichi Chan for help with these notes. The author is also indebted to his department and to his University; their motivated students and their inspiring atmosphere made teaching a worthwhile effort and a great pleasure.

These notes were produced entirely on a Macintosh II computer using the TEX typesetting system, Textures, Mathematica and Adobe Illustrator.

Klaus Schulten
University of Illinois at Urbana–Champaign
August 1991

Contents

Chapter 1

Lagrangian Mechanics

Our introduction to Quantum Mechanics will be based on its correspondence to Classical Mechanics. For this purpose we will review the relevant concepts of Classical Mechanics. An important concept is that the equations of motion of Classical Mechanics can be based on a variational principle, namely, that along a path describing classical motion the action integral assumes a minimal value (Hamiltonian Principle of Least Action).

1.1 Basics of Variational Calculus

The derivation of the Principle of Least Action requires the tools of the calculus of variation which we will provide now.
Definition: A *functional* $S[\,]$ is a map

$$S[\,] \,:\, \mathcal{F} \,\rightarrow \mathbb{R}\,;\; \mathcal{F} \,=\, \{\vec{q}(t);\, \vec{q}:\, [t_0,t_1] \,\subset\, \mathbb{R} \,\rightarrow\, \mathbb{R}^M;\, \vec{q}(t) \text{ differentiable}\} \tag{1.1}$$

from a space $\mathcal{F}$ of vector-valued functions $\vec{q}(t)$ onto the real numbers. $\vec{q}(t)$ is called the *trajectory* of a system of M degrees of freedom described by the *configurational coordinates* $\vec{q}(t) \,=\, (q_1(t), q_2(t), \ldots q_M(t))$.
In case of N classical particles holds $M \,=\, 3N$, i.e., there are $3N$ configurational coordinates, namely, the position coordinates of the particles in any kind of coordianate system, often in the Cartesian coordinate system. It is important to note at the outset that for the description of a classical system it will be necessary to provide information $\vec{q}(t)$ as well as $\frac{d}{dt}\vec{q}(t)$. The latter is the velocity vector of the system.
Definition: A functional $S[\,]$ is *differentiable*, if for any $\vec{q}(t) \,\in \mathcal{F}$ and $\delta\vec{q}(t) \,\in\, \mathcal{F}_\epsilon$ where

$$\mathcal{F}_\epsilon \,=\, \{\delta\vec{q}(t);\, \delta\vec{q}(t) \,\in \mathcal{F},\, |\delta\vec{q}(t)| \,<\, \epsilon,\, |\frac{d}{dt}\delta\vec{q}(t)| \,<\, \epsilon, \forall t,\, t \,\in\, [t_0,t_1] \subset \mathbb{R}\} \tag{1.2}$$

a functional $\delta S[\,\cdot\,,\,\cdot\,]$ exists with the properties

$$\begin{aligned} &(i) \quad S[\vec{q}(t) \,+\, \delta\vec{q}(t)] \,=\, S[\vec{q}(t)] \,+\, \delta S[\vec{q}(t), \delta\vec{q}(t)] \,+\, O(\epsilon^2) \\ &(ii) \quad \delta S[\vec{q}(t), \delta\vec{q}(t)] \text{ is linear in } \delta\vec{q}(t). \end{aligned} \tag{1.3}$$

$\delta S[\,\cdot\,,\,\cdot\,]$ is called the *differential* of $S[\,]$. The linearity property above implies

$$\delta S[\vec{q}(t), \alpha_1\,\delta\vec{q}_1(t) \,+\, \alpha_2\,\delta\vec{q}_2(t)] \,=\, \alpha_1\,\delta S[\vec{q}(t), \delta\vec{q}_1(t)] \,+\, \alpha_2\,\delta S[\vec{q}(t), \delta\vec{q}_2(t)]\,. \tag{1.4}$$

Note: $\delta\vec{q}(t)$ describes small variations around the trajectory $\vec{q}(t)$, i.e. $\vec{q}(t) + \delta\vec{q}(t)$ is a 'slightly' different trajectory than $\vec{q}(t)$. We will later often assume that only variations of a trajectory $\vec{q}(t)$ are permitted for which $\delta\vec{q}(t_0) = 0$ and $\delta\vec{q}(t_1) = 0$ holds, i.e., at the ends of the time interval of the trajectories the variations vanish.
It is also important to appreciate that $\delta S[\,\cdot\,,\,\cdot\,]$ in conventional differential calculus does not correspond to a differentiated function, but rather to a differential of the function which is simply the differentiated function multiplied by the differential increment of the variable, e.g., $df = \frac{df}{dx}dx$ or, in case of a function of M variables, $df = \sum_{j=1}^{M} \frac{\partial f}{\partial x_j} dx_j$.
We will now consider a particular class of functionals $S[\,]$ which are expressed through an integral over the the interval $[t_0, t_1]$ where the integrand is a function $L(\vec{q}(t), \frac{d}{dt}\vec{q}(t), t)$ of the configuration vector $\vec{q}(t)$, the velocity vector $\frac{d}{dt}\vec{q}(t)$ and time t. We focus on such functionals because they play a central role in the so-called action integrals of Classical Mechanics.
In the following we will often use the notation for velocities and other time derivatives $\frac{d}{dt}\vec{q}(t) = \dot{\vec{q}}(t)$ and $\frac{dx_j}{dt} = \dot{x}_j$.
Theorem: Let

$$S[\vec{q}(t)] \;=\; \int_{t_0}^{t_1} dt\, L(\vec{q}(t), \dot{\vec{q}}(t), t) \tag{1.5}$$

where $L(\,\cdot\,,\,\cdot\,,\,\cdot\,)$ is a function differentiable in its three arguments. It holds

$$\delta S[\vec{q}(t), \delta\vec{q}(t)] \;=\; \int_{t_0}^{t_1} dt\, \left\{ \sum_{j=1}^{M} \left[\frac{\partial L}{\partial q_j} \;-\; \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right)\right] \delta q_j(t) \right\} \;+\; \sum_{j=1}^{M} \frac{\partial L}{\partial \dot{q}_j} \delta q_j(t) \Bigg|_{t_0}^{t_1} . \tag{1.6}$$

For a proof we can use conventional differential calculus since the functional (1.6) is expressed in terms of 'normal' functions. We attempt to evaluate

$$S[\vec{q}(t) \,+\, \delta\vec{q}(t)] \;=\; \int_{t_0}^{t_1} dt\, L(\vec{q}(t) \,+ \delta\vec{q}(t), \dot{\vec{q}}(t) \,+ \delta\dot{\vec{q}}(t), t) \tag{1.7}$$

through Taylor expansion and identification of terms linear in $\delta q_j(t)$, equating these terms with $\delta S[\vec{q}(t), \delta\vec{q}(t)]$. For this purpose we consider

$$L(\vec{q}(t) \,+ \delta\vec{q}(t), \dot{\vec{q}}(t) \,+ \delta\dot{\vec{q}}(t), t) \;=\; L(\vec{q}(t), \dot{\vec{q}}(t), t) \;+\; \sum_{j=1}^{M} \left(\frac{\partial L}{\partial q_j}\delta q_j \;+\; \frac{\partial L}{\partial \dot{q}_j}\delta\dot{q}_j \right) \;+\; O(\epsilon^2) \tag{1.8}$$

We note then using $\frac{d}{dt} f(t)g(t) = \dot{f}(t)g(t) + f(t)\dot{g}(t)$

$$\frac{\partial L}{\partial \dot{q}_j}\delta\dot{q}_j \;=\; \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\delta q_j\right) \;-\; \left(\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j}\right)\delta q_j \,. \tag{1.9}$$

This yields for $S[\vec{q}(t) + \delta\vec{q}(t)]$

$$S[\vec{q}(t)] \;+\; \int_{t_0}^{t_1} dt \sum_{j=1}^{M} \left[\frac{\partial L}{\partial q_j} \;-\; \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right)\right] \delta q_j \;+\; \int_{t_0}^{t_1} dt \sum_{j=1}^{M} \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\delta q_j\right) \;+\; O(\epsilon^2) \tag{1.10}$$

From this follows (1.6) immediately.

We now consider the question for which functions the functionals of the type (1.5) assume extreme values. For this purpose we define
Definition: An extremal of a differentiable functional $S[\,]$ is a function $q_e(t)$ with the property

$$\delta S[\vec{q}_e(t), \delta\vec{q}(t)] \;=\; 0 \quad \text{for all } \delta\vec{q}(t) \,\in\, \mathcal{F}_\epsilon. \tag{1.11}$$

The extremals $\vec{q}_e(t)$ can be identified through a condition which provides a suitable differential equation for this purpose. This condition is stated in the following theorem.
Theorem: *Euler–Lagrange Condition*
For the functional defined through (1.5), it holds in case $\delta\vec{q}(t_0) \,=\, \delta\vec{q}(t_1) \,=\, 0$ that $\vec{q}_e(t)$ is an extremal, if and only if it satisfies the conditions ($j \,=\, 1, 2, \ldots, M$)

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) \;-\; \frac{\partial L}{\partial q_j} \;=\; 0 \tag{1.12}$$

The proof of this theorem is based on the property
Lemma: If for a continuous function f(t)

$$f:\; [t_0, t_1] \,\subset \mathbb{R} \;\rightarrow\; \mathbb{R} \tag{1.13}$$

holds

$$\int_{t_0}^{t_1} dt\, f(t)h(t) \;=\; 0 \tag{1.14}$$

for any continuous function $h(t) \,\in\, \mathcal{F}_\epsilon$ with $h(t_0) \,=\, h(t_1) \,=\, 0$, then

$$f(t) \;\equiv\; 0 \quad \text{on } [t_0, t_1]. \tag{1.15}$$

We will not provide a proof for this Lemma.
The proof of the above theorem starts from (1.6) which reads in the present case

$$\delta S[\vec{q}(t), \delta\vec{q}(t)] \;=\; \int_{t_0}^{t_1} dt\, \left\{ \sum_{j=1}^{M} \left[\frac{\partial L}{\partial q_j} \;-\; \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right)\right] \delta q_j(t) \right\} . \tag{1.16}$$

This property holds for any δq_j with $\delta\vec{q}(t) \,\in\, \mathcal{F}_\epsilon$. According to the Lemma above follows then (1.12) for $j \,=\, 1, 2, \ldots M$. On the other side, from (1.12) for $j \,=\, 1, 2, \ldots M$ and $\delta q_j(t_0) \,=\, \delta q_j(t_1) \,=\, 0$ follows according to (1.16) the property $\delta S[\vec{q}_e(t), \cdot\,] \,\equiv\, 0$ and, hence, the above theorem.

An Example

As an application of the above rules of the variational calculus we like to prove the well-known result that a straight line in $\mathbb{R}^{\not\in}$ is the shortest connection (geodesics) between two points (x_1, y_1) and (x_2, y_2). Let us assume that the two points are connected by the path $y(x)$, $y(x_1) \,=\, y_1$, $y(x_2) \,=\, y_2$. The length of such path can be determined starting from the fact that the incremental length ds in going from point $(x, y(x))$ to $(x + dx, y(x + dx))$ is

$$ds \;=\; \sqrt{(dx)^2 \,+\, (\frac{dy}{dx}dx)^2} \;=\; dx\sqrt{1 \,+\, (\frac{dy}{dx})^2}\,. \tag{1.17}$$

The total path length is then given by the integral

$$s \;=\; \int_{x_0}^{x_1} dx \sqrt{1 \,+\, (\frac{dy}{dx})^2}\;. \tag{1.18}$$

s is a functional of $y(x)$ of the type (1.5) with $L(y(x), \frac{dy}{dx}) \,=\, \sqrt{1 \,+\, (dy/dx)^2}$. The shortest path is an extremal of $s[y(x)]$ which must, according to the theorems above, obey the Euler–Lagrange condition. Using $y' \,=\, \frac{dy}{dx}$ the condition reads

$$\frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right) \;=\; \frac{d}{dx}\left(\frac{y'}{\sqrt{1 \,+\, (y')^2}}\right) \;=\; 0\;. \tag{1.19}$$

From this follows $y'/\sqrt{1 \,+\, (y')^2} \,=\, const$ and, hence, $y' \,=\, const$. This in turn yields $y(x) \,=\, ax \,+\, b$. The constants a and b are readily identified through the conditons $y(x_1) \,=\, y_1$ and $y(x_2) \,=\, y_2$. One obtains

$$y(x) \;=\; \frac{y_1 \,-\, y_2}{x_1 \,-\, x_2}\,(x \,-\, x_2) \,+\, y_2\;. \tag{1.20}$$

Exercise 1.1.1: Show that the shortest path between two points on a sphere are great circles, i.e., circles whose centers lie at the center of the sphere.

1.2 Lagrangian Mechanics

The results of variational calculus derived above allow us now to formulate the Hamiltonian Principle of Least Action of Classical Mechanics and study its equivalence to the Newtonian equations of motion.

Threorem: *Hamiltonian Principle of Least Action*

The trajectories $\vec{q}(t)$ of systems of particles described through the Newtonian equations of motion

$$\frac{d}{dt}(m_j \dot{q}_j) \,+\, \frac{\partial U}{\partial q_j} \;=\; 0 \quad ;\; j = 1, 2, \ldots M \tag{1.21}$$

are extremals of the functional, the so-called *action integral*,

$$S[\vec{q}(t)] \;=\; \int_{t_0}^{t_1} dt\, L(\vec{q}(t), \dot{\vec{q}}(t), t) \tag{1.22}$$

where $L(\vec{q}(t), \dot{\vec{q}}(t), t)$ is the so-called *Lagrangian*

$$L(\vec{q}(t), \dot{\vec{q}}(t), t) \;=\; \sum_{j=1}^{M} \frac{1}{2} m_j \dot{q}_j^2 \;-\; U(q_1, q_2, \ldots, q_M)\;. \tag{1.23}$$

Presently we consider only velocity–independent potentials. Velocity–dependent potentials which describe particles moving in electromagnetic fields will be considered below.

For a proof of the Hamiltonian Principle of Least Action we inspect the Euler–Lagrange conditions associated with the action integral defined through (1.22, 1.23). These conditions read in the present case

$$\frac{\partial L}{\partial q_j} \; - \; \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) \; = \; 0 \; \rightarrow \; -\frac{\partial U}{\partial q_j} \; - \; \frac{d}{dt}(m_j \dot{q}_j) \; = \; 0 \tag{1.24}$$

which are obviously equivalent to the Newtonian equations of motion.

Particle Moving in an Electromagnetic Field

We will now consider the Newtonian equations of motion for a single particle of charge q with a trajectory $\vec{r}(t) \; = \; (x_1(t), x_2(t), x_3(t))$ moving in an electromagnetic field described through the electrical and magnetic field components $\vec{E}(\vec{r}, t)$ and $\vec{B}(\vec{r}, t)$, respectively. The equations of motion for such a particle are

$$\frac{d}{dt}(m\dot{\vec{r}}) \; = \; \vec{F}(\vec{r}, t) \; ; \; \vec{F}(\vec{r}, t) \; = \; q\,\vec{E}(\vec{r}, t) \; + \; \frac{q}{c}\,\vec{v} \times \vec{B}(\vec{r}, t) \tag{1.25}$$

where $\frac{d\vec{r}}{dt} \; = \; \vec{v}$ and where $\vec{F}(\vec{r}, t)$ is the Lorentz force.
The fields $\vec{E}(\vec{r}, t)$ and $\vec{B}(\vec{r}, t)$ obey the Maxwell equations

$$\nabla \times \vec{E} \; + \; \tfrac{1}{c}\tfrac{\partial}{\partial t}\vec{B} \; = \; 0 \tag{1.26}$$

$$\nabla \cdot \vec{B} \; = \; 0 \tag{1.27}$$

$$\nabla \times \vec{B} \; - \; \tfrac{1}{c}\tfrac{\partial}{\partial t}\vec{E} \; = \; \frac{4\pi \vec{J}}{c} \tag{1.28}$$

$$\nabla \cdot \vec{E} \; = \; 4\pi\rho \tag{1.29}$$

where $\rho(\vec{r}, t)$ describes the charge density present in the field and $\vec{J}(\vec{r}, t)$ describes the charge current density. Equations (1.27) and (1.28) can be satisfied implicitly if one represents the fields through a scalar potential $V(\vec{r}, t)$ and a vector potential $\vec{A}(\vec{r}, t)$ as follows

$$\vec{B} \; = \; \nabla \times \vec{A} \tag{1.30}$$

$$\vec{E} \; = \; -\nabla V \; - \; \frac{1}{c}\frac{\partial \vec{A}}{\partial t} \; . \tag{1.31}$$

Gauge Symmetry of the Electromagnetic Field

It is well known that the relationship between fields and potentials (1.30, 1.31) allows one to transform the potentials without affecting the fields and without affecting the equations of motion (1.25) of a particle moving in the field. The transformation which leaves the fields invariant is

$$\vec{A}'(\vec{r}, t) \; = \; \vec{A}(\vec{r}, t) \; + \; \nabla K(\vec{r}, t) \tag{1.32}$$

$$V'(\vec{r}, t) \; = \; V(\vec{r}, t) \; - \; \frac{1}{c}\frac{\partial}{\partial t}K(\vec{r}, t) \tag{1.33}$$

Lagrangian of Particle Moving in Electromagnetic Field

We want to show now that the equation of motion (1.25) follows from the Hamiltonian Principle of Least Action, if one assumes for a particle the Lagrangian

$$L(\vec{r},\dot{\vec{r}},t) \;=\; \frac{1}{2}m\vec{v}^2 \;-\; q\,V(\vec{r},t) \;+\; \frac{q}{c}\,\vec{A}(\vec{r},t)\cdot\vec{v}\,. \tag{1.34}$$

For this purpose we consider only one component of the equation of motion (1.25), namely,

$$\frac{d}{dt}(mv_1) \;=\; F_1 \;=\; -q\,\frac{\partial V}{\partial x_1} \;+\; \frac{q}{c}\,[\vec{v}\times\vec{B}]_1\,. \tag{1.35}$$

We notice using (1.30), e.g., $B_3 = \frac{\partial A_2}{\partial x_1} - \frac{\partial A_1}{\partial x_2}$

$$[\vec{v}\times\vec{B}]_1 \;=\; \dot{x}_2B_3 \;-\; \dot{x}_3B_2 \;=\; \dot{x}_2\left(\frac{\partial A_2}{\partial x_1} \;-\; \frac{\partial A_1}{\partial x_2}\right) \;-\; \dot{x}_3\left(\frac{\partial A_1}{\partial x_3} \;-\; \frac{\partial A_3}{\partial x_1}\right)\,. \tag{1.36}$$

This expression allows us to show that (1.35) is equivalent to the Euler–Lagrange condition

$$\frac{d}{dt}\left(\frac{\partial L}{\partial\dot{x}_1}\right) \;-\; \frac{\partial L}{\partial x_1} \;=\; 0\,. \tag{1.37}$$

The second term in (1.37) is

$$\frac{\partial L}{\partial x_1} \;=\; -q\,\frac{\partial V}{\partial x_1} \;+\; \frac{q}{c}\left(\frac{\partial A_1}{\partial x_1}\,\dot{x}_1 \;+\; \frac{\partial A_2}{\partial x_1}\,\dot{x}_2 \;+\; \frac{\partial A_3}{\partial x_1}\,\dot{x}_3\right)\,. \tag{1.38}$$

The first term in (1.37) is

$$\frac{d}{dt}\left(\frac{\partial L}{\partial\dot{x}_1}\right) \;=\; \frac{d}{dt}(m\dot{x}_1) \;+\; \frac{q}{c}\,\frac{dA_1}{dt} \;=\; \frac{d}{dt}(m\dot{x}_1) \;+\; \frac{q}{c}\left(\frac{\partial A_1}{\partial x_1}\,\dot{x}_1 + \frac{\partial A_1}{\partial x_2}\,\dot{x}_2 + \frac{\partial A_1}{\partial x_3}\,\dot{x}_3\right)\,. \tag{1.39}$$

The results (1.38, 1.39) together yield

$$\frac{d}{dt}(m\dot{x}_1) \;=\; -q\,\frac{\partial V}{\partial x_1} \;+\; \frac{q}{c}\,O \tag{1.40}$$

where

$$\begin{aligned} O \;&=\; \frac{\partial A_1}{\partial x_1}\,\dot{x}_1 \;+\; \frac{\partial A_2}{\partial x_1}\,\dot{x}_2 \;+\; \frac{\partial A_3}{\partial x_1}\,\dot{x}_3 \;-\; \frac{\partial A_1}{\partial x_1}\,\dot{x}_1 \;-\; \frac{\partial A_1}{\partial x_2}\,\dot{x}_2 \;-\; \frac{\partial A_1}{\partial x_3}\,\dot{x}_3 \\ &=\; \dot{x}_2\left(\frac{\partial A_2}{\partial x_1} \;-\; \frac{\partial A_1}{\partial x_2}\right) \;-\; \dot{x}_3\left(\frac{\partial A_1}{\partial x_3} \;-\; \frac{\partial A_3}{\partial x_1}\right) \end{aligned} \tag{1.41}$$

which is identical to the term (1.36) in the Newtonian equation of motion. Comparing then (1.40, 1.41) with (1.35) shows that the Newtonian equations of motion and the Euler–Lagrange conditions are, in fact, equivalent.

1.3 Symmetry Properties in Lagrangian Mechanics

Symmetry properties play an eminent role in Quantum Mechanics since they reflect the properties of the elementary constituents of physical systems, and since these properties allow one often to simplify mathematical descriptions.
We will consider in the following two symmetries, gauge symmetry and symmetries with respect to spatial transformations.
The gauge symmetry, encountered above in connection with the transformations (1.32, 1.33) of electromagnetic potentials, appear in a different, surprisingly simple fashion in Lagrangian Mechanics. They are the subject of the following theorem.
Theorem: *Gauge Transformation of Lagrangian*
The equation of motion (Euler–Lagrange conditions) of a classical mechanical system are unaffected by the following transformation of its Lagrangian

$$L'(\vec{q},\dot{\vec{q}},t) \;=\; L(\vec{q},\dot{\vec{q}},t) \;+\; \frac{d}{dt}\frac{q}{c}K(\vec{q},t) \tag{1.42}$$

This transformation is termed gauge transformation. The factor $\frac{q}{c}$ has been introduced to make this transformation equivalent to the gauge transformation (1.32, 1.33) of electyromagnetic potentials. Note that one adds the *total* time derivative of a function $K(\vec{r},t)$ the Lagrangian. This term is

$$\frac{d}{dt}K(\vec{r},t) \;=\; \frac{\partial K}{\partial x_1}\dot{x}_1 \;+\; \frac{\partial K}{\partial x_2}\dot{x}_2 \;+\; \frac{\partial K}{\partial x_3}\dot{x}_3 \;+\; \frac{\partial K}{\partial t} \;=\; (\nabla K)\cdot\vec{v} \;+\; \frac{\partial K}{\partial t}\,. \tag{1.43}$$

To prove this theorem we determine the action integral corresponding to the transformed Lagrangian

$$\begin{aligned} S'[\vec{q}(t)] \;&=\; \int_{t_0}^{t_1} dt L'(\vec{q},\dot{\vec{q}},t) \;=\; \int_{t_0}^{t_1} dt L(\vec{q},\dot{\vec{q}},t) \;+\; \frac{q}{c}K(\vec{q},t)\Big|_{t_0}^{t_1} \\ &=\; S[\vec{q}(t)] \;+\; \frac{q}{c}K(\vec{q},t)\Big|_{t_0}^{t_1} \end{aligned} \tag{1.44}$$

Since the condition $\delta\vec{q}(t_0) \;=\; \delta\vec{q}(t_1) \;=\; 0$ holds for the variational functions of Lagrangian Mechanics, Eq. (1.44) implies that the gauge transformation amounts to adding a constant term to the action integral, i.e., a term not affected by the variations allowed. One can conclude then immediately that any extremal of $S'[\vec{q}(t)]$ is also an extremal of $S[\vec{q}(t)]$.
We want to demonstrate now that the transformation (1.42) is, in fact, equivalent to the gauge transformation (1.32, 1.33) of electromagnetic potentials. For this purpose we consider the transformation of the single particle Lagrangian (1.34)

$$L'(\vec{r},\dot{\vec{r}},t) \;=\; \frac{1}{2}m\vec{v}^2 \;-\; qV(\vec{r},t) \;+\; \frac{q}{c}\vec{A}(\vec{r},t)\cdot\vec{v} \;+\; \frac{q}{c}\frac{d}{dt}K(\vec{r},t)\,. \tag{1.45}$$

Inserting (1.43) into (1.45) and reordering terms yields using (1.32, 1.33)

$$\begin{aligned} L'(\vec{r},\dot{\vec{r}},t) \;&=\; \frac{1}{2}m\vec{v}^2 \;-\; q\left(V(\vec{r},t) \;-\; \frac{1}{c}\frac{\partial K}{\partial t}\right) \;+\; \frac{q}{c}\left(\vec{A}(\vec{r},t) \;+\; \nabla K\right)\cdot\vec{v} \\ &=\; \frac{1}{2}m\vec{v}^2 \;-\; qV'(\vec{r},t) \;+\; \frac{q}{c}\vec{A}'(\vec{r},t)\cdot\vec{v}\,. \end{aligned} \tag{1.46}$$

Obviously, the transformation (1.42) corresponds to replacing in the Lagrangian potentials $V(\vec{r},t)$, $\vec{A}(\vec{r},t)$ by gauge transformed potentials $V'(\vec{r},t)$, $\vec{A}'(\vec{r},t)$. We have proven, therefore, the equivalence of (1.42) and (1.32, 1.33).
We consider now invariance properties connected with coordinate transformations. Such invariance properties are very familiar, for example, in the case of central force fields which are invariant with respect to rotations of coordinates around the center.
The following description of spatial symmetry is important in two respects, for the connection between invariance properties and constants of motion, which has an important analogy in Quantum Mechanics, and for the introduction of infinitesimal transformations which will provide a crucial method for the study of symmetry in Quantum Mechanics. The transformations we consider are the most simple kind, the reason being that our interest lies in achieving familiarity with the principles (just mentioned above) of symmetry properties rather than in providing a general tool in the context of Classical Mechanics. The transformations considered are specified in the following definition.
Definition: Infinitesimal One-Parameter Coordinate Transformations
A *one-parameter coordinate transformation* is decribed through

$$\vec{r}' \;=\; \vec{r}'(\vec{r},\epsilon)\;,\; \vec{r},\,\vec{r}' \;\in\; \mathbb{R}^{\not 4}\;,\; \epsilon \;\in\; \mathbb{R} \tag{1.47}$$

where the origin of ϵ is chosen such that

$$\vec{r}'(\vec{r},0) \;=\; \vec{r}\,. \tag{1.48}$$

The corresponding *infinitesimal transformation* is defined for small ϵ through

$$\vec{r}'(\vec{r},\epsilon) \;=\; \vec{r} \;+\; \epsilon\,\vec{R}(\vec{r}) \;+\; O(\epsilon^2)\;; \quad \vec{R}(\vec{r}) \;=\; \left.\frac{\partial \vec{r}'}{\partial \epsilon}\right|_{\epsilon=0} \tag{1.49}$$

In the following we will denote *unit vectors* as $\hat{a}$, i.e., for such vectors holds $\hat{a}\cdot\hat{a} \;=\; 1$.

Examples of Infinitesimal Transformations

The beauty of infinitesimal transformations is that they can be stated in a very simple manner. In case of a *translation transformation* in the direction $\hat{e}$ nothing new is gained. However, we like to provide the transformation here anyway for later reference

$$\vec{r}' \;=\; \vec{r} \;+\; \epsilon\,\hat{e}\,. \tag{1.50}$$

A non-trivial example is furnished by the infinitesimal rotation around axis $\hat{e}$

$$\vec{r}' \;=\; \vec{r} \;+\; \epsilon\,\hat{e}\times\vec{r}\,. \tag{1.51}$$

We would like to derive this transformation in a somewhat complicated, but nevertheless instructive way considering rotations around the x_3–axis. In this case the transformation can be written in matrix form

$$\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} \;=\; \begin{pmatrix} \cos\epsilon & -\sin\epsilon & 0 \\ \sin\epsilon & \cos\epsilon & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \tag{1.52}$$

In case of small ϵ this transformation can be written neglecting terms $O(\epsilon^2)$ using $\cos\epsilon = 1 + O(\epsilon^2)$, $\sin\epsilon = \epsilon + O(\epsilon^2)$

$$\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} 0 & -\epsilon & 0 \\ \epsilon & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + O(\epsilon^2)\,. \tag{1.53}$$

One can readily verify that in case $\hat{e} = \hat{e}_3$ ($\hat{e}_j$ denoting the unit vector in the direction of the x_j–axis) (1.51) reads

$$\vec{r}' = \vec{r} - x_2\,\hat{e}_1 + x_1\,\hat{e}_2 \tag{1.54}$$

which is equivalent to (1.53).

Anytime, a classical mechanical system is invariant with respect to a coordinate transformation a constant of motion exists, i.e., a quantity $C(\vec{r}, \dot{\vec{r}})$ which is constant along the classical path of the system. We have used here the notation corresponding to single particle motion, however, the property holds for any system.

The property has been shown to hold in a more general context, namely for fields rather than only for particle motion, by Noether. We consider here only the 'particle version' of the theorem. Before the embark on this theorem we will comment on what is meant by the statement that a classical mechanical system is invariant under a coordinate transformation. In the context of Lagrangian Mechanics this implies that such transformation leaves the Lagrangian of the system unchanged.

Theorem: *Noether's Theorem*

If $L(\vec{q}, \dot{\vec{q}}, t)$ is invariant with respect to an infinitesimal transformation $\vec{q}' = \vec{q} + \epsilon\,\vec{Q}(\vec{q})$, then $\sum_{j=1}^{M} Q_j \frac{\partial L}{\partial \dot{x}_j}$ is a constant of motion.

We have generalized in this theorem the definition of infinitesimal coordinate transformation to M–dimensional vectors $\vec{q}$.

In order to prove Noether's theorem we note

$$q_j' = q_j + \epsilon\,Q_j(\vec{q}) \tag{1.55}$$

$$\dot{q}_j' = \dot{q}_j + \epsilon \sum_{k=1}^{M} \frac{\partial Q_j}{\partial q_k}\,\dot{q}_k\,. \tag{1.56}$$

Inserting these infinitesimal changes of q_j and $\dot{q}_j$ into the Lagrangian $L(\vec{q}, \dot{\vec{q}}, t)$ yields after Taylor expansion, neglecting terms of order $O(\epsilon^2)$,

$$L'(\vec{q}, \dot{\vec{q}}, t) = L(\vec{q}, \dot{\vec{q}}, t) + \epsilon \sum_{j=1}^{M} \frac{\partial L}{\partial q_j} Q_j + \epsilon \sum_{j,k=1}^{M} \frac{\partial L}{\partial \dot{q}_j} \frac{\partial Q_j}{\partial q_k} \dot{q}_k \tag{1.57}$$

where we used $\frac{d}{dt} Q_j = \sum_{k=1}^{M} (\frac{\partial}{\partial q_k} Q_j) \dot{q}_k$. Invariance implies $L' = L$, i.e., the second and third term in (1.57) must cancel each other or both vanish. Using the fact, that *along the classical path* holds the Euler-Lagrange condition $\frac{\partial L}{\partial q_j} = \frac{d}{dt}(\frac{\partial L}{\partial \dot{q}_j})$ one can rewrite the sum of the second and third term in (1.57)

$$\sum_{j=1}^{M} \left(Q_j \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_j} \right) + \frac{\partial L}{\partial \dot{q}_j} \frac{d}{dt} Q_j \right) = \frac{d}{dt} \left(\sum_{j=1}^{M} Q_j \frac{\partial L}{\partial \dot{q}_j} \right) = 0 \tag{1.58}$$

From this follows the statement of the theorem.

Application of Noether's Theorem

We consider briefly two examples of invariances with respect to coordinate transformations for the Lagrangian $L(\vec{r},\vec{v}) = \frac{1}{2}m\vec{v}^2 - U(\vec{r})$.
We first determine the constant of motion in case of invariance with respect to translations as defined in (1.50). In this case we have $Q_j = \hat{e}_j \cdot \hat{e}$, $j = 1,2,3$ and, hence, Noether's theorem yields the constant of motion ($q_j = x_j$, $j = 1,2,3$)

$$\sum_{j=1}^{3} Q_j \frac{\partial L}{\partial \dot{x}_j} = \hat{e} \cdot \sum_{j=1}^{3} \hat{e}_j m\dot{x}_j = \hat{e} \cdot m\vec{v}\,. \tag{1.59}$$

We obtain the well known result that in this case the momentum in the direction, for which translational invariance holds, is conserved.
We will now investigate the consequence of rotational invariance as described according to the infinitesimal transformation (1.51). In this case we will use the same notation as in (1.59), except using now $Q_j = \hat{e}_j \cdot (\hat{e} \times \vec{r})$. A calculation similar to that in (1.59) yields the constant of motion $(\hat{e} \times \vec{r}) \cdot m\vec{v}$. Using the cyclic property $(\vec{a} \times \vec{b}) \cdot \vec{c} = (\vec{b} \times \vec{c}) \cdot \vec{a} = (\vec{c} \times \vec{a}) \cdot \vec{b}$ allows one to rewrite the constant of motion $\hat{e} \cdot (\vec{r} \times m\vec{v})$ which can be identified as the component of the angular momentum $m\vec{r} \times \vec{v}$ in the $\hat{e}$ direction. It was, of course, to be expected that this is the constant of motion.
The important result to be remembered for later considerations of symmetry transformations in the context of Quantum Mechanics is that it is sufficient to know the consequences of infinitesimal transformations to predict the symmetry properties of Classical Mechanics. It is not necessary to investigate the consequences of global. i.e, not infinitesimal transformations.

Chapter 2

Quantum Mechanical Path Integral

2.1 The Double Slit Experiment

Will be supplied at later date

2.2 Axioms for Quantum Mechanical Description of Single Particle

Let us consider a particle which is described by a Lagrangian $L(\vec{r}, \dot{\vec{r}}, t)$. We provide now a set of formal rules which state how the probability to observe such a particle at some space–time point $\vec{r}$, t is described in Quantum Mechanics.

1. The particle is described by a wave function $\psi(\vec{r}, t)$

$$\psi : \mathbb{R}^3 \otimes \mathbb{R} \rightarrow \mathbb{C}. \tag{2.1}$$

2. The probability that the particle is detected at space–time point $\vec{r}, t$ is

$$|\psi(\vec{r}, t)|^2 = \overline{\psi(\vec{r}, t)}\psi(\vec{r}, t) \tag{2.2}$$

   where $\overline{z}$ is the conjugate complex of z.

3. The probability to detect the particle with a detector of sensitivity $f(\vec{r})$ is

$$\int_\Omega d^3r\, f(\vec{r})\, |\psi(\vec{r}, t)|^2 \tag{2.3}$$

   where Ω is the space volume in which the particle can exist. At present one may think of $f(\vec{r})$ as a sum over δ–functions which represent a multi–slit screen, placed into the space at some particular time and with a detector behind each slit.

4. The wave function $\psi(\vec{r}, t)$ is normalized

$$\int_\Omega d^3r\, |\psi(\vec{r}, t)|^2 = 1 \qquad \forall t,\, t \in [t_0, t_1]\,, \tag{2.4}$$

   a condition which enforces that the probability of finding the particle somewhere in Ω at any particular time t in an interval $[t_0, t_1]$ in which the particle is known to exist, is unity.

5. The time evolution of $\psi(\vec{r},t)$ is described by a linear map of the type

$$\psi(\vec{r},t) \;=\; \int_\Omega d^3r'\,\phi(\vec{r},t|\vec{r}',t')\,\psi(\vec{r}',t') \qquad t\,>\,t',\; t,t'\,\in\,[t_0,t_1] \tag{2.5}$$

6. Since (2.4) holds for all times, the propagator is unitary, i.e., ($t\,>\,t',\; t,t'\,\in\,[t_0,t_1]$)

$$\begin{gathered}\textstyle\int_\Omega d^3r\,|\psi(\vec{r},t)|^2 \;=\\ \textstyle\int_\Omega d^3r\int_\Omega d^3r'\int_\Omega d^3r''\;\phi(\vec{r},t|\vec{r}',t')\;\overline{\phi(\vec{r},t|\vec{r}'',t')}\;\psi(\vec{r}',t')\;\overline{\psi(\vec{r}'',t')}\\ \textstyle=\;\int_\Omega d^3r\,|\psi(\vec{r},t')|^2\;=\;1\;.\end{gathered} \tag{2.6}$$

This must hold for any $\psi(\vec{r}',t')$ which requires

$$\int_\Omega d^3r'\,\phi(\vec{r},t|\vec{r}',t')\overline{\phi(\vec{r},t|\vec{r}'',t')}\;=\;\delta(\vec{r}'\,-\,\vec{r}'') \tag{2.7}$$

7. The following so-called completeness relationship holds for the propagator ($t\,>\,t' \quad t,t'\,\in\,[t_0,t_1]$)

$$\int_\Omega d^3r\,\phi(\vec{r},t|\vec{r}',t')\;\phi(\vec{r}',t'|\vec{r}_0,t_0)\;=\;\phi(\vec{r},t|\vec{r}_0,t_0) \tag{2.8}$$

This relationship has the following interpretation: Assume that at time t_0 a particle is generated by a source at one point $\vec{r}_0$ in space, i.e., $\psi(\vec{r}_0,t_0)\,=\,\delta(\vec{r}\,-\,\vec{r}_0)$. The state of a system at time t, described by $\psi(\vec{r},t)$, requires then according to (2.8) a knowledge of the state at all space points $\vec{r}'\,\in\Omega$ at some intermediate time t'. This is different from the classical situation where the particle follows a discrete path and, hence, at any intermediate time the particle needs only be known at one space point, namely the point on the classical path at time t'.

8. The generalization of the completeness property to $N-1$ intermediate points $t\,>\,t_{N-1}\,>\,t_{N-2}\,>\,\ldots\,>t_1\,>t_0$ is

$$\begin{gathered}\phi(\vec{r},t|\vec{r}_0,t_0)\;=\;\textstyle\int_\Omega d^3r_{N-1}\int_\Omega d^3r_{N-2}\cdots\int_\Omega d^3r_1\\ \phi(\vec{r},t_|\vec{r}_{N-1},t_{N-1})\;\phi(\vec{r}_{N-1},t_{N-1}|\vec{r}_{N-2},t_{N-2})\cdots\phi(\vec{r}_1,t_1|\vec{r}_0,t_0)\quad .\end{gathered} \tag{2.9}$$

Employing a continuum of intermediate times $t'\,\in\,[t_0,t_1]$ yields an expression of the form

$$\phi(\vec{r},t|\vec{r}_0,t_0)\;=\;\iint_{\vec{r}(t_0)=\vec{r}_0}^{\vec{r}(t_N)=\vec{r}_N} d[\vec{r}(t)]\;\Phi[\vec{r}(t)]\;. \tag{2.10}$$

We have introduced here a new symbol, the path integral

$$\iint_{\vec{r}(t_0)=\vec{r}_0}^{\vec{r}(t_N)=\vec{r}_N} d[\vec{r}(t)]\;\cdots \tag{2.11}$$

which denotes an integral over all paths $\vec{r}(t)$ with end points $\vec{r}(t_0)\,=\,\vec{r}_0$ and $\vec{r}(t_N)\,=\,\vec{r}_N$. This symbol will be defined further below. The definition will actually assume an infinite number of intermediate times and express the path integral through integrals of the type (2.9) for $N\,\to\,\infty$.

9. The functional $\Phi[\vec{r}(t)]$ in (2.11) is

$$\Phi[\vec{r}(t)] \;=\; \exp\left\{ \frac{i}{\hbar}\, S[\vec{r}(t)] \right\} \tag{2.12}$$

where $S[\vec{r}(t)]$ is the classical action integral

$$S[\vec{r}(t)] \;=\; \int_{t_0}^{t_N} dt\, L(\vec{r}, \dot{\vec{r}}, t) \tag{2.13}$$

and

$$\hbar \;=\; 1.0545 \cdot 10^{-27}\, \mathrm{erg\ s}\,. \tag{2.14}$$

In (2.13) $L(\vec{r}, \dot{\vec{r}}, t)$ is the Lagrangian of the classical particle. However, in complete distinction from Classical Mechanics, expressions (2.12, 2.13) are built on action integrals for all possible paths, not only for the classical path. Situations which are well described classically will be distinguished through the property that the classical path gives the dominant, actually often essentially exclusive, contribution to the path integral (2.12, 2.13). However, for microscopic particles like the electron this is by no means the case, i.e., for the electron many paths contribute and the action integrals for non-classical paths need to be known.

The constant $\hbar$ given in (2.14) has the same dimension as the action integral $S[\vec{r}(t)]$. Its value is extremely small in comparision with typical values for action integrals of macroscopic particles. However, it is comparable to action integrals as they arise for microscopic particles under typical circumstances. To show this we consider the value of the action integral for a particle of mass $m = 1$ g moving over a distance of 1 cm/s in a time period of 1 s. The value of $S[\vec{r}(t)]$ is then

$$S_{cl} \;=\; \frac{1}{2}\, m\, v^2\, t \;=\; \frac{1}{2}\mathrm{erg\,s}\,. \tag{2.15}$$

The exponent of (2.12) is then $S_{cl}/\hbar \approx 0.5 \cdot 10^{27}$, i.e., a very large number. Since this number is multiplied by 'i', the exponent is a very large imaginary number. Any variations of S_{cl} would then lead to strong oscillations of the contributions $\exp(\frac{i}{\hbar}S)$ to the path integral and one can expect destructive interference betwen these contributions. Only for paths close to the classical path is such interference ruled out, namely due to the property of the classical path to be an extremal of the action integral. This implies that small variations of the path near the classical path alter the value of the action integral by very little, such that destructive interference of the contributions of such paths does not occur.
The situation is very different for microscopic particles. In case of a proton with mass $m = 1.6725 \cdot 10^{-24}$ g moving over a distance of 1 Å in a time period of 10^{-14} s the value of $S[\vec{r}(t)]$ is $S_{cl} \approx 10^{-26}$ erg s and, accordingly, $S_{cl}/\hbar \approx 8$. This number is much smaller than the one for the macroscopic particle considered above and one expects that variations of the exponent of $\Phi[\vec{r}(t)]$ are of the order of unity for protons. One would still expect significant descructive interference between contributions of different paths since the value calculated is comparable to 2π. However, interferences should be much less dramatic than in case of the macroscopic particle.

2.3 How to Evaluate the Path Integral

In this section we will provide an explicit algorithm which defines the path integral (2.12, 2.13) and, at the same time, provides an avenue to evaluate path integrals. For the sake of simplicity we will consider the case of particles moving in one dimension labelled by the position coordinate x. The particles have associated with them a Lagrangian

$$L(x, \dot{x}, t) \; = \; \frac{1}{2} m \dot{x}^2 \; - \; U(x) \; . \tag{2.16}$$

In order to define the path integral we assume, as in (2.9), a series of times $t_N > t_{N-1} > t_{N-2} > \ldots > t_1 > t_0$ letting N go to infinity later. The spacings between the times t_{j+1} and t_j will all be identical, namely

$$t_{j+1} \; - \; t_j \; = \; (t_N \; - \; t_0)/N \; = \; \epsilon_N \; . \tag{2.17}$$

The discretization in time leads to a discretization of the paths $x(t)$ which will be represented through the series of space–time points

$$\{(x_0, t_0), \, (x_1, t_1), \, \ldots (x_{N-1}, t_{N-1}), \, (x_N, t_N)\} \; . \tag{2.18}$$

The time instances are fixed, however, the x_j values are not. They can be anywhere in the allowed volume which we will choose to be the interval $]-\infty, \infty[$. In passing from one space–time instance (x_j, t_j) to the next (x_{j+1}, t_{j+1}) we assume that kinetic energy and potential energy are constant, namely $\frac{1}{2} m (x_{j+1} - x_j)^2 / \epsilon_N^2$ and $U(x_j)$, respectively. These assumptions lead then to the following Riemann form for the action integral

$$S[x(t)] \; = \; \lim_{N \to \infty} \epsilon_N \sum_{j=0}^{N-1} \left(\frac{1}{2} m \frac{(x_{j+1} - x_j)^2}{\epsilon_N^2} \; - \; U(x_j) \right) \; . \tag{2.19}$$

The main idea is that one can replace the path integral now by a multiple integral over x_1, x_2, etc. This allows us to write the evolution operator using (2.10) and (2.12)

$$\begin{aligned} \phi(x_N, t_N | x_0, t_0) \; = \; & \lim_{N \to \infty} C_N \int_{-\infty}^{+\infty} dx_1 \int_{-\infty}^{+\infty} dx_2 \ldots \int_{-\infty}^{+\infty} dx_{N-1} \\ & \exp \left\{ \frac{i}{\hbar} \epsilon_N \sum_{j=0}^{N-1} \left[\frac{1}{2} m \frac{(x_{j+1} - x_j)^2}{\epsilon_N^2} \; - \; U(x_j) \right] \right\} \; . \end{aligned} \tag{2.20}$$

Here, C_N is a constant which depends on N (actually also on other constant in the exponent) which needs to be chosen to ascertain that the limit in (2.20) can be properly taken. Its value is

$$C_N \; = \; \left[\frac{m}{2 \pi i \hbar \epsilon_N} \right]^{\frac{N}{2}} \tag{2.21}$$

2.4 Propagator for a Free Particle

As a first example we will evaluate the path integral for a free particle following the algorithm introduced above.

Rather then using the integration variables x_j, it is more suitable to define new integration variables y_j, the origin of which coincides with the classical path of the particle. To see the benifit of such approach we define a path $y(t)$ as follows

$$x(t) \;=\; x_{cl}(t) \;+\; y(t) \tag{2.22}$$

where $x_{cl}(t)$ is the classical path which connects the space–time points (x_0, t_0) and (x_N, t_N), namely,

$$x_{cl}(t) \;=\; x_0 \;+\; \frac{x_N \;-\; x_0}{t_N \;-\; t_0}(\,t \;-\; t_0)\;. \tag{2.23}$$

It is essential for the following to note that, since $x(t_0) \;=\; x_{cl}(t_0) \;=\; x_0$ and $x(t_N) \;=\; x_{cl}(t_N) \;=\; x_N$, it holds

$$y(t_0) \;=\; y(t_N) \;=\; 0\;. \tag{2.24}$$

Also we use the fact that the velocity of the classical path $\dot{x}_{cl} \;=\; (x_N - x_0)/(t_n - t_0)$ is constant. The action integral[1] $S[x(t)|x(t_0) = x_0,\, x(t_N) = x_N]$ for any path $x(t)$ can then be expressed through an action integral over the path $y(t)$ relative to the classical path. One obtains

$$\begin{aligned} S[x(t)|x(t_0) = x_0,\, x(t_N) = x_N] \;&=\; \textstyle\int_{t_0}^{t_N} dt \tfrac{1}{2} m(\dot{x}_{cl}^2 \;+\; 2\dot{x}_{cl}\dot{y} \;+\; \dot{y}^2) \;= \\ &\textstyle\int_{t_0}^{t_N} dt \tfrac{1}{2} m\dot{x}_{cl}^2 \;+\; m\dot{x}_{cl} \int_{t_0}^{t_N} dt\dot{y} \;+\; \int_{t_0}^{t_N} dt \tfrac{1}{2} m\dot{y}^2\;. \end{aligned} \tag{2.25}$$

The condition (2.24) implies for the second term on the r.h.s.

$$\int_{t_0}^{t_N} dt\, \dot{y} \;=\; y(t_N) \;-\; y(t_0) \;=\; 0\;. \tag{2.26}$$

The first term on the r.h.s. of (2.25) is, using (2.23),

$$\int_{t_0}^{t_N} dt\, \frac{1}{2}\, m\, \dot{x}_{cl}^2 \;=\; \frac{1}{2} m \frac{(x_N \;-\; x_0)^2}{t_N \;-\; t_0}\;. \tag{2.27}$$

The third term can be written in the notation introduced

$$\int_{t_0}^{t_N} dt\, \frac{1}{2} m\dot{y}^2 \;=\; S[x(t)|x(t_0) = 0,\, x(t_N) = 0]\;, \tag{2.28}$$

i.e., due to (2.24), can be expressed through a path integral with endpoints $x(t_0) = 0$, $x(t_N) = 0$. The resulting expression for $S[x(t)|x(t_0) = x_0,\, x(t_N) = x_N]$ is

$$\begin{aligned} S[x(t)|x(t_0) = x_0,\, x(t_N) = x_N] \;&=\; \frac{1}{2} m \frac{(x_N \;-\; x_0)^2}{t_N \;-\; t_0} \;+\; 0 \;+ \\ &+\; S[x(t)|x(t_0) = 0,\, x(t_N) = 0]\;. \end{aligned} \tag{2.29}$$

This expression corresponds to the action integral in (2.13). Inserting the result into (2.10, 2.12) yields

$$\phi(x_N, t_N|x_0, t_0) \;=\; \exp\left[\frac{im}{2\hbar}\frac{(x_N - x_0)^2}{t_N - t_0}\right] \iint_{x(t_0)=0}^{x(t_N)=0} d[x(t)] \exp\left\{\frac{i}{\hbar}\, S[x(t)]\right\} \tag{2.30}$$

a result, which can also be written

$$\phi(x_N, t_N|x_0, t_0) \;=\; \exp\left[\frac{im}{2\hbar}\frac{(x_N - x_0)^2}{t_N - t_0}\right] \phi(0, t_N|0, t_0) \tag{2.31}$$

[1] We have denoted explicitly that the action integral for a path connecting the space–time points (x_0, t_0) and (x_N, t_N) is to be evaluated.

Evaluation of the necessary path integral

To determine the propagator (2.31) for a free particle one needs to evaluate the following path integral

$$\begin{aligned} \phi(0,t_N|0,t_0) \;&=\; lim_{N\to\infty}\left[\tfrac{m}{2\pi i\hbar\epsilon_N}\right]^{\frac{N}{2}} \times \\ \times\, \textstyle\int_{-\infty}^{+\infty}dy_1\cdots\int_{-\infty}^{+\infty}dy_{N-1}\exp\left[\tfrac{i}{\hbar}\,\epsilon_N\,\sum_{j=0}^{N-1}\,\tfrac{1}{2}m\tfrac{(y_{j+1}-y_j)^2}{\epsilon_N^2}\right] \end{aligned} \tag{2.32}$$

The exponent E can be written, noting $y_0 \,=\, y_N \,=0$, as the quadratic form

$$\begin{aligned} E \;&=\; \frac{im}{2\hbar\epsilon_N}\,(\,2y_1^2 \,-\, y_1y_2 \,-\, y_2y_1 \,+\, 2y_2^2 \,-\, y_2y_3 \,-\, y_3y_2 \\ &\quad +2y_3^2 \,-\cdots-\, y_{N-2}y_{N-1} \,-\, y_{N-1}y_{N-2} \,+\, 2y_{N-1}^2\,) \\ &=\; i\sum_{j,k=1}^{N-1}\, y_j\,a_{jk}\,y_k \end{aligned} \tag{2.33}$$

where a_{jk} are the elements of the following symmetric $(N-1)\times(N-1)$ matrix

$$\left(\,a_{jk}\,\right) \;=\; \frac{m}{2\hbar\epsilon_N}\begin{pmatrix} 2 & -1 & 0 & \dots & 0 & 0 \\ -1 & 2 & -1 & \dots & 0 & 0 \\ 0 & -1 & 2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 2 & -1 \\ 0 & 0 & 0 & & -1 & 2 \end{pmatrix} \tag{2.34}$$

The following integral

$$\int_{-\infty}^{+\infty}dy_1\cdots\int_{-\infty}^{+\infty}dy_{N-1}\exp\left(i\sum_{j,k=1}^{N-1}\,y_ja_{jk}y_k\right) \tag{2.35}$$

must be determined. In the appendix we prove

$$\int_{-\infty}^{+\infty}dy_1\cdots\int_{-\infty}^{+\infty}dy_{N-1}\exp\left(i\sum_{j,k=1}^{d}\,y_jb_{jk}y_k\right) \;=\; \left[\frac{(i\pi)^d}{\det(b_{jk})}\right]^{\frac{1}{2}}\;. \tag{2.36}$$

which holds for a d-dimensional, real, symmetric matrix (b_{jk}) and $\det(b_{jk}) \,\neq 0$.
In order to complete the evaluation of (2.32) we split off the factor $\frac{m}{2\hbar\epsilon_N}$ in the definition (2.34) of (a_{jk}) defining a new matrix (A_{jk}) through

$$a_{jk} \;=\; \frac{m}{2\hbar\epsilon_N}\,A_{jk}\;. \tag{2.37}$$

Using

$$\det(a_{jk}) \;=\; \left[\frac{m}{2\hbar\epsilon_N}\right]^{N-1}\det(A_{jk})\;, \tag{2.38}$$

a property which follows from $\det(c\mathbf{B}) = c^n \det\mathbf{B}$ for any $n \times n$ matrix $\mathbf{B}$, we obtain

$$\phi(0,t_N|0,t_0) \;=\; lim_{N\to\infty} \left[\frac{m}{2\pi i\hbar\epsilon_N}\right]^{\frac{N}{2}} \left[\frac{2\pi i\hbar\epsilon_N}{m}\right]^{\frac{N-1}{2}} \frac{1}{\sqrt{\det(A_{jk})}}\,. \tag{2.39}$$

In order to determine $\det(A_{jk})$ we consider the dimension n of (A_{jk}), presently $N-1$, variable, let say n, $n = 1, 2, \ldots$. We seek then to evaluate the determinant of the $n \times n$ matrix

$$D_n \;=\; \left| \begin{pmatrix} 2 & -1 & 0 & \ldots & 0 & 0 \\ -1 & 2 & -1 & \ldots & 0 & 0 \\ 0 & -1 & 2 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 2 & -1 \\ 0 & 0 & 0 & & -1 & 2 \end{pmatrix} \right| . \tag{2.40}$$

For this purpose we expand (2.40) in terms of subdeterminants along the last column. One can readily verify that this procedure leads to the following recursion equation for the determinants

$$D_n \;=\; 2\,D_{n-1} \;-\; D_{n-2}\,. \tag{2.41}$$

To solve this three term recursion relationship one needs two starting values. Using

$$D_1 \;=\; |(2)| \;=\; 2\,; \quad D_2 \;=\; \left| \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \right| \;=\; 3 \tag{2.42}$$

one can readily verify

$$D_n \;=\; n+1\,. \tag{2.43}$$

We like to note here for further use below that one might as well employ the 'artificial' starting values $D_0 = 1$, $D_1 = 2$ and obtain from (2.41) the same result for D_2, $D_3, \ldots$.
Our derivation has provided us with the value $\det(A_{jk}) = N$. Inserting this into (2.39) yields

$$\phi(0,t_N|0,t_0) \;=\; lim_{N\to\infty} \left[\frac{m}{2\pi i\hbar\epsilon_N N}\right]^{\frac{1}{2}} \tag{2.44}$$

and with $\epsilon_N N = t_N - t_0$, which follows from (2.18) we obtain

$$\phi(0,t_N|0,t_0) \;=\; \left[\frac{m}{2\pi i\hbar(t_N - t_0)}\right]^{\frac{1}{2}}\,. \tag{2.45}$$

Expressions for Free Particle Propagator

We have now collected all pieces for the final expression of the propagator (2.31) and obtain, defining $t = t_N$, $x = x_N$

$$\phi(x,t|x_0,t_0) \;=\; \left[\frac{m}{2\pi i\hbar(t-t_0)}\right]^{\frac{1}{2}} \exp\left[\frac{im}{2\hbar}\frac{(x-x_0)^2}{t-t_0}\right]\,. \tag{2.46}$$

This propagator, according to (2.5) allows us to predict the time evolution of any state function $\psi(x,t)$ of a free particle. Below we will apply this to a particle at rest and a particle forming a so-called wave packet.
The result (2.46) can be generalized to three dimensions in a rather obvious way. One obtains then for the propagator (2.10)

$$\phi(\vec{r},t|\vec{r}_0,t_0) \;=\; \left[\frac{m}{2\pi i\hbar(t-t_0)}\right]^{\frac{3}{2}} \exp\left[\frac{im}{2\hbar}\frac{(\vec{r}-\vec{r}_0)^2}{t-t_0}\right] \;. \tag{2.47}$$

One-Dimensional Free Particle Described by Wave Packet

We assume a particle at time $t = t_o = 0$ is described by the wave function

$$\psi(x_0,t_0) \;=\; \left[\frac{1}{\pi\delta^2}\right]^{\frac{1}{4}} \exp\left(-\frac{x_0^2}{2\delta^2} \;+\; i\frac{p_o}{\hbar}\,x\right) \tag{2.48}$$

Obviously, the associated probability distribution

$$|\psi(x_0,t_0)|^2 \;=\; \left[\frac{1}{\pi\delta^2}\right]^{\frac{1}{2}} \exp\left(-\frac{x_0^2}{\delta^2}\right) \tag{2.49}$$

is Gaussian of width δ, centered around $x_0 = 0$, and describes a single particle since

$$\left[\frac{1}{\pi\delta^2}\right]^{\frac{1}{2}} \int_{-\infty}^{+\infty} dx_0 \exp\left(-\frac{x_0^2}{\delta^2}\right) \;=\; 1 \;. \tag{2.50}$$

One refers to such states as *wave packets*. We want to apply axiom (2.5) to (2.48) as the initial state using the propagator (2.46).
We will obtain, thereby, the wave function of the particle at later times. We need to evaluate for this purpose the integral

$$\begin{aligned}\psi(x,t) \quad &= \quad \left[\frac{1}{\pi\delta^2}\right]^{\frac{1}{4}} \left[\frac{m}{2\pi i\hbar t}\right]^{\frac{1}{2}} \\ &\qquad \int_{-\infty}^{+\infty} dx_0 \exp\underbrace{\left[\frac{im}{2\hbar}\frac{(x-x_0)^2}{t} \;-\; \frac{x_0^2}{2\delta^2} \;+\; i\frac{p_o}{\hbar}\,x_o\right]}_{E_o(x_o,x) \;+\; E(x)}\end{aligned} \tag{2.51}$$

For this evaluation we adopt the strategy of combining in the exponential the terms quadratic ($\sim x_0^2$) and linear ($\sim x_0$) in the integration variable to a complete square

$$ax_0^2 + 2bx_0 \;=\; a\left(x_0 \;+\; \frac{b}{a}\right)^2 \;-\; \frac{b^2}{a} \tag{2.52}$$

and applying (2.247).
We devide the contributions to the exponent $E_o(x_o,x) \;+\; E(x)$ in (2.51) as follows

$$E_o(x_o,x) \;=\; \frac{im}{2\hbar t}\left[x_o^2\left(1 \;+\; i\frac{\hbar t}{m\delta^2}\right) \;-\; 2x_o\left(x \;-\; \frac{p_o}{m}t\right) \;+\; f(x)\right] \tag{2.53}$$

$$E(x) \;=\; \frac{im}{2\hbar t}\,\left[\,x^2 \;-\; f(x)\,\right]\;. \tag{2.54}$$

One chooses then $f(x)$ to complete, according to (2.52), the square in (2.53)

$$f(x) \;=\; \left(\,\frac{x-\frac{p_o}{m}t}{\sqrt{1+i\frac{\hbar t}{m\delta^2}}}\,\right)^2\;. \tag{2.55}$$

This yields

$$E_o(x_o,x) \;=\; \frac{im}{2\hbar t}\,\left(\,x_o\sqrt{1+i\frac{\hbar t}{m\delta^2}}\;-\;\frac{x\;-\;\frac{p_o}{m}\,t}{\sqrt{1+i\frac{\hbar t}{m\delta^2}}}\,\right)^2\;. \tag{2.56}$$

One can write then (2.51)

$$\psi(x,t) \;=\; \left[\frac{1}{\pi\delta^2}\right]^{\frac{1}{4}}\,\left[\frac{m}{2\pi i\hbar t}\right]^{\frac{1}{2}}\,e^{E(x)}\,\int_{-\infty}^{+\infty}dx_0\,e^{E_o(x_o,x)} \tag{2.57}$$

and needs to determine the integral

$$\begin{aligned} I \;&=\; \int_{-\infty}^{+\infty}dx_0\,e^{E_o(x_o,x)} \\ &=\; \int_{-\infty}^{+\infty}dx_0\exp\left[\,\frac{im}{2\hbar t}\,\left(\,x_o\sqrt{1+i\frac{\hbar t}{m\delta^2}}\;-\;\frac{x\;-\;\frac{p_o}{m}\,t}{\sqrt{1+i\frac{\hbar t}{m\delta^2}}}\,\right)^2\,\right] \\ &=\; \int_{-\infty}^{+\infty}dx_0\exp\left[\,\frac{im}{2\hbar t}\,\left(1+i\frac{\hbar t}{m\delta^2}\right)\left(\,x_o\;-\;\frac{x\;-\;\frac{p_o}{m}\,t}{1+i\frac{\hbar t}{m\delta^2}}\,\right)^2\,\right]\;. \end{aligned} \tag{2.58}$$

The integrand is an analytical function everywhere in the complex plane and we can alter the integration path, making certain, however, that the new path does not lead to additional contributions to the integral.
We proceed as follows. We consider a transformation to a new integration variable ρ defined through

$$\sqrt{i\,\left(1\;-\;i\frac{\hbar t}{m\delta^2}\right)}\,\rho \;=\; x_0\;-\;\frac{x\;-\;\frac{p_o}{m}\,t}{1+i\frac{\hbar t}{m\delta^2}}\;. \tag{2.59}$$

An integration path in the complex x_0–plane along the direction

$$\sqrt{i\,\left(1\;-\;i\frac{\hbar t}{m\delta^2}\right)} \tag{2.60}$$

is then represented by real ρ values. The beginning and the end of such path are the points

$$z_1' \;=\; -\infty\;\times\;\sqrt{i\,\left(1\;-\;i\frac{\hbar t}{m\delta^2}\right)}\;,\quad z_2' \;=\; +\infty\;\times\;\sqrt{i\,\left(1\;-\;i\frac{\hbar t}{m\delta^2}\right)} \tag{2.61}$$

whereas the original path in (2.58) has the end points

$$z_1 \;=\; -\infty\,, \quad z_2 \;=\; +\infty\,. \tag{2.62}$$

If one can show that an integration of (2.58) along the path $z_1 \to z_1'$ and along the path $z_2 \to z_2'$ gives only vanishing contributions one can replace (2.58) by

$$I \;=\; \sqrt{i\left(1 \,-\, i\frac{\hbar t}{m\delta^2}\right)}\int_{-\infty}^{+\infty} d\rho \exp\left[-\frac{m}{2\hbar t}\left(1 \,+\, \left(\frac{\hbar t}{m\delta^2}\right)^2\right)\rho^2\right] \tag{2.63}$$

which can be readily evaluated. In fact, one can show that z_1' lies at $-\infty \,-\, i\times\infty$ and z_2' at $+\infty \,+\, i\times\infty$. Hence, the paths between $z_1 \to z_1'$ and $z_2 \to z_2'$ have a real part of x_0 of $\pm\infty$. Since the exponent in (2.58) has a leading contribution in x_0 of $-x_0^2/\delta^2$ the integrand of (2.58) vanishes for $Re\, x_0 \;\to\; \pm\infty$. We can conclude then that (2.63) holds and, accordingly,

$$I \;=\; \sqrt{\frac{2\pi i\hbar t}{m(1 \,+\, i\frac{\hbar t}{m\delta^2})}}\,. \tag{2.64}$$

Equation (2.57) reads then

$$\psi(x,t) \;=\; \left[\frac{1}{\pi\delta^2}\right]^{\frac{1}{4}}\left[\frac{1}{1 \,+\, i\frac{\hbar t}{m\delta^2}}\right]^{\frac{1}{2}} \exp\left[\,E(x)\,\right]\,. \tag{2.65}$$

Seperating the phase factor

$$\left[\frac{1 \,-\, i\frac{\hbar t}{m\delta^2}}{1 \,+\, i\frac{\hbar t}{m\delta^2}}\right]^{\frac{1}{4}}\,, \tag{2.66}$$

yields

$$\psi(x,t) \;=\; \left[\frac{1 \,-\, i\frac{\hbar t}{m\delta^2}}{1 \,+\, i\frac{\hbar t}{m\delta^2}}\right]^{\frac{1}{4}}\left[\frac{1}{\pi\delta^2\,(1 \,+\, \frac{\hbar^2 t^2}{m^2\delta^4})}\right]^{\frac{1}{4}} \exp\left[\,E(x)\,\right]\,. \tag{2.67}$$

We need to determine finally (2.54) using (2.55). One obtains

$$E(x) \;=\; -\,\frac{x^2}{2\delta^2(1 + i\frac{\hbar t}{m\delta^2})} \;+\; \frac{i\frac{p_o}{\hbar}x}{1 + i\frac{\hbar t}{m\delta^2}} \;-\; \frac{\frac{i}{\hbar}\frac{p_o^2}{2m}t}{1 + i\frac{\hbar t}{m\delta^2}} \tag{2.68}$$

and, using

$$\frac{a}{1 \,+\, b} \;=\; a \;-\; \frac{a\,b}{1 \,+\, b}\,, \tag{2.69}$$

finally

$$E(x) \;=\; -\frac{(x \,-\, \frac{p_o}{m}t)^2}{2\delta^2(1 + i\frac{\hbar t}{m\delta^2})} \;+\; i\frac{p_o}{\hbar}x \;-\; \frac{i}{\hbar}\frac{p_o^2}{2m}t \tag{2.70}$$

which inserted in (2.67) provides the complete expression of the wave function at all times t

$$\psi(x,t) = \left[\frac{1 - i\frac{\hbar t}{m\delta^2}}{1 + i\frac{\hbar t}{m\delta^2}}\right]^{\frac{1}{4}} \left[\frac{1}{\pi\delta^2\,(1 + \frac{\hbar^2 t^2}{m^2\delta^4})}\right]^{\frac{1}{4}} \times \tag{2.71}$$

$$\times \exp\left[-\frac{(x - \frac{p_o}{m}t)^2}{2\delta^2(1 + \frac{\hbar^2 t^2}{m^2\delta^4})}(1 - i\frac{\hbar t}{m\delta^2}) + i\frac{p_o}{\hbar}x - \frac{i}{\hbar}\frac{p_o^2}{2m}t\right] .$$

The corresponding probability distribution is

$$|\psi(x,t)|^2 = \left[\frac{1}{\pi\delta^2\,(1 + \frac{\hbar^2 t^2}{m^2\delta^4})}\right]^{\frac{1}{2}} \exp\left[-\frac{(x - \frac{p_o}{m}t)^2}{\delta^2(1 + \frac{\hbar^2 t^2}{m^2\delta^4})}\right] . \tag{2.72}$$

Comparision of Moving Wave Packet with Classical Motion

It is revealing to compare the probability distributions (2.49), (2.72) for the initial state (2.48) and for the final state (2.71), respectively. The center of the distribution (2.72) moves in the direction of the positive x-axis with velocity $v_o = p_o/m$ which identifies p_o as the momentum of the particle. The width of the distribution (2.72)

$$\delta\sqrt{1 + \frac{\hbar^2 t^2}{m^2\delta^4}} \tag{2.73}$$

increases with time, coinciding at $t = 0$ with the width of the initial distribution (2.49). This 'spreading' of the wave function is a genuine quantum phenomenon. Another interesting observation is that the wave function (2.71) conserves the phase factor $\exp[i(p_o/\hbar)x]$ of the original wave function (2.48) and that the respective phase factor is related with the velocity of the classical particle and of the center of the distribution (2.72). The conservation of this factor is particularly striking for the (unnormalized) initial wave function

$$\psi(x_0,t_0) = \exp\left(i\frac{p_o}{m}\,x_o\right) , \tag{2.74}$$

which corresponds to (2.48) for $\delta \to \infty$. In this case holds

$$\psi(x,t) = \exp\left(\, i\frac{p_o}{m}\,x - \frac{i}{\hbar}\frac{p_o^2}{2m}t\right) . \tag{2.75}$$

i.e., the spatial dependence of the initial state (2.74) remains invariant in time. However, a time-dependent phase factor $\exp[-\frac{i}{\hbar}(p_o^2/2m)\,t]$ arises which is related to the energy $\epsilon = p_o^2/2m$ of a particle with momentum p_o. We had assumed above [c.f. (2.48)] $t_o = 0$. the case of arbitrary t_o is recovered iby replacing $t \to t_o$ in (2.71, 2.72). This yields, instead of (2.75)

$$\psi(x,t) = \exp\left(\, i\frac{p_o}{m}\,x - \frac{i}{\hbar}\frac{p_o^2}{2m}(t - t_o)\right) . \tag{2.76}$$

From this we conclude that an initial wave function

$$\psi(x_o,t_0) = \exp\left(\, i\frac{p_o}{m}\,x_o - \frac{i}{\hbar}\frac{p_o^2}{2m}t_o\right) . \tag{2.77}$$

becomes at $t > t_o$

$$\psi(x,t) \;=\; \exp\left(i\frac{p_o}{m}\,x \;-\; \frac{i}{\hbar}\frac{p_o^2}{2m}t\right)\;, \tag{2.78}$$

i.e., the spatial as well as the temporal dependence of the wave function remains invariant in this case. One refers to the respective states as *stationary states*. Such states play a cardinal role in quantum mechanics.

2.5 Propagator for a Quadratic Lagrangian

We will now determine the propagator (2.10, 2.12, 2.13)

$$\phi(x_N,t_N|x_0,t_0) \;=\; \iint_{x(t_0)=x_0}^{x(t_N)=x_N} d[x(t)]\,\exp\left\{\frac{i}{\hbar}\,S[x(t)]\right\} \tag{2.79}$$

for a quadratic Lagrangian

$$L(x,\dot{x},t) \;=\; \frac{1}{2}m\dot{x}^2 \;-\; \frac{1}{2}c(t)x^2 \;-\; e(t)x\;. \tag{2.80}$$

For this purpose we need to determine the action integral

$$S[x(t)] \;=\; \int_{t_0}^{t_N} dt'\,L(x,\dot{x},t) \tag{2.81}$$

for an arbitrary path $x(t)$ with end points $x(t_0) = x_0$ and $x(t_N) = x_N$. In order to simplify this task we define again a new path $y(t)$

$$x(t) \;=\; x_{cl}(t) \;+\; y(t) \tag{2.82}$$

which describes the deviation from the classical path $x_{cl}(t)$ with end points $x_{cl}(t_0) = x_0$ and $x_{cl}(t_N) = x_N$. Obviously, the end points of $y(t)$ are

$$y(t_0) \;=\; 0 \quad ; \quad y(t_N) \;=\; 0\;. \tag{2.83}$$

Inserting (2.80) into (2.82) one obtains

$$L(x_{cl}+y,\dot{x}_{cl}+\dot{y}(t),t) \;=\; L(x_{cl},\dot{x}_{cl},t) \;+\; L'(y,\dot{y}(t),t) \;+\; \delta L \tag{2.84}$$

where

$$\begin{aligned} L(x_{cl},\dot{x}_{cl},t) \quad &= \; \frac{1}{2}m\dot{x}_{cl}^2 \;-\; \frac{1}{2}c(t)x_{cl}^2 \;-\; e(t)x_{cl} \\ L'(y,\dot{y}(t),t) \quad &= \; \frac{1}{2}m\dot{y}^2 \;-\; \frac{1}{2}c(t)y^2 \\ \delta L \qquad &= \; m\dot{x}_{cl}\dot{y}(t) \;-\; c(t)x_{cl}y \;-\; e(t)y\;. \end{aligned} \tag{2.85}$$

We want to show now that the contribution of δL to the action integral (2.81) vanishes[2]. For this purpose we use

$$\dot{x}_{cl}\dot{y} \;=\; \frac{d}{dt}\left(\dot{x}_{cl}\,y\right) \;-\; \ddot{x}_{cl}\,y \tag{2.86}$$

[2]The reader may want to verify that the contribution of δL to the action integral is actually equal to the differential $\delta S[x_{cl},y(t)]$ which vanishes according to the Hamiltonian principle as discussed in Sect. 1.

and obtain

$$\int_{t_0}^{t_N} dt\, \delta L \;=\; m\;[\dot{x}_{cl}\, y]|_{t_0}^{t^N} \;-\; \int_{t_0}^{t_N} dt\;[\, m\ddot{x}_{cl}(t) \;+\; c(t)\, x_{cl}(t) \;+\; e(t)\,]\; y(t)\;. \tag{2.87}$$

According to (2.83) the first term on the r.h.s. vanishes. Applying the Euler–Lagrange conditions (1.24) to the Lagrangian (2.80) yields for the classical path

$$m\ddot{x}_{cl} \;+\; c(t)\, x_{cl} \;+\; e(t) \;=\; 0 \tag{2.88}$$

and, hence, also the second contribution on the r.h.s. of (2.88) vanishes. One can then express the propagator (2.79)

$$\phi(x_N, t_N|x_0, t_0) \;=\; \exp\left\{\, \frac{i}{\hbar}\, S[x_{cl}(t)]\,\right\} \tilde{\phi}(0, t_N|0, t_0) \tag{2.89}$$

where

$$\tilde{\phi}(0, t_N|0, t_0) \;=\; \iint_{y(t_0)=0}^{y(t_N)=0} d[y(t)]\, \exp\left\{\, \frac{i}{\hbar} \int_{t_0}^{t_N} dt\, L'(y, \dot{y}, t)\,\right\}\;. \tag{2.90}$$

Evaluation of the Necessary Path Integral

We have achieved for the quadratic Lagrangian a separation in terms of a classical action integral and a propagator connecting the end points $y(t_0) \;=\; 0$ and $y(t_N) \;=\; 0$ which is analogue to the result (2.31) for the free particle propagator. For the evaluation of $\tilde{\phi}(0, t_N|0, t_0)$ we will adopt a strategy which is similar to that used for the evaluation of (2.32). The discretization scheme adopted above yields in the present case

$$\begin{aligned} \tilde{\phi}(0, t_N|0, t_0) \;&=\; lim_{N\to\infty} \left[\tfrac{m}{2\pi i\hbar\epsilon_N}\right]^{\frac{N}{2}} \;\times \\ \times\; \textstyle\int_{-\infty}^{+\infty} dy_1 \cdots \int_{-\infty}^{+\infty} dy_{N-1} \exp\left[\tfrac{i}{\hbar}\,\epsilon_N \,\textstyle\sum_{j=0}^{N-1} \left(\tfrac{1}{2} m \tfrac{(y_{j+1}-y_j)^2}{\epsilon_N^2} \;-\; \tfrac{1}{2}\, c_j\, y_j^2\right)\right] \end{aligned} \tag{2.91}$$

where $c_j \;=\; c(t_j)$, $t_j \;=\; t_0 + \epsilon_N\, j$. One can express the exponent E in (2.91) through the quadratic form

$$E \;=\; i \sum_{j,k=1}^{N-1} y_j\, a_{jk}\, y_k \tag{2.92}$$

where a_{jk} are the elements of the following $(N-1) \times (N-1)$ matrix

$$\begin{aligned} \left(\, a_{jk} \,\right) \;&=\; \frac{m}{2\hbar\epsilon_N} \begin{pmatrix} 2 & -1 & 0 & \dots & 0 & 0 \\ -1 & 2 & -1 & \dots & 0 & 0 \\ 0 & -1 & 2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 2 & -1 \\ 0 & 0 & 0 & & -1 & 2 \end{pmatrix} \\ &-\; \frac{\epsilon_N}{2\hbar} \begin{pmatrix} c_1 & 0 & 0 & \dots & 0 & 0 \\ 0 & c_2 & 0 & \dots & 0 & 0 \\ 0 & 0 & c_3 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & c_{N-2} & 0 \\ 0 & 0 & 0 & & 0 & c_{N-1} \end{pmatrix} \end{aligned} \tag{2.93}$$

In case $\det(a_{jk}) \neq 0$ one can express the multiple integral in (2.91) according to (2.36) as follows

$$\begin{aligned}\tilde{\phi}(0,t_N|0,t_0) &= lim_{N\to\infty}\left[\frac{m}{2\pi i\hbar\epsilon_N}\right]^{\frac{N}{2}}\left[\frac{(i\pi)^{N-1}}{\det(\mathbf{a})}\right]^{\frac{1}{2}} \\ &= lim_{N\to\infty}\left[\frac{m}{2\pi i\hbar}\,\frac{1}{\epsilon_N\left(\frac{2\hbar\epsilon_N}{m}\right)^{N-1}\det(\mathbf{a})}\right]^{\frac{1}{2}} . \end{aligned} \tag{2.94}$$

In order to determine $\tilde{\phi}(0,t_N|0,t_0)$ we need to evaluate the function

$$f(t_0,t_N) \;=\; lim_{N\to\infty}\left[\epsilon_N\left(\frac{2\hbar\epsilon_N}{m}\right)^{N-1}\det(\mathbf{a})\right] . \tag{2.95}$$

According to (2.93) holds

$$D_{N-1} \;\stackrel{\rm def}{=}\; \left[\tfrac{2\hbar\epsilon_N}{m}\right]^{N-1}\det(\mathbf{a}) \tag{2.96}$$

$$= \left|\begin{pmatrix} 2-\frac{\epsilon_N^2}{m}c_1 & -1 & 0 & \dots & 0 & 0 \\ -1 & 2-\frac{\epsilon_N^2}{m}c_2 & -1 & \dots & 0 & 0 \\ 0 & -1 & 2-\frac{\epsilon_N^2}{m}c_3 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 2-\frac{\epsilon_N^2}{m}c_{N-2} & -1 \\ 0 & 0 & 0 & & -1 & 2-\frac{\epsilon_N^2}{m}c_{N-1} \end{pmatrix}\right|$$

In the following we will asume that the dimension $n = N-1$ of the matrix in (2.97) is variable. One can derive then for D_n the recursion relationship

$$D_n \;=\; \left(2 \,-\, \frac{\epsilon_N^2}{m}\,c_n\right) D_{n-1} \;-\; D_{n-2} \tag{2.97}$$

using the well-known method of expanding a determinant in terms of the determinants of lower dimensional submatrices. Using the starting values [c.f. the comment below Eq. (2.43)]

$$D_0 \;=\; 1\,; \quad D_1 \;=\; 2 \,-\, \frac{\epsilon_N^2}{m}\,c_1 \tag{2.98}$$

this recursion relationship can be employed to determine D_{N-1}. One can express (2.97) through the 2nd order difference equation

$$\frac{D_{n+1} \,-\, 2D_n \,+\, D_{n-1}}{\epsilon_N^2} \;=\; -\,\frac{c_{n+1}\,D_n}{m}\,. \tag{2.99}$$

Since we are interested in the solution of this equation in the limit of vanishing ϵ_N we may interpret (2.99) as a 2nd order differential equation in the continuous variable $t = n\epsilon_N + t_0$

$$\frac{d^2 f(t_0,t)}{dt^2} \;=\; -\,\frac{c(t)}{m}\,f(t_0,t)\,. \tag{2.100}$$

The boundary conditions at $t = t_0$, according to (2.98), are

$$\begin{aligned} f(t_0,t_0) &= \epsilon_N D_0 = 0 ; \\ \left.\frac{df(t_0,t)}{dt}\right|_{t=t_0} &= \epsilon_N \frac{D_1 - D_0}{\epsilon_N} = 2 - \frac{\epsilon_N^2}{m} c_1 - 1 = 1 . \end{aligned} \tag{2.101}$$

We have then finally for the propagator (2.79)

$$\phi(x,t|x_0,t_0) = \left[\frac{m}{2\pi i\hbar f(t_o,t)}\right]^{\frac{1}{2}} \exp\left\{\frac{i}{\hbar} S[x_{cl}(t)]\right\} \tag{2.102}$$

where $f(t_0,t)$ is the solution of (2.100, 2.101) and where $S[x_{cl}(t)]$ is determined by solving first the Euler–Lagrange equations for the Lagrangian (2.80) to obtain the classical path $x_{cl}(t)$ with end points $x_{cl}(t_0) = x_0$ and $x_{cl}(t_N) = x_N$ and then evaluating (2.81) for this path. Note that the required solution $x_{cl}(t)$ involves a solution of the Euler–Lagrange equations for boundary conditions which are different from those conventionally encountered in Classical Mechanics where usually a solution for initial conditions $x_{cl}(t_0) = x_0$ and $\dot{x}_{cl}(t_0) = v_0$ are determined.

2.6 Wave Packet Moving in Homogeneous Force Field

We want to consider now the motion of a quantum mechanical particle, decribed at time $t = t_o$ by a wave packet (2.48), in the presence of a homogeneous force due to a potential $V(x) = -f\,x$. As we have learnt from the study of the time-development of (2.48) in case of free particles the wave packet (2.48) corresponds to a classical particle with momentum p_o and position $x_o = 0$. We expect then that the classical particle assumes the following position and momentum at times $t > t_o$

$$y(t) = \frac{p_o}{m}(t - t_o) + \frac{1}{2}\frac{f}{m}(t - t_o)^2 \tag{2.103}$$

$$p(t) = p_o + f\,(t - t_o) \tag{2.104}$$

The Lagrangian for the present case is

$$L(x,\dot{x},t) = \frac{1}{2}m\,\dot{x}^2 + f\,x . \tag{2.105}$$

This corresponds to the Lagrangian in (2.80) for $c(t) \equiv 0$, $e(t) \equiv -f$. Accordingly, we can employ the expression (2.89, 2.90) for the propagator where, in the present case, holds $L'(y,\dot{y},t) = \frac{1}{2}m\dot{y}^2$ such that $\tilde{\phi}(0,t_N|0,t_0)$ is the free particle propagator (2.45). One can write then the propagator for a particle moving subject to a homogeneous force

$$\phi(x,t|x_0,t_0) = \left[\frac{m}{2\pi i\hbar(t-t_0)}\right]^{\frac{1}{2}} \exp\left[\frac{i}{\hbar} S[x_{cl}(\tau)]\right] . \tag{2.106}$$

Here $S[x_{cl}(\tau)]$ is the action integral over the classical path with end points

$$x_{cl}(t_o) = x_o \quad , \qquad x_{cl}(t) = x . \tag{2.107}$$

The classical path obeys

$$m\,\ddot{x}_{cl} \;=\; f\;. \tag{2.108}$$

The solution of (2.107, 2.108) is

$$x_{cl}(\tau) \;=\; x_o \;+\; \left(\frac{x\,-\,x_o}{t\,-\,t_o} \;-\; \frac{1}{2}\,\frac{f}{m}\,(t\,-\,t_o)\right)\,\tau \;+\; \frac{1}{2}\,\frac{f}{m}\,\tau^2 \tag{2.109}$$

as can be readily verified. The velocity along this path is

$$\dot{x}_{cl}(\tau) \;=\; \frac{x\,-\,x_o}{t\,-\,t_o} \;-\; \frac{1}{2}\,\frac{f}{m}\,(t\,-\,t_o) \;+\; \frac{f}{m}\,\tau \tag{2.110}$$

and the Lagrangian along the path, considered as a function of τ, is

$$\begin{aligned} g(\tau) \;&=\; \frac{1}{2}m\dot{x}_{cl}^2\,(\tau) \;+\; f\,x_{cl}\,(\tau) \\ &=\; \frac{1}{2}m\left(\frac{x\,-\,x_o}{t\,-\,t_o} \;-\; \frac{1}{2}\,\frac{f}{m}\,(t\,-\,t_o)\right)^2 \;+\; f\left(\frac{x\,-\,x_o}{t\,-\,t_o} \;-\; \frac{1}{2}\,\frac{f}{m}\,(t\,-\,t_o)\right)\tau \\ &\quad\; +\,\frac{1}{2}\,\frac{f^2}{m}\,\tau^2 +\; f\,x_o \;+\; f\left(\frac{x\,-\,x_o}{t\,-\,t_o} \;-\; \frac{1}{2}\,\frac{f}{m}\,(t\,-\,t_o)\right)\tau \;+\; \frac{1}{2}\,\frac{f^2}{m}\,\tau^2 \\ &=\; \frac{1}{2}m\left(\frac{x\,-\,x_o}{t\,-\,t_o} \;-\; \frac{1}{2}\,\frac{f}{m}\,(t\,-\,t_o)\right)^2 \;+\; 2f\left(\frac{x\,-\,x_o}{t\,-\,t_o} \;-\; \frac{1}{2}\,\frac{f}{m}\,(t\,-\,t_o)\right)\tau \\ &\quad\; +\,\frac{f^2}{m}\,\tau^2 \;+\; f\,x_o \end{aligned} \tag{2.111}$$

One obtains for the action integral along the classical path

$$\begin{aligned} S[x_{cl}\,(\tau)] \;&=\; \int_{t_o}^{t} d\tau\; g(\tau) \\ &=\; \frac{1}{2}m\left(\frac{x\,-\,x_o}{t\,-\,t_o} \;-\; \frac{1}{2}\,\frac{f}{m}\,(t\,-\,t_o)\right)^2 (t\,-\,t_o) \\ &\quad\; +\,f\left(\frac{x\,-\,x_o}{t\,-\,t_o} \;-\; \frac{1}{2}\,\frac{f}{m}\,(t\,-\,t_o)\right)(t\,-\,t_o)^2 \\ &\quad\; +\,\frac{1}{3}\,\frac{f^2}{m}\,(t\,-\,t_o)^3 \;+\; x_o\,f\,(t\,-\,t_o) \\ &=\; \frac{1}{2}\,m\,\frac{(x\,-\,x_o)^2}{t\,-\,t_o} \;+\; \frac{1}{2}\,(x\,+\,x_o)\,f\,(t\,-\,t_o) \;-\; \frac{1}{24}\,\frac{f^2}{m}\,(t\,-\,t_o)^3 \end{aligned} \tag{2.112}$$

and, finally, for the propagator

$$\begin{aligned} \phi(x,t|x_o,t_o) \;&=\; \left[\frac{m}{2\pi i\hbar(t-t_o)}\right]^{\frac{1}{2}} \;\times \\ &\times\,\exp\left[\,\frac{im}{2\hbar}\,\frac{(x\,-\,x_o)^2}{t\,-\,t_o} \;+\; \frac{i}{2\hbar}\,(x\,+\,x_o)\,f\,(t\,-\,t_o) \;-\; \frac{i}{24}\,\frac{f^2}{\hbar m}\,(t\,-\,t_o)^3\right] \end{aligned} \tag{2.113}$$

The propagator (2.113) allows one to determine the time-evolution of the initial state (2.48) using (2.5). Since the propagator depends only on the time-difference $t - t_o$ we can assume, withoult loss of generality, $t_o = 0$ and are lead to the integral

$$\begin{aligned}\psi(x,t) &= \left[\frac{1}{\pi\delta^2}\right]^{\frac{1}{4}}\left[\frac{m}{2\pi i\hbar t}\right]^{\frac{1}{2}}\int_{-\infty}^{+\infty}dx_0 \\ &\exp\underbrace{\left[\frac{im}{2\hbar}\frac{(x-x_0)^2}{t} - \frac{x_0^2}{2\delta^2} + i\frac{p_o}{\hbar}x_o + \frac{i}{2\hbar}(x+x_o)\,f\,t - \frac{i}{24}\frac{f^2}{m\hbar}t^3\right]}_{E_o(x_o,x)\;+\;E(x)}\end{aligned} \tag{2.114}$$

To evaluate the integral we adopt the same computational strategy as used for (2.51) and divide the exponent in (2.114) as follows [c.f. (2.54)]

$$E_o(x_o,x) = \frac{im}{2\hbar t}\left[x_o^2\left(1 + i\frac{\hbar t}{m\delta^2}\right) - 2x_o\left(x - \frac{p_o}{m}t - \frac{f\,t^2}{2m}\right) + f(x)\right] \tag{2.115}$$

$$E(x) = \frac{im}{2\hbar t}\left[x^2 + \frac{f\,t^2}{m}x - f(x)\right] - \frac{1}{24}\frac{f^2t^3}{\hbar m}\,. \tag{2.116}$$

One chooses then $f(x)$ to complete, according to (2.52), the square in (2.115)

$$f(x) = \left(\frac{x - \frac{p_o}{m}t - \frac{f\,t^2}{2m}}{\sqrt{1 + i\frac{\hbar t}{m\delta^2}}}\right)^2\,. \tag{2.117}$$

This yields

$$E_o(x_o,x) = \frac{im}{2\hbar t}\left(x_o\sqrt{1 + i\frac{\hbar t}{m\delta^2}} - \frac{x - \frac{p_o}{m}t - \frac{f\,t^2}{2m}}{\sqrt{1 + i\frac{\hbar t}{m\delta^2}}}\right)^2\,. \tag{2.118}$$

Following in the footsteps of the calculation on page 18 ff. one can state again

$$\psi(x,t) = \left[\frac{1 - i\frac{\hbar t}{m\delta^2}}{1 + i\frac{\hbar t}{m\delta^2}}\right]^{\frac{1}{4}}\left[\frac{1}{\pi\delta^2\,(1 + \frac{\hbar^2t^2}{m^2\delta^4})}\right]^{\frac{1}{4}}\exp\left[\,E(x)\,\right] \tag{2.119}$$

and is lead to the exponential (2.116)

$$E(x) = -\frac{1}{24}\frac{f^2t^3}{\hbar m} + \frac{im}{2\hbar t(1 + i\frac{\hbar t}{m\delta^2})}\,S(x) \tag{2.120}$$

where

$$\begin{aligned}S(x) &= x^2\left(1 + i\frac{\hbar t}{m\delta^2}\right) + x\frac{ft^2}{m}\left(1 + i\frac{\hbar t}{m\delta^2}\right) - \left(x - \frac{p_o}{m}t - \frac{ft^2}{2m}\right)^2 \\ &= \left(x - \frac{p_o}{m}t - \frac{ft^2}{2m}\right)^2\left(1 + i\frac{\hbar t}{m\delta^2}\right) - \left(x - \frac{p_o}{m}t - \frac{ft^2}{2m}\right)^2 \\ &\quad + \left[x\frac{ft^2}{m} + 2x\left(\frac{p_o}{m}t + \frac{ft^2}{2m}\right) - \left(\frac{p_o}{m}t + \frac{ft^2}{2m}\right)^2\right]\left(1 + i\frac{\hbar t}{m\delta^2}\right)\end{aligned} \tag{2.121}$$

Inserting this into (2.120) yields

$$E(x) = -\frac{\left(x - \frac{p_o}{m}t - \frac{ft^2}{2m}\right)^2}{2\delta^2\left(1 + i\frac{\hbar t}{m\delta^2}\right)} + \frac{i}{\hbar}(p_o + ft)x - \frac{i}{2m\hbar}\left(p_o t + p_o f t^2 + \frac{f^2t^3}{4} + \frac{f^2t^3}{12}\right) \tag{2.122}$$

The last term can be written

$$-\frac{i}{2m\hbar}\left(p_o t + p_o f t^2 + \frac{f^2t^3}{3}\right) = -\frac{i}{2m\hbar}\int_0^t d\tau\,(p_o + f\tau)^2\,. \tag{2.123}$$

Altogether, (2.119, 2.122, 2.123) provide the state of the particle at time $t > 0$

$$\begin{aligned}\psi(x,t) = {} & \left[\frac{1 - i\frac{\hbar t}{m\delta^2}}{1 + i\frac{\hbar t}{m\delta^2}}\right]^{\frac{1}{4}}\left[\frac{1}{\pi\delta^2\left(1 + \frac{\hbar^2t^2}{m^2\delta^4}\right)}\right]^{\frac{1}{4}} \times \\ & \times \exp\left[-\frac{\left(x - \frac{p_o}{m}t - \frac{ft^2}{2m}\right)^2}{2\delta^2\left(1 + \frac{\hbar^2t^2}{m^2\delta^4}\right)}\left(1 - i\frac{\hbar t}{m\delta^2}\right)\right] \times \\ & \times \exp\left[\frac{i}{\hbar}(p_o + ft)x - \frac{i}{\hbar}\int_0^t d\tau\,\frac{(p_o + f\tau)^2}{2m}\right]\,.\end{aligned} \tag{2.124}$$

The corresponding probablity distribution is

$$|\psi(x,t)|^2 = \left[\frac{1}{\pi\delta^2\left(1 + \frac{\hbar^2t^2}{m^2\delta^4}\right)}\right]^{\frac{1}{2}} \exp\left[-\frac{\left(x - \frac{p_o}{m}t - \frac{ft^2}{2m}\right)^2}{\delta^2\left(1 + \frac{\hbar^2t^2}{m^2\delta^4}\right)}\right]\,. \tag{2.125}$$

Comparision of Moving Wave Packet with Classical Motion

It is again [c.f. (2.4)] revealing to compare the probability distributions for the initial state (2.48) and for the states at time t, i.e., (2.125). Both distributions are Gaussians. Distribution (2.125) moves along the x-axis with distribution centers positioned at $y(t)$ given by (2.103), i.e., as expected for a classical particle. The states (2.124), in analogy to the states (2.71) for free particles, exhibit a phase factor $\exp[ip(t)x/\hbar]$, for which $p(t)$ agrees with the classical momentum (2.104). While these properties show a close correspondence between classical and quantum mechanical behaviour, the distribution shows also a pure quantum effect, in that it increases its width . This increase, for the homogeneous force case, is identical to the increase (2.73) determined for a free particle. Such increase of the width of a distribution is not a necessity in quantum mechanics. In fact, in case of so-called bound states, i.e., states in which the classical and quantum mechanical motion is confined to a finite spatial volume, states can exist which do not alter their spatial distribution in time. Such states are called stationary states. In case of a harmonic potential there exists furthermore the possibility that the center of a wave packet follows the classical behaviour and the width remains constant in time. Such states are referred to as coherent states, or Glauber states, and will be

studied below. It should be pointed out that in case of vanishing, linear and quadratic potentials quantum mechanical wave packets exhibit a particularly simple evolution; in case of other type of potential functions and, in particular, in case of higher-dimensional motion, the quantum behaviour can show features which are much more distinctive from classical behaviour, e.g., tunneling and interference effects.

Propagator of a Harmonic Oscillator

In order to illustrate the evaluation of (2.102) we consider the case of a harmonic oscillator. In this case holds for the coefficents in the Lagrangian (2.80) $c(t) = m\omega^2$ and $e(t) = 0$, i.e., the Lagrangian is

$$L(x, \dot{x}) = \frac{1}{2} m \dot{x}^2 - \frac{1}{2} m \omega^2 x^2 . \tag{2.126}$$

.We determine first $f(t_0, t)$. In the present case holds

$$\ddot{f} = -\omega^2 f ; \quad f(t_0, t_0) = 0 ; \quad \dot{f}(t_o, t_o) = 1 . \tag{2.127}$$

The solution which obeys the stated boundary conditions is

$$f(t_0, t) = \frac{\sin\omega(t - t_0)}{\omega} . \tag{2.128}$$

We determine now $S[x_{cl}(\tau)]$. For this purpose we seek first the path $x_{cl}(\tau)$ which obeys $x_{cl}(t_0) = x_0$ and $x_{cl}(t) = x$ and satisfies the Euler–Lagrange equation for the harmonic oscillator

$$m\ddot{x}_{cl} + m\omega^2 x_{cl} = 0 . \tag{2.129}$$

This equation can be written

$$\ddot{x}_{cl} = -\omega^2 x_{cl} . \tag{2.130}$$

the general solution of which is

$$x_{cl}(\tau') = A \sin\omega(\tau - t_0) + B \cos\omega(\tau - t_0) . \tag{2.131}$$

The boundary conditions $x_{cl}(t_0) = x_0$ and $x_{cl}(t) = x$ are satisfied for

$$B = x_0 ; \quad A = \frac{x - x_0 \cos\omega(t - t_o)}{\sin\omega(t - t_0)} , \tag{2.132}$$

and the desired path is

$$x_{cl}(\tau) = \frac{x - x_0\, c}{s} \sin\omega(\tau - t_0) + x_0 \cos\omega(\tau - t_0) \tag{2.133}$$

where we introduced

$$c = \cos\omega(t - t_o) , \quad s = \sin\omega(t - t_o) \tag{2.134}$$

We want to determine now the action integral associated with the path (2.133, 2.134)

$$S[x_{cl}(\tau)] = \int_{t_0}^{t} d\tau \left(\frac{1}{2} m \dot{x}_{cl}^2(\tau) - \frac{1}{2} m\omega^2 x_{cl}^2(\tau) \right) \tag{2.135}$$

For this purpose we assume presently $t_o = 0$. From (2.133) follows for the velocity along the classical path

$$\dot{x}_{cl}(\tau) = \omega \frac{x - x_0\, c}{s} \cos\omega\tau - \omega\, x_0 \sin\omega\tau \tag{2.136}$$

and for the kinetic energy

$$\begin{aligned} \frac{1}{2}m\dot{x}_{cl}^2(\tau) &= \frac{1}{2}m\omega^2 \frac{(x - x_0\, c)^2}{s^2} \cos^2\omega\tau \\ &\quad -m\omega^2 x_o \frac{x - x_0\, c}{s} \cos\omega\tau \sin\omega\tau \\ &\quad +\frac{1}{2}m\omega^2\, x_o^2 \sin^2\omega\tau \end{aligned} \tag{2.137}$$

Similarly, one obtains from (2.133) for the potential energy

$$\begin{aligned} \frac{1}{2}m\omega^2 x_{cl}^2(\tau) &= \frac{1}{2}m\omega^2 \frac{(x - x_0\, c)^2}{s^2} \sin^2\omega\tau \\ &\quad +m\omega^2 x_o \frac{x - x_0\, c}{s} \cos\omega\tau \sin\omega\tau \\ &\quad +\frac{1}{2}m\omega^2\, x_o^2 \cos^2\omega\tau \end{aligned} \tag{2.138}$$

Using

$$\cos^2\omega\tau = \frac{1}{2} + \frac{1}{2}\cos 2\omega\tau \tag{2.139}$$

$$\sin^2\omega\tau = \frac{1}{2} - \frac{1}{2}\cos 2\omega\tau \tag{2.140}$$

$$\cos\omega\tau \sin\omega\tau = \frac{1}{2}\sin 2\omega\tau \tag{2.141}$$

the Lagrangian, considered as a function of τ, reads

$$\begin{aligned} g(\tau) = \frac{1}{2}m\dot{x}_{cl}^2(\tau) - \frac{1}{2}m\omega^2 x_{cl}^2(\tau) &= \frac{1}{2}m\omega^2 \frac{(x - x_0\, c)^2}{s^2} \cos 2\omega\tau \\ &\quad -m\omega^2 x_o \frac{x - x_0\, c}{s} \sin 2\omega\tau \\ &\quad -\frac{1}{2}m\omega^2\, x_o^2 \cos 2\omega\tau \end{aligned} \tag{2.142}$$

Evaluation of the action integral (2.135), i.e., of $S[x_{cl}(\tau)] = \int_0^t d\tau g(\tau)$ requires the integrals

$$\int_0^t d\tau \cos 2\omega\tau = \frac{1}{2\omega}\sin 2\omega t = \frac{1}{\omega}\, s\, c \tag{2.143}$$

$$\int_0^t d\tau \sin 2\omega\tau = \frac{1}{2\omega}\left[1 - \cos 2\omega t\right] = \frac{1}{\omega}\, s^2 \tag{2.144}$$

where we employed the definition (2.134) Hence, (2.135) is, using $s^2 + c^2 = 1$,

$$\begin{aligned} S[x_{cl}(\tau)] &= \frac{1}{2}m\omega \frac{(x - x_0\, c)^2}{s^2}\, s\, c - m\omega x_o \frac{x - x_0\, c}{s}\, s^2 - \frac{1}{2}m\omega^2\, x_o^2\, s\, c \\ &= \frac{m\omega}{2s}\left[(x^2 - 2xx_oc + x_o^2c^2)\, c - 2x_oxs^2 + 2x_o^2cs^2 - x_o^2s^2c\right] \\ &= \frac{m\omega}{2s}\left[(x^2 + x_o^2)\, c - 2x_ox\right] \end{aligned} \tag{2.145}$$

and, with the definitions (2.134),

$$S[x_{cl}(\tau)] \;=\; \frac{m\omega}{2\sin\omega(t\,-\,t_0)}\,\left[(x_0^2\,+\,x^2)\cos\omega(t\,-\,t_0)\;-\;2x_0x\right]\;. \tag{2.146}$$

For the propagator of the harmonic oscillator holds then

$$\begin{aligned}\phi(x,t|x_0,t_0) \;&=\; \left[\tfrac{m\omega}{2\pi i\hbar\sin\omega(t-t_0)}\right]^{\frac{1}{2}}\;\times\\ \times\exp\left\{\tfrac{i\,m\,\omega}{2\hbar\sin\omega(t-t_0)}\,\left[(x_0^2\,+\,x^2)\cos\omega(t\,-\,t_0)\;-\;2x_0x\right]\right\}\quad .\end{aligned} \tag{2.147}$$

Quantum Pendulum or Coherent States

As a demonstration of the application of the propagator (2.147) we use it to describe the time development of the wave function for a particle in an initial state

$$\psi(x_0,t_0) \;=\; \left[\frac{m\omega}{\pi\hbar}\right]^{\frac{1}{4}}\,\exp\left(-\,\frac{m\omega(x_0-b_o)^2}{2\hbar}\;+\;\frac{i}{\hbar}\,p_o\,x_o\right)\;. \tag{2.148}$$

The initial state is decribed by a Gaussian wave packet centered around the position $x = b_o$ and corresponds to a particle with initial momentum p_o. The latter property follows from the role of such factor for the initial state (2.48) when applied to the case of a free particle [c.f. (2.71)] or to the case of a particle moving in a homogeneous force [c.f. (2.124, 2.125)] and will be borne out of the following analysis; at present one may regard it as an assumption.
If one identifies the center of the wave packet with a classical particle, the following holds for the time development of the position (displacement), momentum, and energy of the particle

$$\begin{aligned} b(t) \;&=\; b_o\,\cos\omega(t-t_o)\;+\;\frac{p_o}{m\omega}\,\sin\omega(t-t_o) && \text{displacement}\\ p(t) \;&=\; -\,m\omega b_o\,\sin\omega(t-t_o)\;+\;p_o\,\cos\omega(t-t_o) && \text{momentum}\\ \epsilon_o \;&=\; \frac{p_o^2}{2m}\;+\;\tfrac{1}{2}m\omega^2b_o^2 && \text{energy}\end{aligned} \tag{2.149}$$

We want to explore, using (2.5), how the probability distribution $|\psi(x,t)|^2$ of the quantum particle propagates in time.
The wave function at times $t > t_0$ is

$$\psi(x,t) \;=\; \int_{-\infty}^{\infty}dx_0\,\phi(x,t|x_0,t_0)\,\psi(x_0,t_0)\;. \tag{2.150}$$

Expressing the exponent in (2.148)

$$\frac{i\,m\,\omega}{2\hbar\sin\omega(t-t_o)}\;\left[i\,(x_o-b_o)^2\,\sin\omega(t-t_o)\;+\;\frac{2p_o}{m\omega}\,x_o\,\sin\omega(t-t_o)\right] \tag{2.151}$$

(2.147, 2.150, 2.151) can be written

$$\psi(x,t) \;=\; \left[\frac{m\omega}{\pi\hbar}\right]^{\frac{1}{4}}\;\left[\frac{m}{2\pi i\omega\hbar\sin\omega(t\,-\,t_0)}\right]^{\frac{1}{2}}\;\int_{-\infty}^{\infty}dx_0\exp\left[\,E_0\;+\;E\,\right] \tag{2.152}$$

where

$$E_0(x_o,x) \;=\; \frac{i\,m\,\omega}{2\hbar s}\left[x_o^2 c \,-\, 2x_o x \,+\, i s x_o^2 \,-\, 2is x_o b_o \,+\, \frac{2p_o}{m\omega}\,x_o s \,+\, f(x)\right] \tag{2.153}$$

$$E(x) \;=\; \frac{im\omega}{2\hbar s}\,\left[\,x^2 c \,+\, i s b_o^2 \,-\, f(x)\,\right]\,. \tag{2.154}$$

$$c \;=\; \cos\omega(t\,-\,t_o)\,, \quad s \;=\; \sin\omega(t\,-\,t_o)\,. \tag{2.155}$$

Here $f(x)$ is a function which is introduced to complete the square in (2.153) for simplification of the Gaussian integral in x_0. Since $E(x)$ is independent of x_o (2.152) becomes

$$\psi(x,t) \;=\; \left[\frac{m\omega}{\pi\hbar}\right]^{\frac{1}{4}}\,\left[\frac{m}{2\pi i\omega\hbar\sin\omega(t\,-\,t_0)}\right]^{\frac{1}{2}}\,e^{E(x)}\,\int_{-\infty}^{\infty}dx_0\exp\left[\,E_0(x_o,x)\right] \tag{2.156}$$

We want to determine now $E_o(x_o,x)$ as given in (2.153). It holds

$$E_o \;=\; \frac{i\,m\,\omega}{2\hbar s}\left[x_o^2\,e^{i\omega(t-t_o)} \,-\, 2x_o(x + isb_o - \frac{p_o}{m\omega}\,s) \,+\, f(x)\right] \tag{2.157}$$

For $f(x)$ to complete the square we choose

$$f(x) \;=\; (x \,+\, isb_o \,-\, \frac{p_o}{m\omega}\,s)^2\,e^{-i\omega(t-t_o)}\,. \tag{2.158}$$

One obtains for (2.157)

$$E_0(x_o,x) \;=\; \frac{im\omega}{2\hbar s}\exp\left[i\omega(t-t_0)\right]\,\left[x_0 \,-\, (x + isb_o - \frac{p_o}{m\omega}\,s)\exp\left(-i\omega(t-t_0)\right]^2\,. \tag{2.159}$$

To determine the integral in (2.156) we employ the integration formula (2.247) and obtain

$$\int_{-\infty}^{+\infty}dx_0\,e^{E_0(x_0)} \;=\; \left[\frac{2\pi i\hbar\sin\omega(t-t_0)}{m\omega\exp[i\omega(t-t_0)]}\right]^{\frac{1}{2}} \tag{2.160}$$

Inserting this into (2.156) yields

$$\psi(x,t) \;=\; \left[\frac{m\omega}{\pi\hbar}\right]^{\frac{1}{4}}\;e^{E(x)} \tag{2.161}$$

For $E(x)$ as defined in (2.154) one obtains, using $\exp[\pm i\omega(t-t_o)] \,=\, c \,\pm\, is$,

$$\begin{aligned}
E(x) \;=\;& \tfrac{im\omega}{2\hbar s}\,\left[\,x^2c \,+\, isb_o^2 \,-\, x^2c \,+\, isx^2 \,-\, 2isxb_oc \,-\, 2s^2xb_o\right.\\
&\quad +\, s^2b_o^2c \,-\, is^3b_o^2 \,+\, 2\,\tfrac{p_o}{m\omega}\,xsc \,+\, 2i\,\tfrac{p_o}{m\omega}\,b_os^2c\\
&\quad \left.-\,2i\,\tfrac{p_o}{m\omega}\,xs^2 \,+\, 2\,\tfrac{p_o}{m\omega}\,b_os^3 \,-\, \tfrac{p_o^2}{m^2\omega^2}\,s^2c \,+\, i\,\tfrac{p_o^2}{m^2\omega^2}\,s^3\right]\\
=\;& -\,\tfrac{m\omega}{2\hbar}\,\left[\,x^2 \,+\, c^2b_o^2 \,-\, 2xb_oc \,+\, 2ixsb_o \,-\, ib_o^2sc\right.\\
&\quad -2\,\tfrac{p_o}{m\omega}\,xs \,+\, 2\,\tfrac{p_o}{m\omega}\,b_osc \,+\, \tfrac{p_o^2}{m^2\omega^2}\,s^2 \,-\, 2i\,\tfrac{p_o}{m\omega}\,xc\\
&\quad \left.-2i\,\tfrac{p_o}{m\omega}\,b_os^2 \,+\, i\,\tfrac{p_o^2}{m^2\omega^2}\,sc\right]\\
=\;& -\,\tfrac{m\omega}{2\hbar}\,(x \,-\, cb_o \,-\, \tfrac{p_o}{m\omega}\,s)^2 \,+\, \frac{i}{\hbar}\,(-\,m\omega b_os \,+\, p_oc)\,x\\
&\quad -\,\frac{i}{\hbar}\,(\tfrac{p_o^2}{2m\omega} \,-\, \tfrac{1}{2}m\omega b_o^2)sc \,+\, \frac{i}{\hbar}\,p_ob_os^2
\end{aligned} \tag{2.162}$$

We note the following identities

$$\int_{t_o}^{t} d\tau \, \frac{p^2(\tau)}{2m} = \frac{1}{2}\,\epsilon_o(t-t_o) \,+\, \frac{1}{2}\left(\frac{p_o^2}{2m\omega} \,-\, \frac{m\omega b_o^2}{2}\right) sc \,-\, \frac{1}{2}\,b_o p_o s^2 \tag{2.163}$$

$$\int_{t_o}^{t} d\tau \, \frac{m\omega^2 b^2(\tau)}{2} = \frac{1}{2}\,\epsilon_o(t-t_o) \,-\, \frac{1}{2}\left(\frac{p_o^2}{2m\omega} \,-\, \frac{m\omega b_o^2}{2}\right) sc \,+\, \frac{1}{2}\,b_o p_o s^2 \tag{2.164}$$

where we employed $b(\tau)$ and $p(\tau)$ as defined in (2.149). From this follows, using $p(\tau) = m\dot{b}(\tau)$ and the Lagrangian (2.126),

$$\int_{t_o}^{t} d\tau \, L[b(\tau), \dot{b}(\tau)] \;=\; \left(\frac{p_o^2}{2m\omega} \,-\, \frac{m\omega b_o^2}{2}\right) sc \,-\, b_o p_o s \tag{2.165}$$

such that $E(x)$ in (2.162) can be written, using again (2.149)),

$$E(x) \;=\; -\,\frac{m\omega}{2\hbar}\,[x-b(t)]^2 \,+\, \frac{i}{\hbar}\,p(t)\,x \,-\, i\,\frac{1}{2}\,\omega\,(t-t_o) \,-\, \frac{i}{\hbar}\int_{t_o}^{t} d\tau \, L[b(\tau), \dot{b}(\tau)] \tag{2.166}$$

Inserting this into (2.161) yields,

$$\psi(x,t) \;=\; \left[\frac{m\omega}{\pi\hbar}\right]^{\frac{1}{4}} \,\times\, \exp\left\{-\,\frac{m\omega}{2\hbar}\,[x-b(t)]^2\right\} \,\times \tag{2.167}$$
$$\times\, \exp\left\{\frac{i}{\hbar}\,p(t)\,x \,-\, i\,\frac{1}{2}\,\omega\,(t-t_o) \,-\, \frac{i}{\hbar}\int_{t_o}^{t} d\tau \, L[b(\tau), \dot{b}(\tau)]\right\}$$

where $b(t)$, $p(t)$, and ϵ_o are the classical displacement, momentum and energy, respectively, defined in (2.149).

Comparision of Moving Wave Packet with Classical Motion

The probability distribution associated with (2.167)

$$|\psi(x,t)|^2 \;=\; \left[\frac{m\omega}{\pi\hbar}\right]^{\frac{1}{2}} \,\exp\left\{-\frac{m\omega}{\hbar}\,[x \,-\, b(t)]^2\right\} \tag{2.168}$$

is a Gaussian of time-independent width, the center of which moves as described by $b(t)$ given in (2.148) , i.e., the center follows the motion of a classical oscillator (pendulum) with initial position b_o and initial momentum p_o. It is of interest to recall that propagating wave packets in the case of vanishing [c.f. (2.72)] or linear [c.f. (2.125)] potentials exhibit an increase of their width in time; in case of the quantum oscillator for the particular width chosen for the initial state (2.148) the width, actually, is conserved. One can explain this behaviour as arising from constructive interference due to the restoring forces of the harmonic oscillator. We will show in Chapter 4 [c.f. (4.166, 4.178) and Fig. 4.1] that an initial state of arbitrary width propagates as a Gaussian with oscillating width.

In case of the free particle wave packet (2.48, 2.71) the factor $\exp(ip_o x)$ gives rise to the translational motion of the wave packet described by $p_o t/m$, i.e., p_o also corresponds to initial classical momentum. In case of a homogeneous force field the phase factor $\exp(ip_o x)$ for the initial state (2.48) gives rise to a motion of the center of the propagating wave packet [c.f. (2.125)] described by $(p_o/m)t + \frac{1}{2}ft^2$ such that again p_o corresponds to the classical momentum. Similarly, one observes for all three cases (free particle, linear and quadratic potential) a phase factor $\exp[ip(t)x/\hbar]$ for the propagating wave packet where $fp(t)$ corresponds to the initial classical momentum at time t. One can, hence, summarize that for the three cases studied (free particle, linear and quadratic potential) propagating wave packets show remarkably close analogies to classical motion.
We like to consider finally the propagation of an initial state as in (2.148), but with $b_o = 0$ and $p_o = 0$. Such state is given by the wave function

$$\psi(x_0, t_0) \;=\; \Big[\frac{m\omega}{\pi\hbar}\Big]^{\frac{1}{4}} \exp\left(-\frac{m\omega x_0^2}{2\hbar} \;-\; \frac{i\omega}{2}\, t_o \right) . \tag{2.169}$$

where we added a phase factor $\exp(-i\omega t_o/2)$. According to (2.167) the state (2.169) reproduces itself at later times t and the probablity distribution remains at all times equal to

$$\Big[\frac{m\omega}{\pi\hbar}\Big]^{\frac{1}{2}} \exp\left(-\frac{m\omega x_0^2}{\hbar} \right) , \tag{2.170}$$

i.e., the state (2.169) is a stationary state of the system. The question arises if the quantum oscillator posesses further stationary states. In fact, there exist an infinite number of such states which will be determined now.

2.7 Stationary States of the Harmonic Oscillator

In order to find the stationary states of the quantum oscillator we consider the function

$$W(x,t) \;=\; \exp\left(2\sqrt{\frac{m\omega}{\hbar}}\, x\, e^{-i\omega t} \;-\; e^{-2i\omega t} \;-\; \frac{m\omega}{2\hbar}\, x^2 \;-\; \frac{i\omega t}{2} \right) . \tag{2.171}$$

We want to demonstrate that $w(x,t)$ is invariant in time, i.e., for the propagator (2.147) of the harmonic oscillator holds

$$W(x,t) \;=\; \int_{-\infty}^{+\infty} dx_o\, \phi(x,t|x_o,t_o)\, W(x_o,t_o) . \tag{2.172}$$

We will demonstrate further below that (2.172) provides us in a nutshell with all the stationary states of the harmonic oscillator, i.e., with all the states with time-independent probability distribution.
In order to prove (2.172) we express the propagator, using (2.147) and the notation $T = t - t_o$

$$\begin{aligned} \phi(x,t|x_o,t_o) \;=\; & e^{-\frac{1}{2}i\omega T} \left[\frac{m\omega}{\pi\hbar(1-e^{-2i\omega T})} \right]^{\frac{1}{2}} \;\times \\ & \times \exp\left[-\frac{m\omega}{2\hbar}\,(x_o^2 + x^2)\, \frac{1 + e^{-2i\omega T}}{1 - e^{-2i\omega T}} \;-\; \frac{m\omega}{\hbar\omega}\, \frac{2x\, x_o e^{-i\omega T}}{1 - e^{-2i\omega T}} \right] \end{aligned} \tag{2.173}$$

One can write then the r.h.s. of (2.172)

$$I \;=\; e^{-\frac{1}{2}i\omega t}\left[\frac{m\omega}{\pi\hbar(1-e^{-2i\omega T})}\right]^{\frac{1}{2}}\;\int_{-\infty}^{+\infty}dx_o\exp[\,E_o(x_o,x)\;+\;E(x)\,] \tag{2.174}$$

where

$$\begin{aligned} E_o(x_o,x) \;&=\; -\frac{m\omega}{2\hbar}\left[x_o^2\left(\frac{1+e^{-2i\omega T}}{1-e^{-2i\omega T}}+1\right)\right. \\ &\qquad \left. +\,2x_o\left(\frac{2xe^{-i\omega T}}{1-e^{-2i\omega T}}\;+\;2\sqrt{\frac{\hbar}{m\omega}}\,e^{-i\omega t_o}\right)\;+\;f(x)\right] \end{aligned} \tag{2.175}$$

$$E(x) \;=\; -\frac{m\omega}{2\hbar}\left[x^2\,\frac{1+e^{-2i\omega T}}{1-e^{-2i\omega T}}\;+\;\frac{2\hbar}{m\omega}\,e^{-2i\omega t_o}\;-\;f(x)\right] \tag{2.176}$$

Following the by now familiar strategy one choses $f(x)$ to complete the square in (2.175), namely,

$$f(x)\;=\;\frac{1}{2}\,(1\;-\;e^{-2i\omega T})\left(\frac{2xe^{-i\omega T}}{1-e^{-2i\omega T}}\;+\;2\sqrt{\frac{\hbar}{m\omega}}\,e^{-i\omega t_o}\right)^2\,. \tag{2.177}$$

This choice of $f(x)$ results in

$$\begin{aligned} E_o(x_o,x) \;&=\; -\frac{m\omega}{2\hbar}\left[x_o\sqrt{\frac{2}{1-e^{-2i\omega T}}}\right. \\ &\qquad \left. +\,\sqrt{\frac{1-e^{-2i\omega T}}{2}}\left(\frac{2xe^{-i\omega T}}{1-e^{-2i\omega T}}\;+\;2\sqrt{\frac{\hbar}{m\omega}}\,e^{-i\omega t_o}\right)\right]^2 \\ &=\; i\,\frac{m\,\omega}{i\hbar(e^{-2i\omega T}-1)}\,(x_o\;+\;z_o)^2 \end{aligned} \tag{2.178}$$

for some constant $z_o\in\mathbb{C}$. Using (2.247) one obtains

$$\int_{-\infty}^{+\infty}dx_o\,e^{E_o(x_o,x)}\;=\;\left[\frac{\pi\hbar(1\;-\;e^{-2i\omega T}}{m\omega}\right]^{\frac{1}{2}} \tag{2.179}$$

and, therefore, one obtains for (2.174)

$$I\;=\;e^{-\frac{1}{2}i\omega t}\,e^{E(x)}\,. \tag{2.180}$$

For $E(x)$, as given in (2.176, 2.177), holds

$$\begin{aligned} E(x) \;&=\; -\frac{m\omega}{2\hbar}\left[x^2\,\frac{1+e^{-2i\omega T}}{1-e^{-2i\omega T}}\;+\;\frac{2\hbar}{m\omega}\,e^{-2i\omega t_o}\;-\;\frac{2x^2e^{-2i\omega T}}{1-e^{-2i\omega T}}\right. \\ &\qquad \left. -4\sqrt{\frac{\hbar}{m\omega}}\,x\,e^{-i\omega T}\;-\;2\,(1-e^{-2i\omega T})\,\frac{\hbar}{m\omega}\,e^{-2i\omega t_o}\right] \\ &=\; -\frac{m\omega}{2\hbar}\left[x^2\;-\;4\sqrt{\frac{\hbar}{m\omega}}\,x\,e^{-i\omega t}\;+\;\frac{2\hbar}{m\omega}\,e^{-2i\omega t}\right] \\ &=\; -\frac{m\omega}{2\hbar}\,x^2\;+\;2\sqrt{\frac{m\omega}{\hbar}}\,x\,e^{-i\omega t}\;-\;e^{-2i\omega t} \end{aligned} \tag{2.181}$$

Altogether, one obtains for the r.h.s. of (2.172)

$$I \;=\; \exp\left(2\sqrt{\frac{m\omega}{\hbar}}\, x\, e^{-i\omega t} \;-\; e^{-2i\omega t} \;-\; \frac{m\omega}{2\hbar}\, x^2 \;-\; \frac{1}{2} i\omega t \right) . \tag{2.182}$$

Comparision with (2.171) concludes the proof of (2.172).
We want to inspect the consequences of the invariance property (2.171, 2.172). We note that the factor $\exp(2\sqrt{m\omega/\hbar}\, xe^{-i\omega t} - e^{-2i\omega t})$ in (2.171) can be expanded in terms of $e^{-in\omega t}$, $n = 1, 2, \ldots$. Accoordingly, one can expand (2.171)

$$W(x,t) \;=\; \sum_{n=0}^{\infty} \frac{1}{n!}\, \exp[-i\omega(n+\tfrac{1}{2})\, t\,]\; \tilde{\phi}_n(x) \tag{2.183}$$

where the expansion coefficients are functions of x, but not of t. Noting that the propagator (2.147) in (2.172) is a function of $t - t_o$ and defining accordingly

$$\Phi(x, x_o; t - t_o) \;=\; \phi(x, t | x_o, t_o) \tag{2.184}$$

we express (2.172) in the form

$$\begin{aligned} &\sum_{n=0}^{\infty} \frac{1}{n!}\, \exp[-i\omega(n+\tfrac{1}{2})\, t]\; \tilde{\phi}_n(x) \\ &\qquad = \sum_{m=0}^{\infty} \int_{-\infty}^{+\infty} dx_o\, \Phi(x, x_o; t - t_o)\, \frac{1}{m!}\, \exp[-i\omega(m+\tfrac{1}{2})\, t_o]\; \tilde{\phi}_m(x_o) \end{aligned} \tag{2.185}$$

Replacing $t \;\rightarrow t + t_o$ yields

$$\begin{aligned} &\sum_{n=0}^{\infty} \frac{1}{n!}\, \exp[-i\omega(n+\tfrac{1}{2})\, (t + t_o)]\; \tilde{\phi}_n(x) \\ &\qquad = \sum_{m=0}^{\infty} \int_{-\infty}^{+\infty} dx_o\, \Phi(x, x_o; t)\, \frac{1}{m!}\, \exp[-i\omega(m+\tfrac{1}{2})\, t_o]\; \tilde{\phi}_m(x_o) \end{aligned} \tag{2.186}$$

Fourier transform, i.e., $\int_{-\infty}^{+\infty} dt_o\, \exp[i\omega(n+\frac{1}{2})\, t_o] \cdots$, results in

$$\begin{aligned} &\frac{1}{n!}\, \exp[-i\omega(n+\tfrac{1}{2})\, t]\; \tilde{\phi}_n(x) \\ &\qquad = \int_{-\infty}^{+\infty} dx_o\, \Phi(x, x_o; t - t_o)\, \frac{1}{n!}\, \tilde{\phi}_n(x_o) \end{aligned} \tag{2.187}$$

or

$$\begin{aligned} &\exp[-i\omega(n+\tfrac{1}{2})\, t]\; \tilde{\phi}_n(x) \\ &\qquad = \int_{-\infty}^{+\infty} dx_o\, \phi(x, t | x_o, t_o)\; \exp[-i\omega(n+\tfrac{1}{2})\, t_o]\; \tilde{\phi}_n(x_o) \, . \end{aligned} \tag{2.188}$$

Equation (2.188) identifies the functions $\tilde{\psi}_n(x,t) = \exp[-i\omega(n+\frac{1}{2})\, t]\; \tilde{\phi}_n(x)$ as invariants under the action of the propagator $\phi(x, t | x_o, t_o)$. In contrast to $W(x,t)$, which also exhibits such invariance,

the functions $\tilde{\psi}_n(x,t)$ are associated with a time-independent probablity density $|\tilde{\psi}_n(x,t)|^2 = |\tilde{\phi}_n(x)|^2$. Actually, we have identified then, through the expansion coefficients $\tilde{\phi}_n(x)$ in (2.183), stationary wave functions $\psi_n(x,t)$ of the quantum mechanical harmonic oscillator

$$\psi_n(x,t) \;=\; \exp[-i\omega(n+\tfrac{1}{2})\,t]\; N_n\; \tilde{\phi}_n(x)\;, \quad n \;=\; 0,1,2,\ldots \tag{2.189}$$

Here N_n are constants which normalize $\psi_n(x,t)$ such that

$$\int_{-\infty}^{+\infty} dx\,|\psi(x,t)|^2 \;=\; N_n^2 \int_{-\infty}^{+\infty} dx\,\tilde{\phi}_n^2(x) \;=\; 1 \tag{2.190}$$

is obeyed. In the following we will characterize the functions $\tilde{\phi}_n(x)$ and determine the normalization constants N_n. We will also argue that the functions $\psi_n(x,t)$ provide all stationary states of the quantum mechanical harmonic oscillator.

The Hermite Polynomials

The function (2.171), through expansion (2.183), characterizes the wave functions $\tilde{\phi}_n(x)$. To obtain closed expressions for $\tilde{\phi}_n(x)$ we simplify the expansion (2.183). For this purpose we introduce first the new variables

$$y \;=\; \sqrt{\frac{m\omega}{\hbar}}\,x \tag{2.191}$$

$$z \;=\; e^{-i\omega t} \tag{2.192}$$

and write (2.171)

$$W(x,t) \;=\; z^{\frac{1}{2}} e^{-y^2/2}\, w(y,z) \tag{2.193}$$

where

$$w(y,z) \;=\; \exp(2yz \,-\, z^2)\;. \tag{2.194}$$

Expansion (2.183) reads then

$$w(y,z)\, z^{\frac{1}{2}} e^{-y^2/2} \;=\; z^{\frac{1}{2}} \sum_{n=0}^{\infty} \frac{z^n}{n!}\, \tilde{\phi}_n(y) \tag{2.195}$$

or

$$w(y,z) \;=\; \sum_{n=0}^{\infty} \frac{z^n}{n!}\, H_n(y) \tag{2.196}$$

where

$$H_n(y) \;=\; e^{y^2/2}\, \tilde{\phi}_n(y)\;. \tag{2.197}$$

The expansion coefficients $H_n(y)$ in (2.197) are called *Hermite polynomials* which are polynomials of degree n which will be evaluated below. Expression (2.194) plays a central role for the *Hermite* polynomials since it contains, according to (2.194), in a 'nutshell' all information on the *Hermite* polynomials. This follows from

$$\frac{\partial^n}{\partial z^n}\, w(y,z)\bigg|_{z=0} \;=\; H_n(y) \tag{2.198}$$

which is a direct consequence of (2.196). One calls $w(y,z)$ the *generating function* for the *Hermite* polynomials. As will become evident in the present case generating functions provide an extremely elegant access to the special functions of Mathematical Physics[3]. We will employ (2.194, 2.196) to derive, among other properties, closed expressions for $H_n(y)$, normalization factors for $\tilde{\phi}(y)$, and recursion equations for the efficient evaluation of $H_n(y)$.
The identity (2.198) for the Hermite polynomials can be expressed in a more convenient form employing definition (2.196)

$$\begin{aligned} \frac{\partial^n}{\partial z^n}\, w(y,z)\Big|_{z=0} &= \frac{\partial^n}{\partial z^n}\, e^{2\,y\,z\,-\,z^2}\Big|_{z=0} \quad e^{y^2} \quad \frac{\partial^n}{\partial z^n}\, e^{-(y-z)^2}\Big|_{z=0} \\ &= (-1)^n\, e^{y^2} \quad \frac{\partial^n}{\partial y^n}\, e^{-(y-z)^2}\Big|_{z=0} = (-1)^n\, e^{y^2}\, \frac{\partial^n}{\partial y^n}\, e^{-y^2} \end{aligned} \tag{2.199}$$

Comparision with (2.196) results in the so-called Rodrigues formula for the *Hermite* polynomials

$$H_n(y) = (-1)^n\, e^{y^2}\, \frac{\partial^n}{\partial y^n}\, e^{-y^2}\ . \tag{2.200}$$

One can deduce from this expression the polynomial character of $H_n(y)$, i.e., that $H_n(y)$ is a polynomial of degree n. (2.200) yields for the first *Hermite* polynomials

$$H_0(y) = 1, \quad H_1(y) = 2y, \quad H_2(y) = 4y^2 - 2, \quad H_3(y) = 8y^3 - 12y, \ldots \tag{2.201}$$

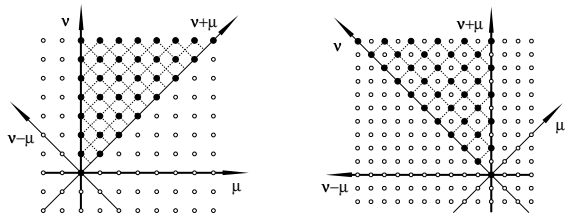


Figure 2.1: Schematic representation of change of summation variables ν and μ to $n = \nu + \mu$ and $m = \nu - \mu$. The diagrams illustrate that a summation over all points in a ν, μ lattice (left diagram) corresponds to a summation over only every other point in an n, m lattice (right diagram). The diagrams also identify the areas over which the summation is to be carried out.

We want to derive now explicit expressions for the *Hermite* polynomials. For this purpose we expand the generating function (2.194) in a Taylor series in terms of $y^p\, z^q$ and identify the corresponding coefficient c_{pq} with the coefficient of the p–th power of y in $H_q(y)$. We start from

$$\begin{aligned} e^{2yz\,-\,z^2} &= \sum_{\nu=0}^{\infty} \sum_{\mu=0}^{\nu} \frac{1}{\nu!} \left(\begin{array}{c} \nu \\ \mu \end{array} \right) z^{2\mu} (-1)^{\mu} (2y)^{\nu-\mu} z^{\nu-\mu} \\ &= \sum_{\nu=0}^{\infty} \sum_{\mu=0}^{\nu} \frac{1}{\nu} \left(\begin{array}{c} \nu \\ \mu \end{array} \right) (-1)^{\mu} (2y)^{\nu-\mu} z^{\nu+\mu} \end{aligned} \tag{2.202}$$

[3] *generatingfunctionology* by H.S.Wilf (Academic Press, Inc., Boston, 1990) is a useful introduction to this tool as is a chapter in the eminently useful *Concrete Mathematics* by R.L.Graham, D.E.Knuth, and O.Patashnik (Addison-Wesley, Reading, Massachusetts, 1989).

and introduce now new summation variables

$$n \;=\; \nu + \mu\,, \quad m \;=\; \nu - \mu \quad 0 \leq n < \infty\,,\; 0 \leq m \leq n\,. \tag{2.203}$$

The old summation variables ν, μ expressend in terms of n, m are

$$\nu \;=\; \frac{n+m}{2}\,, \quad \mu \;=\; \frac{n-m}{2}\,. \tag{2.204}$$

Since ν, μ are integers the summation over n, m must be restricted such that either both n and m are even or both n and m are odd. The lattices representing the summation terms are shown in Fig. 2.1. *With this restriction in mind* one can express (2.202)

$$e^{2yz\,-\,z^2} \;=\; \sum_{n=0}^{\infty} \frac{z^n}{n!} \sum_{m\geq 0}^{\leq n} \frac{n!\,(-1)^{\frac{n-m}{2}}}{\left(\frac{n-m}{2}\right)!\,m!}\,(2y)^m\,. \tag{2.205}$$

Since $(n-m)/2$ is an integer we can introduce now the summation variable $k \;=\; (n-m)/2\,,\; 0 \leq k \leq [n/2]$ where $[x]$ denotes the largest integer p, $p \leq x$. One can write then using $m \;=\; n-2k$

$$e^{2yz\,-\,z^2} \;=\; \sum_{n=0}^{\infty} \frac{z^n}{n!} \underbrace{\sum_{k=0}^{[n/2]} \frac{n!\,(-1)^k}{k!\,(n-2k)!}\,(2y)^{n-2k}}_{=\,H_n(y)}\,. \tag{2.206}$$

From this expansion we can identify $H_n(y)$

$$H_n(y) \;=\; \sum_{k=0}^{[n/2]} \frac{(-1)^k\,n!}{k!\,(n-2k)!}\,(2y)^{n-2k}\,. \tag{2.207}$$

This expression yields for the first four *Hermite* polynomials

$$H_0(y) \;=\; 1, \quad H_1(y) \;=\; 2y, \quad H_2(y) \;=\; 4y^2 \,-\, 2, \quad H_3(y) \;=\; 8y^3 \,-\, 12y, \ldots \tag{2.208}$$

which agrees with the expressions in (2.201).
From (2.207) one can deduce that $H_n(y)$, in fact, is a polynomial of degree n. In case of even n , the sum in (2.207) contains only even powers, otherwise, i.e., for odd n, it contains only odd powers. Hence, it holds

$$H_n(-y) \;=\; (-1)^n\,H_n(y)\,. \tag{2.209}$$

This property follows also from the generating function. According to (2.194) holds $w(-y,z) \;=\; w(y,-z)$ and, hence, according to 2.197)

$$\sum_{n=0}^{\infty} \frac{z^n}{n!} H_n(-y) \;=\; \sum_{n=0}^{\infty} \frac{(-z)^n}{n!} H_n(y) \;=\; \sum_{n=0}^{\infty} \frac{z^n}{n!}\,(-1)^n\,H_n(y) \tag{2.210}$$

from which one can conclude the property (2.209).

The generating function allows one to determine the values of $H_n(y)$ at $y = 0$. For this purpose one considers $w(0,z) = \exp(-z^2)$ and carries out the Taylor expansion on both sides of this expression resulting in

$$\sum_{m=0}^{\infty} \frac{(-1)^m\, z^{2m}}{m!} \;=\; \sum_{n=0}^{\infty} H_n(0)\, \frac{z^n}{n!}\,. \tag{2.211}$$

Comparing terms on both sides of the equation yields

$$H_{2n}(0) \;=\; (-1)^n\, \frac{(2n)!}{n!}\,, \quad H_{2n+1}(0) \;=\; 0\,, \quad n \;=\; 0,1,2,\ldots \tag{2.212}$$

This implies that stationary states of the harmonic oscillator $\phi_{2n+1}(x)$, as defined through (2.188, 2.197) above and given by (2.233) below, have a node at $y = 0$, a property which is consistent with (2.209) since odd functions have a node at the origin.

Recursion Relationships

A useful set of properties for special functions are the so-called recursion relationships. For *Hermite* polynomials holds, for example,

$$H_{n+1}(y) \;-\; 2y\, H_n(y) \;+\; 2n\, H_{n-1}(y) \;=\; 0\,, \quad n \;=\; 1,2,\ldots \tag{2.213}$$

which allow one to evaluate $H_n(y)$ from $H_0(y)$ and $H_1(y)$ given by (2.208). Another relationship is

$$\frac{d}{dy}\, H_n(y) \;=\; 2n\, H_{n-1}(y), \quad n \;=\; 1,2,\ldots \tag{2.214}$$

We want to derive (**??**) using the *generating function*. Starting point of the derivation is the property of $w(y,z)$

$$\frac{\partial}{\partial z}\, w(y,z) \;-\; (2y \;-\; 2z)\, w(y,z) \;=\; 0 \tag{2.215}$$

which can be readily verified using (2.194). Substituting expansion (2.196) into the differential equation (2.215) yields

$$\sum_{n=1}^{\infty} \frac{z^{n-1}}{(n-1)!}\, H_n(y) \;-\; 2y \sum_{n=0}^{\infty} \frac{z^n}{n!}\, H_n(y) \;+\; 2 \sum_{n=0}^{\infty} \frac{z^{n+1}}{n!}\, H_n(y) \;=\; 0\,. \tag{2.216}$$

Combining the sums and collecting terms with identical powers of z

$$\sum_{n=1}^{\infty} \frac{z^n}{n!} \Big[\, H_{n+1}(y) \;-\; 2y\, H_n(y) \;+\; 2n\, H_{n-1}(y) \Big] \;+\; H_1(y) \;-\; 2yH_0(y) \;=\; 0 \tag{2.217}$$

gives

$$H_1(y) \;-\; 2y\, H_0(y) \;=\; 0,\; H_{n+1}(y) \;-2y\, H_n(y) \;+\; 2n\, H_{n-1}(y) \;=\; 0,\; n = 1,2,\ldots \tag{2.218}$$

The reader should recognize the connection between the pattern of the differential equation (**??**) and the pattern of the recursion equation (**??**): a differential operator d/dz increases the order n of H_n by one, a factor z reduces the order of H_n by one and introduces also a factor n. One can then readily state which differential equation of $w(y,z)$ should be equivalent to the relationship (**??**), namely, $dw/dy - 2zw = 0$. The reader may verify that $w(y,z)$, as given in (2.194), indeed satisfies the latter relationship.

Integral Representation of Hermite Polynomials

An integral representation of the *Hermite polynomials* can be derived starting from the integral

$$I(y) \;=\; \int_{-\infty}^{+\infty} dt\; e^{2iy\,t\,-\,t^2}\;. \tag{2.219}$$

which can be written

$$I(y) \;=\; e^{-y^2}\;\int_{-\infty}^{+\infty} dt e^{-(t-iy)^2} \;=\; e^{-y^2}\;\int_{-\infty}^{+\infty} dz e^{-z^2}\;. \tag{2.220}$$

Using (2.247) for $a = i$ one obtains

$$I(y) \;=\; \sqrt{\pi}\; e^{-y^2} \tag{2.221}$$

and, acording to the definition (2.226a),

$$e^{-y^2} \;=\; \frac{1}{\sqrt{\pi}}\;\int_{-\infty}^{+\infty} dt\; e^{2iyt\,-\,t^2}\;. \tag{2.222}$$

Employing this expression now on the r.h.s. of the *Rodrigues* formula (2.200) yields

$$H_n(y) \;=\; \frac{(-1)^n}{\sqrt{\pi}}\; e^{y^2}\;\int_{-\infty}^{+\infty} dt\;\frac{d^n}{dy^n}\; e^{2iyt\,-\,t^2}\;. \tag{2.223}$$

The identity

$$\frac{d^n}{dy^n}\, e^{2iy\,t\,-\,t^2} \;=\; (2\,i\,t)^n\; e^{2iy\,t\,-\,t^2} \tag{2.224}$$

results, finally, in the integral representation of the *Hermite* polynomials

$$H_n(y) \;=\; \frac{2^n\,(-i)^n e^{y^2}}{\sqrt{\pi}}\;\int_{-\infty}^{+\infty} dt\, t^n\, e^{2iy\,t\,-\,t^2}\;, \quad n \;=\; 0, 1, 2, \ldots \tag{2.225}$$

Orthonormality Properties

We want to derive from the generating function (2.194, 2.196) the orthogonality properties of the *Hermite* polynomials. For this purpose we consider the integral

$$\begin{aligned}\textstyle\int_{-\infty}^{+\infty} dy\, w(y,z)\, w(y,z')\, e^{-y^2} \quad &=\; e^{2\,z\,z'}\;\int_{-\infty}^{+\infty} dy\, e^{-(y-z-z')^2} \;=\; \sqrt{\pi}\, e^{2\,z\,z'} \\ &=\; \sqrt{\pi}\;\sum_{n=0}^{\infty}\;\frac{2^n\; z^n\; z'^n}{n!}\;.\end{aligned} \tag{2.226}$$

Expressing the l.h.s. through a double series over *Hermite* polynomials using (2.194, 2.196) yields

$$\sum_{n,n'=0}^{\infty}\;\int_{-\infty}^{+\infty} dy\, H_n(y)\, H_{n'}(y)\, e^{-y^2}\;\frac{z^n\; z'^{n'}}{n!\; n'!} \;=\; \sum_{n=0}^{\infty}\; 2^n n!\,\sqrt{\pi}\;\frac{z^n\; z'^n}{n!\; n!} \tag{2.227}$$

Comparing the terms of the expansions allows one to conclude the orthonormality conditions

$$\int_{-\infty}^{+\infty} dy\, H_n(y)\, H_{n'}(y)\; e^{-y^2} \;=\; 2^n\, n!\,\sqrt{\pi}\;\;\delta_{n\,n'}\;. \tag{2.228}$$

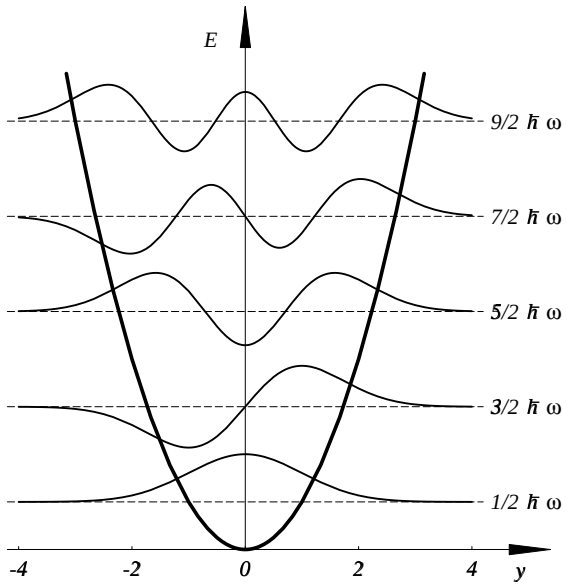


Figure 2.2: Stationary states $\phi_n(y)$ of the harmonic oscillator for $n = 0, 1, 2, 3, 4$.

Normalized Stationary States

The orthonormality conditions (2.228) allow us to construct normalized stationary states of the harmonic oscillator. According to (2.197) holds

$$\tilde{\phi}_n(y) \;=\; e^{-y^2/2}\, H_n(y)\;. \tag{2.229}$$

The normalized states are [c.f. (2.189, 2.190)]

$$\phi_n(y) \;=\; N_n\, e^{-y^2/2}\, H_n(y)\;. \tag{2.230}$$

and for the normalization constants N_n follows from (2.228)

$$N_n^2 \int_{-\infty}^{+\infty} dy\, e^{-y^2}\, H_n^2(y) \;=\; N_n^2\, 2^n\, n!\, \sqrt{\pi} \;=\; 1 \tag{2.231}$$

We conclude

$$N_n \;=\; \frac{1}{\sqrt{2^n\, n!\, \sqrt{\pi}}} \tag{2.232}$$

and can finally state the explicit form of the normalized stationary states

$$\phi_n(y) \;=\; \frac{1}{\sqrt{2^n\, n!\, \sqrt{\pi}}}\; e^{-y^2/2}\, H_n(y)\;. \tag{2.233}$$

The stationary states (2.233) are presented for $n = 0, 1, 2, 3, 4$ in Fig. 2.2. One can recognize, in agreement with our above discussions, that the wave functions are even for $n = 0, 2, 4$ and odd for $n = 1, 3$. One can also recognize that n is equal to the number of nodes of the wave function. Furthermore, the value of the wave function at $y = 0$ is positive for $n = 0, 4$, negative for $n = 2$ and vanishes for $n = 1, 3$, in harmony with (**??**).

The normalization condition (2.231) of the wave functions differs from that postulated in (2.189) by the Jacobian dx/dy, i.e., by

$$\sqrt{\left|\frac{dx}{dy}\right|} \;=\; \left[\frac{m\omega}{\hbar}\right]^{\frac{1}{4}} \;. \tag{2.234}$$

The explicit form of the stationary states of the harmonic oscillator in terms of the position variable x is then, using (2.233) and (2.189)

$$\phi_n(x) \;=\; \frac{1}{\sqrt{2^n\, n!}}\, \left[\frac{m\omega}{\pi\hbar}\right]^{\frac{1}{4}} e^{-\frac{m\omega x^2}{2\hbar}}\, H_n(\sqrt{\frac{m\omega}{\hbar}}x) \;. \tag{2.235}$$

Completeness of the Hermite Polynomials

The *Hermite* polynomials are the first members of a large class of *special functions* which one encounters in the course of describing stationary quantum states for various potentials and in spaces of different dimensions. The *Hermite* polynomials are so-called orthonogal polynomials since they obey the conditions (2.228). The various orthonogal polynomials differ in the spaces $\Omega \subset \mathbb{R}$ over which they are defined and differ in a weight function $w(y)$ which enter in their orthonogality conditions. The latter are written for polynomials $p_n(x)$ in the general form

$$\int_\Omega dx\; p_n(x)\; p_m(x)\; w(x) \;=\; I_n\; \delta_{nm} \tag{2.236}$$

where $w(x)$ is a so-called weight function with the property

$$w(x) \;\geq\; 0,\; w(x) \;=\; 0 \quad \text{only at a discrete set of points } x_k \;\in\; \Omega \tag{2.237}$$

and where I_n denotes some constants. Comparision with (2.228) shows that the orthonogality condition of the *Hermite* polynomials is in complience with (2.236 , 2.237) for $\Omega \;=\; \mathbb{R}$, $w(x) \;=\; \exp(-x^2)$, and $I_n \;=\; 2^n n!\sqrt{\pi}$.

Other examples of orthogonal polynomials are the *Legendre* and *Jacobi* polynomials which arise in solving three-dimensional stationary Schrödinger equations, the *ultra-spherical harmonics* which arise in n–dimensional Schrödinger equations and the associated *Laguerre* polynomials which arise for the stationary quantum states of particles moving in a Coulomb potential. In case of the Legendre polynomials, denoted by $P_\ell(x)$ and introduced in Sect. 5 below [c.f. (5.150 , 5.151, 5.156, 5.179] holds $\Omega \;=\; [-1,1]$, $w(x) \;\equiv\; 1$, and $I_\ell \;=\; 2/(2\ell+1)$. In case of the associated *Laguerre* polynomials, denoted by $L_n^{(\alpha)}(x)$ and encountered in case of the stationary states of the non-relativistic [see Sect. ??? and eq. ???] and relativistic [see Sect. 10.10 and eq. (10.459] hydrogen atom, holds $\Omega \;=\; [0,+\infty[$, $w(x) \;=\; x^\alpha e^{-x}$, $I_n \;=\; \Gamma(n+\alpha+1)/n!$ where $\Gamma(z)$ is the so-called *Gamma function*.

The orthogonal polynomials p_n mentioned above have the important property that they form a *complete* basis in the space $\mathcal{F}$ of normalizable functions, i.e., of functions which obey

$$\int_\Omega dx\; f^2(x)\; w(x) \;=< \infty \;, \tag{2.238}$$

where the space is endowed with the scalar product

$$(f|g) \;=\; \int_\Omega dx\; f(x)\, g(x)\; w(x) \;=< \infty \;, \quad f,\, g \;\in\; \mathcal{F} \;. \tag{2.239}$$

As a result holds for any $f \in \mathcal{F}$

$$f(x) \;=\; \sum_n c_n \, p_n(x) \tag{2.240}$$

where

$$c_n \;=\; \frac{1}{I_n} \int_\Omega dx \; w(x) \, f(x) \, p_n(x) \; . \tag{2.241}$$

The latter identity follows from (2.236). If one replaces for all $f \in \mathcal{F}$: $f(x) \rightarrow \sqrt{w(x)}\, f(x)$ and, in particular, $p_n(x) \rightarrow \sqrt{w(x)}\, p_n(x)$ the scalar product (2.239) becomes the conventional scalar product of quantum mechanics

$$\langle f|g\rangle \;=\; \int_\Omega dx \; f(x) \, g(x) \; . \tag{2.242}$$

Let us assume now the case of a function space governed by the norm (2.242) and the existence of a normalizable state $\psi(y,t)$ which is stationary under the action of the harmonic oscillator propagator (2.147), i.e., a state for which (2.172) holds. Since the *Hermite* polynomials form a complete basis for such states we can expand

$$\psi(y,t) \;=\; \sum_{n=0}^{\infty} c_n(t) \, e^{-y^2/2} \, H_n(y) \; . \tag{2.243}$$

To be consistent with(2.188, 2.197) it must hold $c_n(t) \;=\; d_n \, \exp[-i\omega(n+\frac{1}{2})t]$ and, hence, the stationary state $\psi(y,t)$ is

$$\psi(y,t) \;=\; \sum_{n=0}^{\infty} d_n \, \exp[-i\omega(n+\frac{1}{2})t] \, e^{-y^2/2} \, H_n(y) \; . \tag{2.244}$$

For the state to be stationary $|\psi(x,t)|^2$, i.e.,

$$\sum_{n,m=0}^{\infty} d_n^* d_m \, \exp[i\omega(m-n)t] \, e^{-y^2} \, H_n(y) H_m(y) \; , \tag{2.245}$$

must be time-independent. The only possibility for this to be true is $d_n \;=\; 0$, except for a single $n \;=\; n_o$, i.e., $\psi(y,t)$ must be identical to one of the stationary states (2.233). Therefore, the states (2.233) exhaust all stationary states of the harmonic oscillator.

Appendix: Exponential Integral

We want to prove

$$I \;=\; \int_{-\infty}^{+\infty} dy_1 \; \ldots \int_{-\infty}^{+\infty} dy_n \;\; e^{i \sum_{j,k}^n y_j a_{jk} y_k} \;=\; \sqrt{\frac{(i\pi)^n}{det(\mathbf{a})}} \; , \tag{2.246}$$

for $det(\mathbf{a}) \neq 0$ and real, symmetric $\mathbf{a}$, i.e. $\mathbf{a}^T = \mathbf{a}$. In case of $n \;=\; 1$ this reads

$$\int_{-\infty}^{+\infty} dx \, e^{i\,a\,x^2} \;=\; \sqrt{\frac{i\,\pi}{a}} \; , \tag{2.247}$$

which holds for $a \in \mathbb{C}$ as long as $a \neq 0$.
The proof of (2.246) exploits that for any real, symmetric matrix exists a similarity transformation such that

$$\mathbf{S}^{-1}\mathbf{a}\,\mathbf{S} \;=\; \tilde{\mathbf{a}} \;=\; \begin{pmatrix} \tilde{a}_{11} & 0 & \dots & 0 \\ 0 & \tilde{a}_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \tilde{a}_{nn} \end{pmatrix} . \tag{2.248}$$

where $\mathbf{S}$ can be chosen as an orthonormal transformation, i.e.,

$$\mathbf{S}^T\mathbf{S} \;=\; \mathbb{1} \qquad \text{or} \qquad \mathbf{S} \;=\; \mathbf{S}^{-1} . \tag{2.249}$$

The $\tilde{a}_{kk}$ are the eigenvalues of $\mathbf{a}$ and are real. This property allows one to simplify the bilinear form $\sum_{j,k}^{n} y_j a_{jk} y_k$ by introducing new integration variables

$$\tilde{y}_j \;=\; \sum_{k}^{n} (S^{-1})_{jk} y_k \; ; \quad y_k \;=\; \sum_{k}^{n} S_{kj} \tilde{y}_j \; . \tag{2.250}$$

The bilinear form in (2.246) reads then in terms of $\tilde{y}_j$

$$\begin{aligned} \textstyle\sum_{j,k}^{n} y_j a_{jk} y_k \quad &= \sum_{j,k}^{n} \sum_{\ell m}^{n} \tilde{y}_\ell S_{j\ell} a_{jk} S_{km} \tilde{y}_m \\ &= \sum_{j,k}^{n} \sum_{\ell m}^{n} \tilde{y}_\ell (S^T)_{\ell j} a_{jk} S_{km} \tilde{y}_m \\ &= \sum_{j,k}^{n} \tilde{y}_j \tilde{a}_{jk} \tilde{y}_k \end{aligned} \tag{2.251}$$

where, according to (2.248, 2.249)

$$\tilde{a}_{jk} \;=\; \sum_{l,m}^{n} (S^T)_{jl} a_{lm} S_{mk} \; . \tag{2.252}$$

For the determinant of $\tilde{\mathbf{a}}$ holds

$$det(\tilde{\mathbf{a}}) = \prod_{j=1}^{n} \tilde{a}_{jj} \tag{2.253}$$

as well as

$$\begin{aligned} det(\tilde{\mathbf{a}}) \;&=\; det(\mathbf{S}^{-1}\mathbf{a}\mathbf{S}) \;=\; det(\mathbf{S}^{-1})\, det(\mathbf{a})\; , det(\mathbf{S}) \\ &=\; (det(\mathbf{S}))^{-1}\; det(\mathbf{a})\; det(\mathbf{S}) \;=\; det(\mathbf{a}) \; . \end{aligned} \tag{2.254}$$

One can conclude

$$det(\mathbf{a}) \;=\; \prod_{j=1}^{n} \tilde{a}_{jj} \; . \tag{2.255}$$

We have assumed $det(\mathbf{a}) \neq 0$. Accordingly, holds

$$\prod_{j=1}^{n} \tilde{a}_{jj} \neq 0 \tag{2.256}$$

such that none of the eigenvalues of $\mathbf{a}$ vanishes, i.e.,

$$\tilde{a}_{jj} \neq 0 \, , \quad \text{for } j \, = \, 1, \, 2, \, \ldots , \, n \tag{2.257}$$

Substitution of the integration variables (2.250) allows one to express (2.250)

$$I \;\; = \;\; \int_{-\infty}^{+\infty} d\tilde{y}_1 \;\; \ldots \;\; \int_{-\infty}^{+\infty} d\tilde{y}_n \;\; \left| det \left(\frac{\partial (y_1, \ldots , y_n)}{\partial (\tilde{y}_1, \ldots , \tilde{y}_n)} \right) \right| \;\; e^{i \sum_k^n \tilde{a}_{kk} \tilde{y}_k^2} \; . \tag{2.258}$$

where we introduced the Jacobian matrix

$$\mathbf{J} \;\; = \;\; \frac{\partial (y_1, \ldots , y_n)}{\partial (\tilde{y}_1, \ldots , \tilde{y}_n)} \tag{2.259}$$

with elements

$$J_{js} \;\; = \;\; \frac{\partial y_j}{\partial \tilde{y}_s} \; . \tag{2.260}$$

According to (2.250) holds

$$\mathbf{J} \;\; = \;\; \mathbf{S} \tag{2.261}$$

and, hence,

$$det(\frac{\partial (y_1, \ldots , y_n)}{\partial (\tilde{y}_1, \ldots , \tilde{y}_n)}) \;\; = \;\; det(\mathbf{S}) \; . \tag{2.262}$$

From (2.249) follows

$$1 \;\; = \;\; \det \left(\mathbf{S}^T \mathbf{S} \right) \;\; = \;\; (\det \mathbf{S})^2 \tag{2.263}$$

such that one can conclude

$$\det \mathbf{S} \;\; = \;\; \pm 1 \tag{2.264}$$

One can right then (2.258)

$$\begin{aligned} I \;\; &= \;\; \int_{-\infty}^{+\infty} d\tilde{y}_1 \;\; \ldots \;\; \int_{-\infty}^{+\infty} d\tilde{y}_n \;\; e^{i \sum_k^n \tilde{a}_{kk} \tilde{y}_k^2} \\ &= \;\; \int_{-\infty}^{+\infty} d\tilde{y}_1 \; e^{i \tilde{a}_{11} \tilde{y}_1^2} \;\; \ldots \;\; \int_{-\infty}^{+\infty} d\tilde{y}_n \; e^{i \tilde{a}_{nn} \tilde{y}_n^2} \;\; = \;\; \prod_{k=1}^{n} \int_{-\infty}^{+\infty} d\tilde{y}_k \; e^{i \tilde{a}_{kk} \tilde{y}_k^2} \end{aligned} \tag{2.265}$$

which leaves us to determine integrals of the type

$$\int_{-\infty}^{+\infty} dx \; e^{icx^2} \tag{2.266}$$

where, according to (2.257) holds $c \neq 0$.
We consider first the case $c > 0$ and discuss the case $c < 0$ further below. One can relate integral (2.266) to the well-known Gaussian integral

$$\int_{-\infty}^{+\infty} dx\ e^{-cx^2} \quad = \quad \sqrt{\frac{\pi}{c}}\ , c > 0\ . \tag{2.267}$$

by considering the contour integral

$$J \ = \ \oint_{\gamma} dz\ e^{icz^2} = 0 \tag{2.268}$$

along the path $\gamma = \gamma_1 + \gamma_2 + \gamma_3 + \gamma_4$ displayed in Figure 2.3. The contour integral (2.268) vanishes, since e^{icz^2} is a holomorphic function, i.e., the integrand does not exhibit any singularities anywhere in $\mathbb{C}$. The contour intergral (2.268) can be written as the sum of the following path integrals

$$J \ = \ J_1 \ + \ J_2 \ + \ J_3 \ + \ J_4\ ; \quad J_k \ = \ \oint_{\gamma_k} dz\ e^{icz^2} \tag{2.269}$$

The contributions J_k can be expressed through integrals along a real coordinate axis by realizing that the paths γ_k can be parametrized by real coordinates x

$$\begin{aligned}
&\gamma_1 : z = x && J_1 = \int_{-p}^{p} dx\ e^{icx^2} \\
&\gamma_2 : z = ix + p && J_2 = \int_{0}^{p} i\ dx\ e^{ic(ix+p)^2} \\
&\gamma_3 : z = \sqrt{i}\ x && J_3 = \int_{\sqrt{2}p}^{-\sqrt{2}p} \sqrt{i}\ dx\ e^{ic(\sqrt{i}x)^2} \\
& && \quad = \ -\sqrt{i} \int_{-\sqrt{2}p}^{\sqrt{2}p} dx\ e^{-cx^2} \\
&\gamma_4 : z = ix - p && J_4 = \int_{-p}^{0} i\ dx\ e^{ic(ix-p)^2}\ , \\
& && \text{for } x, p\ \in\ \mathbb{R}.
\end{aligned} \tag{2.270}$$

Substituting $-x$ for x into integral J_4 one obtains

$$\begin{aligned}
J_4 \ &= \ \int_{p}^{0} (-i)\ dx\ e^{ic(-ix+p)^2} \\
&= \ \int_{0}^{p} i\ dx\ e^{ic(ix-p)^2} \quad = \quad J_2\ .
\end{aligned} \tag{2.271}$$

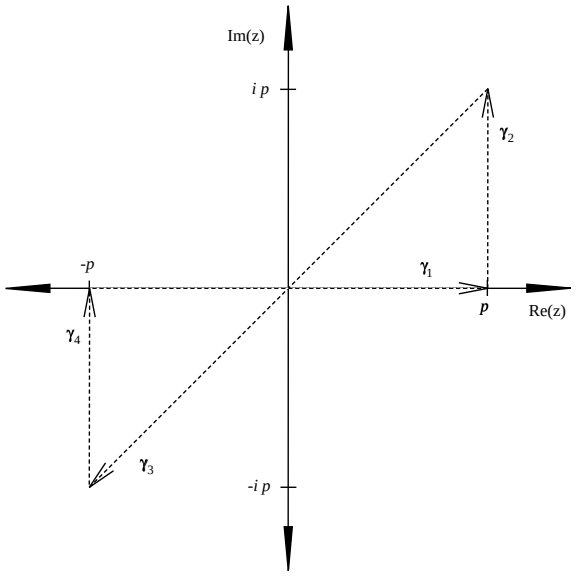


Figure 2.3: Contour path γ in the complex plain.

We will now show that the two integrals J_2 and J_4 vanish for $p \to +\infty$. This follows from the following calculation

$$\begin{aligned} \lim_{p\to+\infty} |J_{2 \text{ or } 4}| &= \lim_{p\to+\infty} |\int_0^p i\ dx\ e^{ic(ix+p)^2}| \\ &\leq \lim_{p\to+\infty} \int_0^p |i|\ dx\ |e^{ic(p^2-x^2)}|\ |e^{-2cxp}|\ . \end{aligned} \tag{2.272}$$

It holds $|e^{ic(p^2-x^2)}| = 1$ since the exponent of e is purely imaginary. Hence,

$$\begin{aligned} \lim_{p\to+\infty} |J_{2 \text{ or } 4}| &\leq \lim_{p\to+\infty} \int_0^p dx\ |e^{-2cxp}| \\ &= \lim_{p\to+\infty} \frac{1-e^{-2cp}}{2\ c\ p} = 0\ . \end{aligned} \tag{2.273}$$

J_2 and J_4 do not contribute then to integral (2.268) for $p = +\infty$. One can state accordingly

$$J = \int_{-\infty}^{\infty} dx\ e^{icx^2} - \sqrt{i} \int_{-\infty}^{\infty} dx\ e^{-cx^2} = 0\ . \tag{2.274}$$

Using 2.267) one has shown then

$$\int_{-\infty}^{\infty} dx\ e^{icx^2} = \sqrt{\frac{i\pi}{c}}\ . \tag{2.275}$$

One can derive the same result for $c < 0$, if one chooses the same contour integral as (2.268), but with a path γ that is reflected at the real axis. This leads to

$$J \;=\; \int_{-\infty}^{\infty} dx\; e^{icx^2} \;+\; \sqrt{-i}\; \int_{\infty}^{-\infty} dx\; e^{cx^2} \;=\; 0 \tag{2.276}$$

and ($c < 0$)

$$\int_{-\infty}^{\infty} dx\; e^{icx^2} \;=\; \sqrt{\frac{-i\pi}{-|c|}} \;=\; \sqrt{\frac{i\pi}{c}} \quad . \tag{2.277}$$

We apply the above results (2.275, 2.277) to (2.265). It holds

$$I \;=\; \prod_{k=1}^{n} \sqrt{\frac{i\pi}{\tilde{a}_{kk}}} = \sqrt{\frac{(i\pi)^n}{\prod_{j=1}^{n} \tilde{a}_{jj}}} \quad . \tag{2.278}$$

Noting (2.255) this result can be expressed in terms of the matrix $\mathbf{a}$

$$I \;=\; \sqrt{\frac{(i\pi)^n}{det(\mathbf{a})}} \tag{2.279}$$

which concludes our proof.

Chapter 3

The Schrödinger Equation

3.1 Derivation of the Schrödinger Equation

We will consider now the propagation of a wave function $\psi(\vec{r},t)$ by an infinitesimal time step ϵ. It holds then according to (2.5)

$$\psi(\vec{r},t+\epsilon) \;=\; \int_\Omega d^3r_0\, \phi(\vec{r},t+\epsilon|\vec{r}_0,t)\, \psi(\vec{r}_0,t)\;. \tag{3.1}$$

We will expand the l.h.s. and the r.h.s. of this equation in terms of powers of ϵ and we will demonstrate that the terms of order ϵ require that $\psi(\vec{r},t)$ satisfies a partial differential equation, namely the Schrödinger equation. For many situations, but by no means all, the Schrödinger equation provides the simpler avenue towards describing quantum systems than the path ingral formulation of Section 2. Notable exceptions are non-stationary systems involving time-dependent linear and quadratic Lagrangians.
The propagator in (3.1) can be expressed through the discretization scheme (2.20, 2.21). In the limit of very small ϵ it is sufficient to employ a single discretization interval in (2.20) to evaluate the propagator. Generalizing (2.20) to $\mathbb{R}^{\not\vdash}$ one obtains then for small ϵ

$$\phi(\vec{r},t+\epsilon|\vec{r}_0,t) \;=\; \Big[\frac{m}{2\pi i\hbar\epsilon}\Big]^{\frac{3}{2}}\, \exp\left\{\frac{i}{\hbar}\left[\frac{m}{2}\frac{(\vec{r}-\vec{r}_0)^2}{\epsilon} \;-\; \epsilon\, U(\vec{r},t)\right]\right\}\;. \tag{3.2}$$

From this follows

$$\psi(\vec{r},t+\epsilon) \;=\; \int_\Omega d^3r_0\, \Big[\frac{m}{2\pi i\hbar\epsilon}\Big]^{\frac{3}{2}}\, \exp\left\{\frac{i}{\hbar}\left[\frac{m}{2}\frac{(\vec{r}-\vec{r}_0)^2}{\epsilon} \;-\; \epsilon\, U(\vec{r},t)\right]\right\}\, \psi(\vec{r}_0,t)\;. \tag{3.3}$$

In order to carry out the integration we set $\vec{r}_0 \;=\; \vec{r}+\vec{s}$ and use $\vec{s}$ as the new integration variable. We will denote the components of $\vec{s}$ by $(x_1,x_2,x_3)^T$. This yields

$$\begin{aligned}\psi(\vec{r},t+\epsilon) \;&=\; \textstyle\int_{-\infty}^{+\infty} dx_1 \int_{-\infty}^{+\infty} dx_2 \int_{-\infty}^{+\infty} dx_3\, \Big[\frac{m}{2\pi i\hbar\epsilon}\Big]^{\frac{3}{2}} \;\times \\ \times\; \underbrace{\exp\left(\frac{i}{\hbar}\,\frac{m}{2}\frac{x_1^2+x_2^2+x_3^2}{\epsilon} \;-\; \epsilon\, U(\vec{r},t)\right)}_{\text{even in } x_1,x_2,,x_3}\; &\psi(\vec{r}+\vec{s},t)\;.\end{aligned} \tag{3.4}$$

It is important to note that the integration is **not** over $\vec{r}$, but over $\vec{s} = (x_1, x_2, x_3)^T$, e.g. $U(\vec{r}, t)$ is a constant with respect to this integration. The integration involves only the wave function $\psi(\vec{r}+\vec{s}, t)$ and the kinetic energy term. Since the latter contributes to (3.4) only for small $x_1^2 + x_2^2 + x_3^2$ values we expand

$$\psi(\vec{r}+\vec{s}, t) \;=\; \psi(\vec{r}, t) \;+\; \sum_{j=1}^{3} x_j \frac{\partial}{\partial x_j}\psi(\vec{r}, t) \;+\; \frac{1}{2}\sum_{j,k=1}^{3} x_j x_k \frac{\partial^2}{\partial x_j \partial x_k}\psi(\vec{r}, t) \;+\; \ldots \tag{3.5}$$

assuming that only the leading terms contribute, a supposition which will be examined below. Since the kinetic energy contribution in (3.4) is even in all three coordinates x_1, x_2, x_3, only terms of the expansion (3.5) which are even separately in all three coordinates yield non-vanishing contributions. It is then sufficient to consider the terms

$$\psi(\vec{r}, t) \quad ; \quad \tfrac{1}{2}\textstyle\sum_{j=1}^{3} x_j^2 \frac{\partial^2}{\partial x_j^2}\psi(\vec{r}, t) \quad ; \quad \tfrac{1}{4}\textstyle\sum_{j,k=1}^{3} x_j^2 x_k^2 \frac{\partial^4}{\partial x_j^2 \partial x_k^2}\psi(\vec{r}, t) \quad ;$$
$$\tfrac{1}{12}\textstyle\sum_{j=1}^{3} x_j^4 \frac{\partial^4}{\partial x_j^4}\psi(\vec{r}, t) \quad ; \ldots \tag{3.6}$$

of the expansion of $\psi(\vec{r}+\vec{s}, t)$.
Obviously, we need then to evaluate integrals of the type

$$I_n(a) \;=\; \int_{-\infty}^{+\infty} dx\, x^{2n} \exp\left(i\, a\, x^2\right) \;, \quad n \;=\; 0, 1, 2 \tag{3.7}$$

According to (2.36) holds

$$I_0(a) \;=\; \sqrt{\frac{i\pi}{a}} \tag{3.8}$$

Inspection of (3.7) shows

$$I_{n+1}(a) \;=\; \frac{1}{i}\,\frac{\partial}{\partial a} I_n(a). \tag{3.9}$$

Starting from (3.8) one can evaluate recursively all integrals $I_n(a)$. It holds

$$I_1(a) \;=\; \frac{i}{2a}\sqrt{\frac{i\pi}{a}}\,; \quad I_2(a) \;=\; -\frac{3}{4a^2}\sqrt{\frac{i\pi}{a}}\,, \ldots \tag{3.10}$$

It is now important to note that in case of integral (3.4) one identifies

$$\frac{1}{a} \;=\; \frac{2\,\epsilon\,\hbar}{m} \;=\; \mathcal{O}(\epsilon) \tag{3.11}$$

and, consequently, the terms collected in (3.6) make contributions of the order

$$\mathcal{O}(\epsilon^{\frac{3}{2}})\;,\; \mathcal{O}(\epsilon^{\frac{5}{2}})\;,\; \mathcal{O}(\epsilon^{\frac{7}{2}})\;,\; \mathcal{O}(\epsilon^{\frac{7}{2}})\;. \tag{3.12}$$

Here one needs to note that we are actually dealing with a three-fold integral. According to (3.11) holds

$$\left[\frac{m}{2\pi i\hbar\epsilon}\right]^{\frac{3}{2}} \times \left[\frac{i\,\pi}{a}\right]^{\frac{3}{2}} \;=\; 1 \tag{3.13}$$

and one can conclude, using (3.10),

$$\psi(\vec{r}, t+\epsilon) \;=\; \exp\left[-\frac{i\epsilon}{\hbar}\,U(\vec{r},t)\right]\,\left[\psi(\vec{r},t) \;+\; \frac{1}{4}\,\frac{2i\epsilon\hbar}{m}\,\nabla^2\psi(\vec{r},t) \;+\; \mathcal{O}(\epsilon^2)\right]\,. \tag{3.14}$$

This expansion in terms of powers of ϵ suggests that we also expand

$$\psi(\vec{r}, t+\epsilon) \;=\; \psi(\vec{r},t) \;+\; \epsilon\frac{\partial}{\partial t}\psi(\vec{r},t) \;+\; \mathcal{O}(\epsilon^2) \tag{3.15}$$

and

$$\exp\left[-\frac{i\epsilon}{\hbar}\,U(\vec{r},t)\right] \;=\; 1 \;-\; \frac{i\epsilon}{\hbar}\,U(\vec{r},t) \;+\; \mathcal{O}(\epsilon^2)\,. \tag{3.16}$$

Inserting this into (3.14) results in

$$\begin{aligned}\psi(\vec{r},t) \;+\; \epsilon\tfrac{\partial}{\partial t}\psi(\vec{r},t) \quad &=\; \psi(\vec{r},t) \;-\; \frac{i\epsilon}{\hbar}\,U(\vec{r},t)\psi(\vec{r},t) \\ &\quad+\; \frac{i\epsilon}{\hbar}\,\frac{\hbar^2}{2m}\,\nabla^2\psi(\vec{r},t) \;+\; \mathcal{O}(\epsilon^2)\,.\end{aligned} \tag{3.17}$$

Obviously, this equation is trivially satisfied to order $\mathcal{O}(\epsilon^0)$. In order $\mathcal{O}(\epsilon)$ the equation reads

$$i\hbar\frac{\partial}{\partial t}\,\psi(\vec{r},t) \;=\; \left[-\frac{\hbar^2}{2m}\,\nabla^2 \;+\; U(\vec{r},t)\right]\,\psi(\vec{r},t)\,. \tag{3.18}$$

This is the celebrated *time-dependent Schrödinger equation*. This equation is often written in the form

$$i\hbar\frac{\partial}{\partial t}\,\psi(\vec{r},t) \;=\; \hat{H}\,\psi(\vec{r},t) \tag{3.19}$$

where

$$\hat{H} \;=\; -\frac{\hbar^2}{2m}\,\nabla^2 \;+\; U(\vec{r},t)\,. \tag{3.20}$$

3.2 Boundary Conditions

The time-dependent Schrödinger equation is a partial differential equation, 1st order in time, 2nd order in the spatial variables and linear in the solution $\psi(\vec{r},t)$. The following general remarks can be made about the solution.

Due to its linear character any linear combination of solutions of the time-dependent Schrödinger equation is also a solution.

The 1st order time derivative requires that for any solution a single temporal condition needs to be specified, e.g., $\psi(\vec{r},t_1) = f(\vec{r})$. Usually, one specifies the so-called initial condition, i.e., a solution is thought for $t \geq t_0$ and the solution is specified at the intial time t_0.

The 2nd order spatial derivatives require that one specifies also properties of the solution on a closed boundary $\partial\Omega$ surrounding the volume Ω in which a solution is to be determined. We will derive briefly the type of boundary conditions encountered. As we will discuss in Chapter 5 below the solutions of the Schrödinger equation are restricted to particular Hilbert spaces $\mathbb{H}$ which are

linear vector spaces of functions $f(\vec{r})$ in which a scalar product between two elements $f, g \in \mathbb{H}$ is defined as follows

$$\langle f|g\rangle_\Omega \;=\; \int_\Omega d^3r f^*(\vec{r}) g(\vec{r}) \tag{3.21}$$

This leads one to consider the integral

$$\langle f|\, H\, |g\rangle_\Omega \;=\; \int_\Omega d^3r f^*(\vec{r})\, \hat{H}\, g(\vec{r}) \tag{3.22}$$

where $\hat{H}$ is defined in (3.20). Interchanging $f^*(\vec{r})$ and $g(\vec{r})$ results in

$$\overline{\langle g|\, H\, |f\rangle_\Omega} \;=\; \int_\Omega d^3r\, g(\vec{r})\, \hat{H} f^*(\vec{r})\;. \tag{3.23}$$

Since $\hat{H}$ is a differential operator the expressions (3.22) and (3.23), in principle, differ from each other. The difference between the integrals is

$$\begin{aligned}
\langle g|\, H\, |f\rangle_\Omega \;&= -\,\overline{\langle g|\, H\, |f\rangle_\Omega} \\
&= \int_\Omega d^3r f^*(\vec{r}) \left(-\frac{\hbar^2}{2m}\nabla^2\right) g(\vec{r}) \;-\; \int_\Omega d^3r g(\vec{r}) \left(-\frac{\hbar^2}{2m}\nabla^2\right) f^*(\vec{r}) \\
&\quad + \int_\Omega d^3r f^*(\vec{r})\, U(\vec{r},t)\, g(\vec{r}) \;-\; \int_\Omega d^3r g(\vec{r})\, U(\vec{r},t), f^*(\vec{r}) \\
&= -\frac{\hbar^2}{2m} \int_\Omega d^3r f^*(\vec{r}) \left(\nabla^2\right) g(\vec{r}) \;-\; \int_\Omega d^3r g(\vec{r}) \left(\nabla^2\right) f^*(\vec{r})
\end{aligned} \tag{3.24}$$

Using Green's theorem[1]

$$\begin{aligned}
&\textstyle\int_\Omega d^3r \,\left(f^*(\vec{r})\,\nabla^2 g(\vec{r}) \;-\; g(\vec{r})\,\nabla^2 f^*(\vec{r})\right) \\
&= \textstyle\int_{\partial\Omega} d\vec{a} \cdot \left(f^*(\vec{r})\nabla g(\vec{r}) \;-\; g(\vec{r})\nabla f^*(\vec{r})\right)
\end{aligned} \tag{3.25}$$

one obtains the identity

$$\langle f|\, \hat{H}\, |g\rangle_\Omega \;=\; \overline{\langle g|\, \hat{H}\, |f\rangle_\Omega} \;+\; \int_{\partial\Omega} d\vec{a} \cdot \vec{P}(f^*, g|\vec{r}) \tag{3.26}$$

where $\int_{\partial\Omega} d\vec{a} \cdot \vec{A}(\vec{r})$ denotes an integral over the surface $\partial\Omega$ of the volume Ω, the surface elements $d\vec{a}$ pointing out of the surface in a direction normal to the surface and the vector–valued function $\vec{A}(\vec{r})$ is taken at points $\vec{r} \in \partial\Omega$. In (3.26) the vector–valued function $\vec{P}(f^*, g|\vec{r})$ is called the *concomitant* of $\hat{H}$ and is

$$\vec{P}(f^*, g|\vec{r}) \;=\; -\frac{\hbar^2}{2m} \left(f^*(\vec{r})\,\nabla g(\vec{r}) \;-\; g(\vec{r})\,\nabla f^*(\vec{r})\right) \tag{3.27}$$

We will postulate below that $\hat{H}$ is an operator in $\mathbb{H}$ which represents energy. Since energy is a real quantity one needs to require that the eigenvalues of the operator $\hat{H}$ are real and, hence, that $\hat{H}$ is hermitian[2]. The hermitian property, however, implies

$$\langle f|\, \hat{H}\, |g\rangle_\Omega \;=\; \overline{\langle g|\, \hat{H}\, |f\rangle_\Omega} \tag{3.28}$$

[1] See, for example, *Calssical Electrodynamics, 2nd Ed.*, J. D. Jackson, (John Wiley, New York, 1975), Chapter 1.

[2] The reader is advised to consult a reference text on 'Linear Algebra' to follow this argument.

and, therefore, we can only allow functions which make the differential $d\vec{a} \cdot \vec{P}(f^*, g|\vec{r})$ vanish on $\partial\Omega$. It must hold then for all $f \in \mathbb{H}$

$$f(\vec{r}) \;=\; 0 \quad \forall \vec{r},\, \vec{r} \in \partial\Omega \tag{3.29}$$

or

$$d\vec{a} \cdot \nabla f(\vec{r}) \;=\; 0 \quad \forall \vec{r},\, \vec{r} \in \partial\Omega \tag{3.30}$$

Note that these boundary conditions are linear in f, i.e., if f and g satisfy these conditions than also does any linear combination $\alpha f + \beta g$. Often the closed surface of a volume $\partial\Omega$ is the union of disconnected surfaces[3] $\partial\Omega_j$, i.e., $\partial\Omega = \partial\Omega_1 \cap \partial\Omega_2 \cap \partial\Omega_3 \cap \ldots$ In this case one can postulate both conditions (3.29, 3.30) each condition holding on an entire surface $\partial\Omega_j$. However, to avoid discontinuities in $\psi(\vec{r}, t)$ on a single connected surface $\partial\Omega_j$ only either one of the conditions (3.29, 3.30) can be postulated.
A most common boundary condition is encountered for the volume $\Omega = \vec{R}^3$ in which case one postulates

$$\lim_{|\vec{r}|\to\infty} f(\vec{r}) \;=\; 0 \quad \text{“natural boundary condition”} \; . \tag{3.31}$$

In fact, in this case also all derivatives of $f(\vec{r})$ vanish at infinity. The latter property stems from the fact that the boundary condition (3.31) usually arises when a particle existing in a bound state is described. In this case one can expect that the particle density is localized in the area where the energy of the particle exceeds the potential eenrgy, and that the density decays rapidly when one moves away from that area. Since the total probability of finding the particle anywhere in space is

$$\int d^3r |f(\vec{r})|^2 \;=\; 1 \tag{3.32}$$

the wave function must decay for $|\vec{r}| \to \infty$ rapidly enough to be *square integrable*, i.e., obey (3.32), e.g., like $\exp(-\kappa r)$, $\kappa > 0$ or like $r^{-\alpha}$, $\alpha > 2$. In either case does $f(\vec{r})$ and all of its derivatives vanish asymptotically.

3.3 Particle Flux and Schrödinger Equation

The solution of the Schrödinger equation is the wave function $\psi(\vec{r}, t)$ which describes the state of a particle moving in the potential $U(\vec{r}, t)$. The observable directly linked to the wave function is the probability to find the particle at position $\vec{r}$ at time t, namely, $|\psi(\vec{r}, t)|^2$. The probability to observe the particle anywhere in the subvolume $\omega \subset \Omega$ is

$$p(\omega, t) \;=\; \int_\omega d^3r \, |\psi(\vec{r}, t)|^2 \; . \tag{3.33}$$

The time derivative of $p(\omega, t)$ is

$$\partial_t p(\omega, t) \;=\; \int_\omega d^3r \; [\, \psi^*(\vec{r}, t)\partial_t\psi(\vec{r}, t) \;+\; \psi(\vec{r}, t)\partial_t\psi^*(\vec{r}, t) \,] \; . \tag{3.34}$$

[3] An example is a volume between two concentric spheres, in which case $\partial\Omega_1$ is the inner sphere and $\partial\Omega_2$ is the outer sphere.

Using (3.19) and its conjugate complex[4]

$$-i\hbar\frac{\partial}{\partial t}\,\psi^*(\vec{r},t) \;=\; \hat{H}\,\psi^*(\vec{r},t) \tag{3.35}$$

yields

$$i\hbar\partial_t p(\omega,t) \;=\; \int_\omega d^3r\,\left[\,\psi^*(\vec{r},t)\,\hat{H}\,\psi(\vec{r},t) \;-\; \psi(\vec{r},t)\,\hat{H}\,\psi^*(\vec{r},t)\,\right]\,. \tag{3.36}$$

According to (3.26, 3.27) this can be written

$$i\hbar\partial_t p(\omega,t) \;=\; \int_{\partial\omega} d\vec{a}\cdot\vec{P}(\psi^*(\vec{r},t),\psi(\vec{r},t)\vec{r},t)\,. \tag{3.37}$$

If one applies this identity to $\omega = \Omega$ one obtains according to (3.29, 3.30) $\partial_t p(\Omega,t) = 0$. Accordingly the probability to observe the particle anywhere in the total volume Ω is constant. A natural choice for this constant is 1. One can multiply the solution of (3.18) by any complex number and accordingly one can define $\psi(\vec{r},t)$ such that

$$\int_\Omega d^3r|\psi(\vec{r},t)|^2 \;=\; 1 \tag{3.38}$$

holds. One refers to such solution as *normalized*. We will assume in the remainder of this Section that the solutions discussed are normalized. Note that for a normalized wave function the quantity

$$\rho(\vec{r},t) \;=\; |\psi(\vec{r},t)|^2 \tag{3.39}$$

is a probability density with units 1/volume.
The surface integral (3.37) can be expressed through a volume integral according to

$$\int_{\partial\omega} d\vec{a}\cdot\vec{A}(\vec{r}) \;=\; \int_\omega d^3r\,\nabla\cdot\vec{A}(\vec{r}) \tag{3.40}$$

One can rewrite then (3.37)

$$\int_\omega d^3r\,\Big(\,\partial_t\rho(\vec{r},t) \;+\; \nabla\cdot\vec{j}(\vec{r},t)\Big) \;=\; 0 \tag{3.41}$$

where

$$\vec{j}(\vec{r},t) \;=\; \vec{P}(\psi^*,\psi|\vec{r},t)\,. \tag{3.42}$$

Using (3.27) one can express this

$$\vec{j}(\vec{r},t) \;=\; \frac{\hbar}{2mi}\,[\,\psi(\vec{r},t)\,\nabla\psi^*(\vec{r},t) \;-\; \psi^*(\vec{r},t)\,\nabla\psi(\vec{r},t)\,]\,. \tag{3.43}$$

Since (3.41) holds for any volume $\omega\subset\Omega$ one can conclude

$$\partial_t\rho(\vec{r},t) \;+\; \nabla\cdot\vec{j}(\vec{r},t) \;=\; 0\,. \tag{3.44}$$

[4]Note that the Hamiltonian $\hat{H}$ involves only real terms.

The interpretation of $\vec{j}(\vec{r},t)$ is that of density flux. This follows directly from an inspection of Eq. (3.41) written in the form

$$\partial_t \int_\omega d^3r\, \rho(\vec{r},t) \;=\; -\int_{\partial\omega} d\vec{a}\cdot\vec{j}(\vec{r},t)\;. \tag{3.45}$$

Obviously, $\vec{j}(\vec{r},t)$ gives rise to a descrease of the total probability in volume ω due to the disappearence of probability density at the surface $\partial\omega$. Note that $\vec{j}(\vec{r},t)$ points in the direction *to the outside* of volume ω.
It is of interest to note from (3.43) that any real wave function $\psi(\vec{r},t)$ has vanishing flux anywhere. One often encounters wave functions of the type

$$\phi(\vec{r}) \;=\; f(\vec{r})\, e^{i\,\vec{k}\cdot\vec{r}} \quad , \qquad \text{for } f(\vec{r}) \in \mathbb{R}. \tag{3.46}$$

The corresponding flux is

$$\vec{j}(\vec{r}) \;=\; \frac{\hbar\vec{k}}{m}\, f^2(\vec{r})\;, \tag{3.47}$$

i.e., arises solely from the complex factor $\exp(i\,\vec{k}\cdot\vec{r})$. Such case arose in Sect. 2 for a free particle [c.f. (2.48, 2.71)], and for particles moving in a linear [c.f. (2.105, 2.125)] and in a quadratic [c.f. (2.148, 2.167)] potential. In Sect. 2 we had demonstrated that a factor $\exp(ip_ox_o/\hbar)$ induces a motion of 1-dimensional wave packets such that p_o/m corresponds to the initial velocity. This finding is consistent with the present evaluation of the particle flux: a factor $\exp(ip_ox_o/\hbar)$ gives rise to a flux p_o/m, i.e., equal to the velocity of the particle. The generalization to three dimenisons implies then that the factor $\exp(i\,\vec{k}\cdot\vec{r})$ corresponds to an intial velocit $\hbar\vec{k}/m$ and a flux of the same magnitude.

3.4 Solution of the Free Particle Schrödinger Equation

We want to consider now solutions of the Schrödinger equation (3.18) in $\Omega_\infty = \mathbb{R}^{\not\vdash}$ in the case $U(\vec{r},t) = 0$

$$i\hbar\frac{\partial}{\partial t}\,\psi(\vec{r},t) \;=\; -\frac{\hbar^2}{2m}\,\nabla^2\,\psi(\vec{r},t) \tag{3.48}$$

which describes the motion of free particles. One can readily show by insertion into (3.48) that the general solution is of the form

$$\psi(\vec{r},t) \;=\; [2\pi]^{-\frac{3}{2}} \int_{\Omega_\infty} d^3k\;\tilde{\phi}(\vec{k})\exp\Big(i(\vec{k}\cdot\vec{r}\;-\;\omega t)\Big) \tag{3.49}$$

where the dispersion relationship holds

$$\omega \;=\; \frac{\hbar k^2}{2m}\;. \tag{3.50}$$

Obviously, the initial condition at $\psi(\vec{r},t_0)$ determines $\tilde{\phi}(\vec{k})$. Equation (3.49) reads at $t = t_0$

$$\psi(\vec{r},t_0) \;=\; [2\pi]^{-\frac{3}{2}} \int_{\Omega_\infty} d^3k\;\tilde{\phi}(\vec{k})\exp\Big(i(\vec{k}\cdot\vec{r}-\omega t_0)\Big)\;. \tag{3.51}$$

The inverse Fourier transform yields

$$\tilde{\phi}(\vec{k}) \;=\; [2\pi]^{-\frac{3}{2}} \int_{\Omega_\infty} d^3 r_0 \exp(-i\vec{k}\cdot\vec{r}_0)\,\psi(\vec{r}_0,t_0)\;. \tag{3.52}$$

We have not specified the spatial boundary condition in case of (3.49). The solution as stated is defined in the infinte space $\Omega_\infty = \mathbb{R}^{\not\#}$. If one chooses the initial state $f(\vec{r})$ defined in (3.51) to be square integrable it follows according to the properties of the Fourier–transform that $\psi(\vec{r},t)$ as given by (3.49) is square integrable at all subsequent times and, hence, that the "natural boundary condition" (3.31) applies. The ensuing solutions are called wave packets.

Comparision with Path Integral Formulation

One can write solution (3.49, 3.51, 3.52) above

$$\psi(\vec{r},t) \;=\; \int_{\Omega_\infty} d^3 r_0\,\phi(\vec{r},t|\vec{r}_0,t_0)\,\psi(\vec{r}_0,t_0) \tag{3.53}$$

where

$$\phi(\vec{r},t|\vec{r}_0,t_0) \;=\; \left[\frac{1}{2\pi}\right]^3 \int_{\Omega_\infty} d^3k \exp\left(i\vec{k}\cdot(\vec{r}\,-\,\vec{r}_0) \;-\frac{i}{\hbar}\,\frac{\hbar^2k^2}{2m}\,(t\,-\,t_0)\right)\;. \tag{3.54}$$

This expression obviously has the same form as postulated in the path integral formulation of Quantum Mechanics introduced above, i.e., in (2.5). We have identified then with (3.54) an alternative representation of the propagator (2.47). In fact, evaluating the integral in (3.54) yields (2.47). To show this one needs to note

$$\begin{aligned}&\frac{1}{2\pi}\int_{-\infty}^{+\infty} dk_1 \exp\left(i\vec{k}_1(x\,-\,x_0) \;-\frac{i}{\hbar}\,\frac{\hbar^2k_1^2}{2m}\,(t\,-\,t_0)\right)\\ &=\;\left[\frac{m}{2\pi i\hbar(t-t_0)}\right]^{\frac{1}{2}} \exp\left[\frac{im}{2\hbar}\frac{(x-x_0)^2}{t-t_0}\right]\;.\end{aligned} \tag{3.55}$$

This follows from completion of the square

$$\begin{aligned}&i\vec{k}_1(x\,-\,x_0) \;-\frac{i}{\hbar}\,\frac{\hbar^2k_1^2}{2m}\,(t\,-\,t_0)\\ &\quad=\;-i\,\frac{\hbar(t-t_o)}{2m}\;\left[k_1\,-\,\frac{m}{\hbar}\,\frac{x-x_o}{t-t_o}\right]^2 \;+\;\frac{i}{\hbar}\,\frac{m}{2}\,\frac{(x-x_o)^2}{t-t_o}\end{aligned} \tag{3.56}$$

and using (2.247).
Below we will generalize the propagator (3.54) to the case of non-vanishing potentials $U(\vec{r})$, i.e., derive an expression similar to (3.54) valid for this case. The general form for this propagator involves an expansion in terms of a complete set of eigenfunctions as in (3.114) and (4.70) derived below for a particle in a box and the harmonic oscillator, respectively. In case of the harmonoc oscillator the expansion can be stated in a closed form given in (4.81)

Free Particle at Rest

We want to apply solution (3.49, 3.52) to the case that the initial state of a 1-dimensional free particle $\psi(x_0, t)$ is given by (**??**). The 1-dimensional version of (3.53, 3.54) is

$$\psi(x,t) \;=\; \int_{-\infty}^{+\infty} dx_0\, \phi(x,t|x_0,t_0)\, \psi(x_0,t) \tag{3.57}$$

where

$$\phi(x,t|x_0,t_0) \;=\; \frac{1}{2\pi} \int_{-\infty}^{+\infty} dk\, \exp\left(ik\,(x \,-\, x_0) \;-\, \frac{i}{\hbar}\,\frac{\hbar^2 k^2}{2m}\,(t \,-\, t_0) \right) . \tag{3.58}$$

Integration over x_0 leads to the integral

$$\int_{-\infty}^{+\infty} dx_0 \exp\left(-ikx_0 \;+\; \frac{i}{\hbar}\,p_o x_0 \;-\; \frac{x_0^2}{2\delta^2} \right) \;=\; \sqrt{2\pi\delta^2}\, \exp\left(-\frac{(k - \frac{p_o}{\hbar})^2 \delta^2}{2} \right) \tag{3.59}$$

which is solved through completion of the square in the exponent [c.f. (3.55, (3.56)]. The remaining integration over k leads to the integral

$$\begin{array}{c} \int_{-\infty}^{+\infty} dk\, \exp\left[ikx \;-\; \frac{1}{2}\,(k - \frac{p_o}{\hbar})^2\, \delta^2 \;-\; i\,\frac{\hbar k^2}{m}(t - t_0) \right] \\ =???? \end{array} \tag{3.60}$$

Combining (3.57–3.60) yields with $t_0 \,=\, 0$

$$\begin{aligned} \psi(x,t) \quad = \quad & \left[\frac{1 \,-\, i\frac{\hbar t}{m\delta^2}}{1 \,+\, i\frac{\hbar t}{m\delta^2}} \right]^{\frac{1}{4}} \left[\frac{1}{\pi\delta^2\,(1 \,+\, \frac{\hbar^2 t^2}{m^2\delta^4})} \right]^{\frac{1}{4}} \;\times \\ & \times \exp\left[-\frac{(x \,-\, \frac{p_o}{m}\,t)^2}{2\delta^2(1 \,+\, \frac{\hbar^2 t^2}{m^2\delta^4})}(1 \,-\, i\frac{\hbar t}{m\delta^2}) \;+\; i\frac{p_o}{\hbar}x \;-\; \frac{i}{\hbar}\frac{p_o^2}{2m}t \right] . \end{aligned} \tag{3.61}$$

a result which is identical to the expression (**??**) obtained by means of the path integral propagator (2.46). We have demonstrated then in this case that the Schrödinger formulation of Quantum Mechanics is equivalent to the Feynman path integral formulation.

Stationary States

We consider now solutions of the time-dependent Schrödinger equation (3.19, 3.20) which are of the form

$$\psi(\vec{r},t) \;=\; f(t)\,\phi(\vec{r})\,. \tag{3.62}$$

We will restrict the space of allowed solutions to a volume Ω such that the functions also make the concomitant (3.27) vanish on the surface $\partial\Omega$ of Ω, i.e., the functions obey boundary conditions of the type (3.29, 3.30). Accordingly, the boundary conditions are

$$\phi(\vec{r}) \;=\; 0 \quad \forall \vec{r},\, \vec{r} \,\in\, \partial\Omega \tag{3.63}$$

or

$$d\vec{a} \cdot \nabla\phi(\vec{r}) \;=\; 0 \quad \forall \vec{r},\, \vec{r} \,\in\, \partial\Omega \tag{3.64}$$

and affect only the spatial wave function $\phi(\vec{r})$. As pointed out above, a common case is $\Omega = \Omega_\infty$ and the 'natural boundary condition' (3.31). We will demonstrate that solutions of the type (3.62) do exist and we will characterize the two factors of the solution $f(t)$ and $\phi(\vec{r})$. We may note in passing that solutions of the type (3.62) which consist of two factors, one factor depending only on the time variable and the other only on the space variables are called *separable in space and time*. It is important to realize that the separable solutions (3.62) are special solutions of the time-dependent Schrödinger equation, by no means all solutions are of this type. In fact, the solutions (3.62) have the particular property that the associated probability distributions are independent of time. We want to demonstrate this property now. It follows from the observation that for the solution space considered (3.27) holds and, hence, according to (3.42) the flux $\vec{j}(\vec{r},t)$ vanishes on the surface of $\partial\Omega$. It follows then from (3.45) that the total probability

$$\int_\Omega d^3r\, \rho(\vec{r},t) \;=\; \int_\Omega d^3r\, |\psi(\vec{r},t)|^2 \;=\; |f(t)|^2 \int_\Omega d^3r\, |\phi(\vec{r})|^2 \tag{3.65}$$

is constant. This can hold only if $|f(t)|$ is time-independent, i.e., if

$$f(t) \;=\; e^{i\alpha},\; \alpha \in \mathbb{R}\,. \tag{3.66}$$

One can conclude that the probability density for the state (3.62) is

$$|\psi(\vec{r},t)|^2 \;=\; |\phi(\vec{r})|^2\,, \tag{3.67}$$

i.e., is time-independent. One calls such states *stationary states*.
In order to further characterize the solution (3.62) we insert it into (3.19). This yields an expression

$$g_1(t)\, h_1(\vec{r}) \;=\; g_2(t)\, h_2(\vec{r}) \tag{3.68}$$

where $g_1(t) = i\hbar\partial_t f(t)$, $g_2(t) = f(t)$, $h_1(\vec{r}) = \phi(\vec{r})$, and $h_2(\vec{r}) = \hat{H}\,\phi(\vec{r})$. The identity (3.68) can hold only for all t and all $\vec{r}$ if $g_1(t) = E\, g_2(t)$ and $E\, h_1(\vec{r}) = h_2(\vec{r})$ for some $E \in \mathbb{C}$. We must postulate therefore

$$\begin{aligned} \partial_t\, f(t) \;&=\; E\, f(t) \\ \hat{H}\,\phi(\vec{r}) \;&=\; E\,\phi(\vec{r})\,. \end{aligned} \tag{3.69}$$

If these two equations can be solved simultaneously a solution of the type (3.62) exists.
It turns out that a solution for $f(t)$ can be found for any E, namely

$$f(t) \;=\; f(0) \exp\left(-\frac{i}{\hbar}\, E\, t\right)\,. \tag{3.70}$$

The task of finding solutions $\phi(\vec{r})$ which solve (3.69) is called an eigenvalue problem. We will encounter many such problems in the subsequent Sections. At this point we state without proof that, in general, for the eigenvalue problems in the confined function space, i.e., for functions required to obey boundary conditions (3.63, 3.64), solutions exist only for a set of discrete E values, the eigenvalues of the operator $\hat{H}$. At this point we will accept that solutions $\phi(\vec{r})$ of the type (3.69) exist, however, often only for a discrete set of values E_n, $n = 1, 2, \ldots$ We denote the corresponding solution by $\phi_E(\vec{r})$. We have then shown that

$$\psi(\vec{r},t) \;=\; f(0) \exp\left(-\frac{i}{\hbar}\, E\right) \phi_E(\vec{r}) \quad \text{where} \quad \hat{H}\,\phi_E(\vec{r}) \;=\; E\,\phi_E(\vec{r})\,. \tag{3.71}$$

is a solution of the time-dependent Schrödinger equation (3.19, 3.20).
According to (3.66) E must be real. We want to prove now that the eigenvalues E which arise in the eigenvalue problem (3.71) are, in fact, real. We start our proof using the property (3.28) for the special case that f and g in (3.28) both represent the state $\phi_E(\vec{r})$, i.e.,

$$\int_\Omega d^3r\, \phi_E^*(\vec{r})\, \hat{H}\, \phi_E(\vec{r}) \;=\; \overline{\int_\Omega d^3r\, \phi_E(\vec{r})\, \hat{H}\, \phi_E^*(\vec{r})}\,. \tag{3.72}$$

According to (3.71) this yields

$$E \int_\Omega d^3r\, \phi_E^*(\vec{r})\, \phi_E(\vec{r}) \;=\; E^* \overline{\int_\Omega d^3r\, \phi_E(\vec{r})\, \phi_E^*(\vec{r})}\,. \tag{3.73}$$

from which follows $E = E^*$ and, hence, $E \in \mathbb{R}$. We will show in Section 5 that E can be interprerted as the total energy of a stationary state.

Stationary State of a Free Particle

We consider now the stationary state of a free particle described by

$$\psi(\vec{r},t) \;=\; \exp\left(-\frac{i}{\hbar}\,E\,t\right)\,\phi_E(\vec{r}) \quad , \quad -\frac{\hbar^2}{2m}\,\nabla^2\,\phi(\vec{r}) \;=\; E\,\phi(\vec{r})\,. \tag{3.74}$$

The classical free particle with constant energy $E > 0$ moves without bounds in the space Ω_∞. As a result we cannot postulate in the present case that wave functions are localized and normalizable. We will wave this assumption as we always need to do later whenever we deal with unbound particles, e.g. particles scattered of a potential.
The solution $\phi_E(\vec{r})$ corresponding to the eigenvalue problem posed by (3.74) is actually best labelled by an index $\vec{k}$, $\vec{k} \in \mathbb{R}^{\not\vDash}$

$$\phi_{\vec{k}}(\vec{r}) \;=\; N \exp\Big(\,i\,\vec{k}\cdot\vec{r}\,\Big) \;, \quad \frac{\hbar^2k^2}{2m} \;=\; E\,. \tag{3.75}$$

One can ascertain this statement by inserting the expression for $\phi_{\vec{k}}(\vec{r})$ into the eigenvalue problem posed in (3.74) using $\nabla \exp\Big(\,i\,\vec{k}\cdot\vec{r}\,\Big) \;=\; i\vec{k}\exp\Big(\,i\,\vec{k}\cdot\vec{r}\,\Big)$. The resulting total energy values E are positive, a property which is to be expected since the energy is purely kinetic energy which, of course, should be positive.
The corresponding stationary solution

$$\psi(\vec{r},t) \;=\; \exp\left(-\frac{i}{\hbar}\,\frac{\hbar^2k^2}{2m}\,t\right)\,\exp\Big(\,i\,\vec{k}\cdot\vec{r}\,\Big) \tag{3.76}$$

has kinetic energy $\hbar^2k^2/2m$. Obviously, one can interpret then $\hbar k$ as the magnitude of the momentum of the particle. The flux corresponding to (3.76) according to (3.43) is

$$\vec{j}(\vec{r},t) \;=\; |N|^2\,\frac{\hbar\vec{k}}{m}\,. \tag{3.77}$$

Noting that $\hbar k$ can be interpreted as the magnitude of the momentum of the particle the flux is equal to the velocity of the particle $\vec{v} = \hbar\vec{k}/m$ multiplied by $|N|^2$.

3.5 Particle in One-Dimensional Box

As an example of a situation in which only bound states exist in a quantum system we consider the stationary states of a particle confined to a one-dimensional interval $[-a, a] \subset \mathbb{R}$ assuming that the potential outside of this interval is infinite. We will refer to this as a particle in a one-dimensional 'box'.

Setting up the Space $\mathcal{F}_1$ of Proper Spatial Functions

The presence of the infinite energy wall is accounted for by restricting the spatial dependence of the solutions to functions $f(x)$ defined in the domain $\Omega_1 = [-a, a] \subset \mathbb{R}$ which vanish on the surface $\partial\Omega_1 = \{-a, a\}$, i.e.,

$$f \in \mathcal{F}_1 = \{f : [-a, a] \subset \mathbb{R} \to \mathbb{R},\ f \text{ continuous},\ \mho(\pm\supset) = \nvDash\} \tag{3.78}$$

Solutions of the Schrödinger Equation in $\mathcal{F}_1$

The time–dependent solutions satisfy

$$i\hbar\partial_t\psi(x,t) = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi(x,t)\ . \tag{3.79}$$

The stationary solutions have the form $\psi(x,t) = \exp(-iEt/\hbar)\phi_E(x)$ where $\phi_E(x)$ is determined by

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\phi_E(x) = E\,\phi_E(x)\ , \quad \phi(\pm a) = 0\ . \tag{3.80}$$

We note that the box is symmetric with respect to the origin. We can expect, hence, that the solutions obey this symmetry as well. We assume, therefore, two types of solutions, so-called *even* solutions obeying $\phi(x) = \phi(-x)$

$$\phi_E^{(e)}(x) = A\cos kx\ , \quad \frac{\hbar^2k^2}{2m} = E \tag{3.81}$$

and so-called *odd* solutions obeying $\phi(x) = -\phi(-x)$

$$\phi_E^{(o)}(x) = A\sin kx;\ , \quad \frac{\hbar^2k^2}{2m} = E\ . \tag{3.82}$$

One can readily verify that (3.81, 3.82) satisfy the differential equation in (3.80).
The boundary conditions which according to (3.80) need to be satisfied are

$$\phi_E^{(e,o)}(a) = 0 \text{ and } \phi_E^{(e,o)}(-a) = 0 \tag{3.83}$$

The solutions (3.81, 3.82) have the property that either both boundary conditions are satisfied or none. Hence, we have to consider only one boundary condition, let say the one at $x = a$. It turns out that this boundary condition can only be satisfied for a discrete set of k–values $k_n, n \in \mathbb{N}$. In case of the even solutions (3.81) they are

$$k_n = \frac{n\pi}{2a}\ , \quad n = 1, 3, 5\ldots \tag{3.84}$$

since for such k_n

$$\cos(k_n a) \;=\; \cos\left(\frac{n\pi a}{2a}\right) \;=\; \cos\left(\frac{n\pi}{2}\right) \;=\; 0\,. \tag{3.85}$$

In case of the odd solutions (3.82) only the k_n–values

$$k_n \;=\; \frac{n\pi}{2a}\,, \quad n \;=\; 2,4,6\ldots \tag{3.86}$$

satisfy the boundary condition since for such k_n

$$\sin(k_n a) \;=\; \sin\left(\frac{n\pi a}{2a}\right) \;=\; \sin\left(\frac{n\pi}{2}\right) \;=\; 0\,. \tag{3.87}$$

(Note that, according to (3.86), n is assumed to be even.)

The Energy Spectrum and Stationary State Wave Functions

The energy values corresponding to the k_n–values in (3.84, 3.86), according to the dispersion relationships given in (3.81, 3.82), are

$$E_n \;=\; \frac{\hbar^2\pi^2}{8ma^2}\,n^2\,, \quad n \;=\; 1,2,3\ldots \tag{3.88}$$

where the energies for odd (even) n–values correspond to the even (odd) solutions given in (3.81) and (3.82), respectively, i.e.,

$$\phi_n^{(e)}(a;x) \;=\; A_n \cos\frac{n\pi x}{2a}\,, \quad n \;=\; 1,3,5\ldots \tag{3.89}$$

and

$$\phi_n^{(o)}(a;x) \;=\; A_n \sin\frac{n\pi x}{2a}\,, \quad n \;=\; 2,4,6\ldots\,. \tag{3.90}$$

The wave functions represent stationary states of the particle in a one-dimensional box. The wave functions for the five lowest energies E_n are presented in Fig. (3.1). Notice that the number of nodes of the wave functions increase by one in going from one state to the state with the next higher energy E_n. By counting the number of their nodes one can determine the energy ordering of the wave functions.
It is desirable to normalize the wave functions such that

$$\int_{-a}^{+a} dx\,|\phi_n^{(e,o)}(a;x)|^2 \;=\; 1 \tag{3.91}$$

holds. This condition implies for the even states

$$|A_n|^2 \int_{-a}^{+a} dx\cos^2\frac{n\pi x}{2a} \;=\; |A_n|^2\,a \;=\; 1\,, \quad n \;=\; 1,3,5\ldots \tag{3.92}$$

and for the odd states

$$|A_n|^2 \int_{-a}^{+a} dx\sin^2\frac{n\pi x}{2a} \;=\; |A_n|^2\,a \;=\; 1\,, \quad n \;=\; 2,4,6\ldots \tag{3.93}$$

The normalization constants are then

$$A_n \;=\; \sqrt{\frac{1}{a}}\,. \tag{3.94}$$

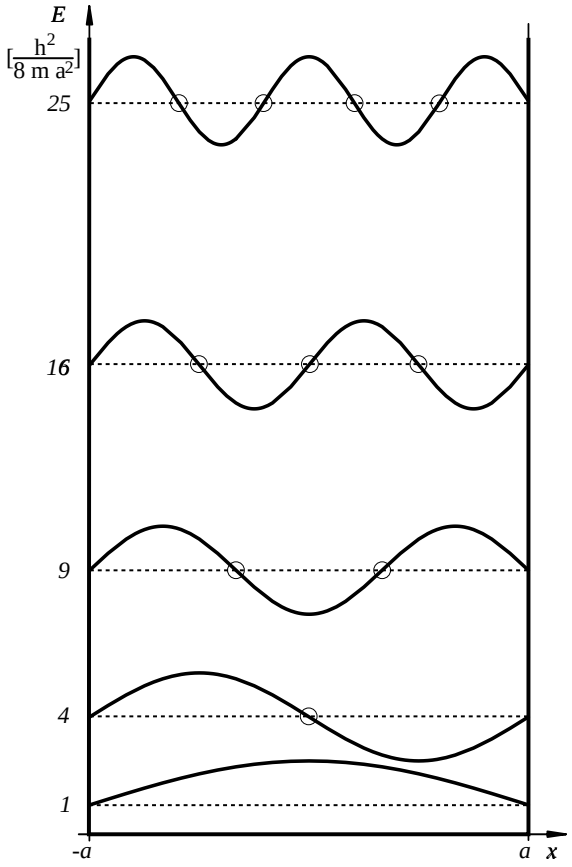


Figure 3.1: Eigenvalues E_n and eigenfunctions $\phi_n^{(e,o)}(a;x)$ for $n = 1,2,3,4,5$ of particle in a box.

The Stationary States form a Complete Orthonormal Basis of $\mathcal{F}_1$

We want to demonstrate now that the set of solutions (3.89, 3.90, 3.94)

$$\mathcal{B}_1 \;=\; \{\phi_n(a;x)\,,\; n \,=\, 1,2,3,\ldots\} \tag{3.95}$$

where

$$\phi_n(a;x) \;=\; \sqrt{\frac{1}{a}} \left\{ \begin{array}{ll} \cos\frac{n\pi x}{2a} & \text{for } n \,=\, 1,3,5\ldots \\ \sin\frac{n\pi x}{2a} & \text{for } n \,=\, 2,4,6\ldots \end{array} \right. \tag{3.96}$$

together with the scalar product[5]

$$\langle f|g\rangle_{\Omega_1} \;=\; \int_{-a}^{+a} dx f(x)\, g(x)\,, \quad f,\, g \,\in\, \mathcal{F}_1 \tag{3.97}$$

form an *orthonormal basis* set, i.e., it holds

$$\langle \phi_n|\phi_m\rangle_{\Omega_1} \;=\; \delta_{n\,m}\,. \tag{3.98}$$

The latter property is obviously true for $n \,=\, m$. In case of $n \,\neq\, m$ we have to consider three cases, (i) n, m both odd, (ii) n, m both even, and (iii) the mixed case. The latter case leads to integrals

$$\langle \phi_n|\phi_m\rangle_{\Omega_1} \;=\; \frac{1}{a} \int_{-a}^{+a} dx \cos\frac{n\pi x}{2a} \sin\frac{m\pi x}{2a}\,. \tag{3.99}$$

Since in this case the integrand is a product of an even and of an odd function, i.e., the integrand is odd, the integral vanishes. Hence we need to consider only the first two cases. In case of n, m odd, $n \,\neq\, m$, the integral arises

$$\begin{array}{l} \langle \phi_n|\phi_m\rangle_{\Omega_1} \;=\; \frac{1}{a} \int_{-a}^{+a} dx \cos\frac{n\pi x}{2a} \cos\frac{m\pi x}{2a} \;= \\ \frac{1}{a} \int_{-a}^{+a} dx \left[\cos\frac{(n-m)\pi x}{2a} \,+\, \cos\frac{(n+m)\pi x}{2a} \right] \end{array} \tag{3.100}$$

The periods of the two cos-functions in the interval $[-a, a]$ are N, $N \,\geq\, 1$. Obviously, the integrals vanish. Similarly, one obtains for n, m even

$$\begin{array}{l} \langle \phi_n|\phi_m\rangle_{\Omega_1} \;=\; \frac{1}{a} \int_{-a}^{+a} dx \sin\frac{n\pi x}{2a} \sin\frac{m\pi x}{2a} \;= \\ \frac{1}{a} \int_{-a}^{+a} dx \left[\cos\frac{(n-m)\pi x}{2a} \,-\, \cos\frac{(n+m)\pi x}{2a} \right] \end{array} \tag{3.101}$$

and, hence, this integral vanishes, too.
Because of the property (3.98) the elements of $\mathcal{B}_1$ must be linearly independent. In fact, for

$$f(x) \;=\; \sum_{n=1}^{\infty} d_n\, \phi_n(a;x) \tag{3.102}$$

holds according to (3.98)

$$\langle f|f\rangle_{\Omega_1} \;=\; \sum_{n=1}^{\infty} d_n^2\,. \tag{3.103}$$

[5]We will show in Section 5 that the property of a scalar product do indeed apply. In particular, it holds: $\langle f|f\rangle_{\Omega_1} \,=\, 0 \,\rightarrow\, f(x) \,\equiv\, 0$.

$f(x) \equiv 0$ implies $\langle f|f\rangle_{\Omega_1} = 0$ which in turn implies $d_n = 0$ since $d_n^2 \geq 0$. It follows that $\mathcal{B}_1$ defined in (3.95) is an orthonormal basis.
We like to show finally that the basis (3.95) is also *complete*, i.e., any element of the function space $\mathcal{F}_1$ defined in (3.78) can be expressed as a linear combination of the elements of $\mathcal{B}_1$ defined in (3.95, 3.96). Demonstration of completeness is a formidable task. In the present case, however, such demonstration can be based on the theory of Fourier series. For this purpose we extend the definition of the elements of $\mathcal{F}_1$ to the whole real axis through

$$\tilde{f} : \mathbb{R} \to \mathbb{R}; \quad \tilde{f}(x) = f([x + a]_{2a} - a) \tag{3.104}$$

where $[y]_a = y \bmod 2a$. The functions $\tilde{f}$ are periodic with period $2a$. Hence, they can be expanded in terms of a Fourier series, i.e., there exist real constants $\{a_n,\, n = 0, 1, 2, \ldots\}$ and $\{b_n,\, n = 1, 2, \ldots\}$ such that

$$\tilde{f} = \sum_{n=0}^{\infty} a_n \cos\frac{n\pi x}{2a} + \sum_{n=1}^{\infty} a_n \sin\frac{n\pi x}{2a} \tag{3.105}$$

The functions $\tilde{f}$ corresponding to the functions in the space $\mathcal{F}_1$ have zeros at $x = \pm m\, a$, $m = 1, 3, 5 \ldots$ Accordingly, the coefficients a_n, $n = 2, 4, \ldots$ and b_n, $n = 1, 3, \ldots$ in (3.105) must vanish. This implies that only the trigonometric functions which are elements of $\mathcal{B}_1$ enter into the Fourier series. We have then shown that any $\tilde{f}$ corresponding to elements of $\mathcal{F}_1$ can be expanded in terms of elements in $\mathcal{B}_1$. Restricting the expansion (3.105) to the interval $[-a, a]$ yields then also an expansion for any element in $\mathcal{F}_1$ and $\mathcal{B}_1$ is a complete basis for $\mathcal{F}_1$.

Evaluating the Propagator

We can now use the expansion of any initial wave function $\psi(x, t_0)$ in terms of eigenfunctions $\phi_n(a; x)$ to obtain an expression for $\psi(x, t)$ at times $t > t_0$. For this purpose we expand

$$\psi(x, t_0) = \sum_{n=1}^{\infty} d_n\, \phi_n(a; x)\ . \tag{3.106}$$

Using the orthonormality property (3.98) one obtains

$$\int_{-a}^{+a} dx_0 \phi_m(a; x_0)\, \psi(x_0, t_0) = d_m\ . \tag{3.107}$$

Inserting this into (3.106) and generalizing to $t \geq t_0$ one can write

$$\psi(x, t) = \sum_{n=1}^{\infty} \phi_n(a; x)\, c_n(t) \int_{-a}^{+a} dx_0 \phi_n(a; x_0)\, \psi(x_0, t_0) \tag{3.108}$$

where the functions $c_n(t)$ are to be determined from the Schrödinger equation (3.79) requiring the initial condition

$$c_n(t_0) = 1\ . \tag{3.109}$$

Insertion of (3.108) into the Schrödinger equation yields

$$\begin{aligned} &\textstyle\sum_{n=1}^{\infty} \phi_n(a; x)\, \partial_t c_n(t) \int_{-a}^{+a} dx_0 \phi_n(a; x_0)\, \psi(x_0, t_0) = \\ &\textstyle\sum_{n=1}^{\infty} \left(-\frac{i}{\hbar} E_n\right) \phi_n(a; x)\, c_n(t) \int_{-a}^{+a} dx_0 \phi_n(a; x_0)\, \psi(x_0, t_0) \quad . \end{aligned} \tag{3.110}$$

Multiplying both sides by $\phi_m(a;x)$ and integrating over $[-a,a]$ yields, according to (3.98),

$$\partial_t\, c_m(t) \;=\; -\frac{i}{\hbar}\, E_m\, c_m(t)\;, \quad c_m(t_0) \;=\; 1\;. \tag{3.111}$$

The solutions of these equations which satisfy (3.109) are

$$c_m(t) \;=\; \exp\left(\, -\frac{i}{\hbar}\, E_m\,(t \,-\, t_0)\,\right)\;. \tag{3.112}$$

Equations (3.108, 3.112) determine now $\psi(x,t)$ for any initial condition $\psi(x,t_0)$. This solution can be written

$$\psi(x,t) \;=\; \int_{-a}^{+a} dx_0\, \phi(x,t|x_0,t_0)\, \psi(x_0,t_0) \tag{3.113}$$

where

$$\phi(x,t|x_0,t_0) \;=\; \sum_{n=1}^{\infty} \phi_n(a;x) \exp\left(\, -\frac{i}{\hbar}\, E_n\,(t \,-\, t_0)\,\right)\, \phi_n(a;x_0)\;. \tag{3.114}$$

This expression has the same form as postulated in the path integral formulation of Quantum Mechanics introduced above, i.e., in (2.5). We have identified then with (3.114) the representation of the propagator for a particle in a box with infinite walls.
It is of interest to note that $\phi(x,t|x_0,t_0)$ itself is a solution of the time-dependent Schrödinger equation (3.79) which lies in the proper function space (3.78). The respective initial condition is $\phi(x,t_0|x_0,t_0) \;=\; \delta(x \,-\, x_0)$ as can be readily verified using (3.113). Often the propagator $\phi(x,t|x_0,t_0)$ is also referred to as a *Greens function*. In the present system which is composed solely of bound states the propagator is given by a sum, rather than by an integral (3.54) as in the case of the free particle system which does not exhibit any bound states.
Note that the propagator has been evaluated in terms of elements of a particular function space $\mathcal{F}_1$, the elements of which satisfy the appropriate boundary conditions. In case that different boundary conditions hold the propagator will be different as well.

Example of a Non-Stationary State

As an illustration of a non-stationary state we consider a particle in an initial state

$$\psi(x_0,t_0) \;=\; \left[\frac{1}{2\pi\sigma^2}\right]^{\frac{1}{4}} \exp\left(\, -\frac{x_0^2}{2\sigma^2} \;+\; i\,k_0\,x_0\,\right)\;, \quad \sigma \;=\; \frac{a}{4}\,,\; k_0 \;=\; \frac{15}{a}\;. \tag{3.115}$$

This initial state corresponds to the particle being localized initially near $x_0 \;=\; 0$ with a velocity $v_0 \;=\; 15\hbar/m\,a$ in the direction of the positive x-axis. Figure 3.2 presents the probability distribution of the particle at subsequent times. One can recognize that the particle moves first to the left and that near the right wall of the box interference effects develop. The particle moves then to the left, being reflected at the right wall. The interference pattern begins to 'smear out' first, but the collision with the left wall leads to new interference effects. The last frame shows the wave front reaching again the right wall and the onset of new interference.

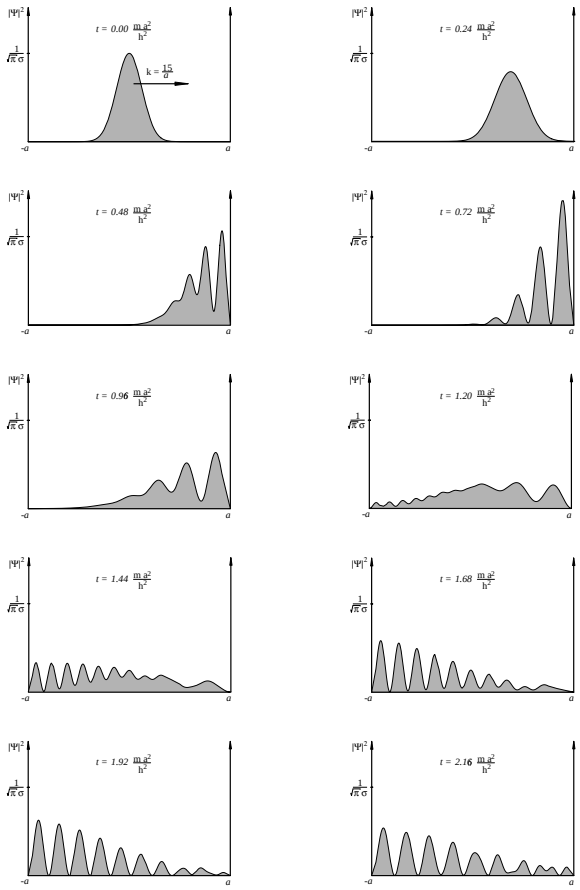


Figure 3.2: Stroboscopic views of the probability distribution $|\psi(x,t)|^2$ for a particle in a box starting in a Gaussian distribution with momentum $15\hbar/a$.

Summary: Particle in One-Dimensional Box

We like to summarize our description of the particle in the one-dimensional box from a point of view which will be elaborated further in Section 5. The description employed a space of functions $\mathcal{F}_1$ defined in (3.78). A complete basis of $\mathcal{F}_1$ is given by the infinite set $\mathcal{B}_1$ (3.95). An important property of this basis and, hence, of the space $\mathcal{F}_1$ is that the elements of the basis set can be enumerated by integer numbers, i.e., can be counted. We have defined in the space $\mathcal{F}_1$ a scalar product (3.97) with respect to which the eigenfunctions are orthonormal. This property allowed us to evaluate the propagator (3.114) in terms of which the solutions for all initial conditions can be expressed.

3.6 Particle in Three-Dimensional Box

We consider now a particle moving in a three-dimensional rectangular box with side lengths $2a_1$, $2a_2$, $2a_3$. Placing the origin at the center and aligning the x_1, x_2, x_3–axes with the edges of the box yields spatial boundary conditions which are obeyed by the elements of the function space

$$\mathcal{F}_3 \,=\, \{f\,:\Omega\,\rightarrow\,\mathbb{R},\, f \text{ continuous},\, \mho(\curvearrowright\!\!\!\!\!\!\not\,,\curvearrowright\!\!\!\!\not\,,\curvearrowright\!\!\!\!\!\!\not\,) \,=\!\not\vdash \forall\,(\curvearrowright\!\!\!\!\!\!\not\,,\curvearrowright\!\!\!\!\not\,,\curvearrowright\!\!\!\!\!\!\not\,)^{\mathbb{T}} \in \partial\not\leqslant\}\,. \tag{3.116}$$

where Ω is the interior of the box and $\partial\Omega$ its surface

$$\begin{aligned}\Omega \quad &= [-a_1,a_1]\otimes[-a_2,a_2]\otimes[-a_3,a_3] \,\subset \mathbb{R}^{\not\vdash}\\ \partial\Omega \quad &= \{(x_1,x_2,x_3)^T\in\Omega,\, x_1=\pm a_1\}\,\cup\,\{(x_1,x_2,x_3)^T\in\Omega,\, x_2=\pm a_2\}\\ &\quad \cup\{(x_1,x_2,x_3)^T\in\Omega,\, x_3=\pm a_3\}\,.\end{aligned} \tag{3.117}$$

We seek then solutions of the time-dependent Schrödinger equation

$$i\hbar\partial_t\psi(x,t) \,=\, \hat{H}\,\psi(x_1,x_2,x_3,t)\,,\quad \hat{H} \,=\, -\frac{\hbar^2}{2m}\,\left(\,\partial_1^2\,+\,\partial_2^2\,+\,\partial_3^2\,\right) \tag{3.118}$$

which are stationary states. The corresponding solutions have the form

$$\psi(x_1,x_2,x_3,t) \,=\, \exp\left(-\frac{i}{\hbar}\,E\,t\right)\,\phi_E(x_1,x_2,x_3) \tag{3.119}$$

where $\phi_E(x_1,x_2,x_3)$ is an element of the function space $\mathcal{F}_3$ defined in (3.116) and obeys the partial differential equation

$$\hat{H}\,\phi_E(x_1,x_2,x_3) \,=\, E\,\phi_E(x_1,x_2,x_3)\,. \tag{3.120}$$

Since the Hamiltonian $\hat{H}$ is a sum of operators $O(x_j)$ each dependent only on a single variable, i.e., $\hat{H} \,=\, O(x_1)+O(x_2)+O(x_3)$, one can express

$$\phi(x_1,x_2,x_3) \,=\, \prod_{j=1}^{3}\phi^{(j)}(x_j) \tag{3.121}$$

where

$$-\frac{\hbar^2}{2m}\,\partial_j^2\,\phi^{(j)}(x_j) \,=\, E_j\,\phi^{(j)}(x_j)\,,\quad \phi^{(j)}(\pm a_j) \,=\, 0\,,\quad j\,=\,1,2,3 \tag{3.122}$$

and $E_1 + E_2 + E_3 = E$. Comparing (3.122) with (3.80) shows that the solutions of (3.122) are given by (3.96) and, hence, the solutions of (3.120) can be written

$$\begin{aligned} \phi_{(n_1,n_2,n_3)}(a_1,a_2,a_3;x_1,x_2,x_3) &= \phi_{n_1}(a_1;x_1)\,\phi_{n_2}(a_2;x_2)\,\phi_{n_3}(a_3;x_3) \\ n_1,n_2,n_3 &= 1,2,3,\ldots \end{aligned} \tag{3.123}$$

and

$$E_{(n_1,n_2,n_3)} = \frac{\hbar^2\pi^2}{8m}\left(\frac{n_1^2}{a_1^2}+\frac{n_2^2}{a_2^2}+\frac{n_3^2}{a_3^2}\right), n_1,n_2,n_3 = 1,2,3,\ldots \tag{3.124}$$

The same considerations as in the one-dimensional case allow one to show that

$$\mathcal{B}_3 = \{\phi_{(n_1,n_2,n_3)}(a_1,a_2,a_3;x_1,x_2,x_3)\,,\; n_1,n_2,n_3 = 1,2,3,\ldots\} \tag{3.125}$$

is a complete orthonormal basis of $\mathcal{F}_3$ and that the propagator for the three-dimensional box is

$$\begin{aligned} \phi(\vec{r},t|\vec{r}_0,t_0) = \sum_{n_1,n_2,n_3=1}^{\infty} & \phi_{(n_1,n_2,n_3)}(a_1,a_2,a_3;\vec{r}) \\ & \exp\left(-\frac{i}{\hbar}E_{(n_1n_2n_3)}\,(t-t_0)\right)\phi_{(n_1,n_2,n_3)}(a_1,a_2,a_3;\vec{r}_0)\,. \end{aligned} \tag{3.126}$$

Symmetries

The three-dimensional box confining a particle allows three symmetry operations that leave the box unchanged, namely rotation by π around the x_1, x_2, x_3–axes. This symmetry has been exploited in deriving the stationary states. If two or all three orthogonal sides of the box have the same length further symmetry operations leave the system unaltered. For example, if all three lengths a_1, a_2, a_3 are identical, i.e., $a_1 = a_2 = a_3 = a$ then rotation around the x_1, x_2, x_3–axes by $\pi/2$ also leaves the system unaltered. This additional symmetry is also reflected by degeneracies in the energy levels. The energies and corresponding degeneracies of the particle in the three-dimensional box with all side lengths equal to $2a$ are given in the following Table:

n_1	n_2	n_3	$E/[\hbar^2\pi^2/8ma^2]$	degeneracy
1	1	1	3	single
1	1	2	6	three-fold
1	2	2	9	three-fold
1	1	3	11	three-fold
2	2	2	12	single
1	2	3	14	six-fold
2	2	3	17	three-fold
1	1	4	18	three-fold
2	3	3	22	three-fold
3	3	3	27	single
1	1	5	27	three-fold

One can readily verify that the symmetry of the box leads to three-fold and six-fold degeneracies. Such degeneracies are always a signature of an underlying symmetry. Actually, in the present case

'accidental' degeneracies also occur, e.g., for $(n_1, n_2, n_3) = (3,3,3)\,(1,1,5)$ as shown in the Table above. The origin of this degeneracy is, however, the identity $3^2 + 3^2 + 3^2 = 1^2 + 1^2 + 5^2$.
One particular aspect of the degeneracies illustrated in the Table above is worth mentioning. We consider the degeneracy of the energy E_{122} which is due to the identity $E_{122} = E_{212} = E_{221}$. Any linear combination of wave functions

$$\tilde{\phi}(\vec{r}) \;=\; \alpha\,\phi_{(1,2,2)}(a,a,a;\vec{r}) \;+\; \beta\,\phi_{(2,1,2)}(a,a,a;\vec{r}) \;+\; \gamma\,\phi_{(2,2,1)}(a,a,a;\vec{r}) \tag{3.127}$$

obeys the stationary Schrödinger equation $\hat{H}\,\tilde{\phi}(\vec{r}) = E_{122}\,\tilde{\phi}(\vec{r})$. However, this linear combination is not necessarily orthogonal to other degeneraste states, for example, $\phi_{(1,2,2)}(a,a,a;\vec{r})$. Hence, in case of degenerate states one cannot necessarily expect that stationary states are orthogonal. However, in case of an n–fold degeneracy it is always possible, due to the hermitian character of $\hat{H}$, to construct n orthogonal stationary states[6].

[6]This is a result of linear algebra which the reader may find in a respective textbook.

Chapter 4

Linear Harmonic Oscillator

The linear harmonic oscillator is described by the Schrödinger equation

$$i\,\hbar\,\partial_t\,\psi(x,t) \;=\; \hat{H}\,\psi(x,t) \tag{4.1}$$

for the Hamiltonian

$$\hat{H} \;=\; -\,\frac{\hbar^2}{2m}\,\frac{\partial^2}{\partial x^2} \;+\; \frac{1}{2}m\,\omega^2 x^2\;. \tag{4.2}$$

It comprises one of the most important examples of elementary Quantum Mechanics. There are several reasons for its pivotal role. The linear harmonic oscillator describes vibrations in molecules and their counterparts in solids, the phonons. Many more physical systems can, at least approximately, be described in terms of linear harmonic oscillator models. However, the most eminent role of this oscillator is its linkage to the boson, one of the conceptual building blocks of microscopic physics. For example, bosons describe the modes of the electromagnetic field, providing the basis for its quantization. The linear harmonic oscillator, even though it may represent rather non-elementary objects like a solid and a molecule, provides a window into the most elementary structure of the physical world. The most likely reason for this connection with fundamental properties of matter is that the harmonic oscillator Hamiltonian (4.2) is symmetric in momentum and position, both operators appearing as quadratic terms.

We have encountered the harmonic oscillator already in Sect. 2 where we determined, in the context of a path integral approach, its propagator, the motion of coherent states, and its stationary states. In the present section we approach the harmonic oscillator in the framework of the Schrödinger equation. The important role of the harmonic oscillator certainly justifies an approach from two perspectives, i.e., from the path integral (propagotor) perspective and from the Schrödinger equation perspective. The path integral approach gave us a direct route to study time-dependent properties, the Schrödinger equation approach is suited particularly for stationary state properties. Both approaches, however, yield the same stationary states and the same propagator, as we will demonstrate below.

The Schrödinger equation approach will allow us to emphasize the algebraic aspects of quantum theory. This Section will be the first in which an algebraic formulation will assume center stage. In this respect the material presented provides an important introduction to later Sections using Lie algebra methods to describe more elementary physical systems. Due to the pedagodical nature of this Section we will link carefully the algebraic treatment with the differential equation methods used so far in studying the Schrödinger equation description of quantum systems.

In the following we consider first the stationary states of the linear harmonic oscillator and later consider the propagator which describes the time evolution of any initial state. The stationary states of the harmonic oscillator have been considered already in Chapter 2 where the corresponding wave functions (2.235) had been determined. In the framework of the Schrödinger equation the stationary states are solutions of (4.1) of the form $\psi(x,t) = \exp(-iEt/\hbar)\phi_E(x,t)$ where

$$\hat{H}\,\phi_E(x) \;=\; E\,\phi_E(x)\;. \tag{4.3}$$

Due to the nature of the harmonic potential which does not allow a particle with finite energy to move to arbitrary large distances, all stationary states of the harmonic oscillator must be bound states and, therefore, the natural boundary conditions apply

$$\lim_{x\to\pm\infty}\phi_E(x) \;=\; 0\;. \tag{4.4}$$

Equation (4.3) can be solved for any $E \in \mathbb{R}$, however, only for a discrete set of E values can the boundary conditions (4.4) be satisfied. In the following algebraic solution of (4.3) we restrict the Hamiltonian $\hat{H}$ and the operators appearing in the Hamiltonian from the outset to the space of functions

$$\mathcal{N}_1 \;=\; \{f:\mathbb{R}\to\mathbb{R},\, \mho\in\mathbb{C}_\infty,\, \lim_{\curvearrowleft\to\pm\infty}\mho(\curvearrowleft) = \not\curvearrowleft\} \tag{4.5}$$

where C_∞ denotes the set of functions which together with all of their derivatives are continuous. It is important to keep in mind this restriction of the space, in which the operators used below, act. We will point out explicitly where assumptions are made which built on this restriction. If this restriction would not apply and all functions $f:\mathbb{R}\to\mathbb{R}$ would be admitted, the spectrum of $\hat{H}$ in (4.3) would be continuous and the eigenfunctions $\phi_E(x)$ would not be normalizable.

4.1 Creation and Annihilation Operators

The Hamiltonian operator (4.2) can be expressed in terms of the two operators

$$\hat{p} \;=\; \frac{\hbar}{i}\frac{d}{dx}\;,\quad \hat{x} \;=\; x \tag{4.6}$$

the first being a differential operator and the second a multiplicative operator. The operators act on the space of functions $\mathcal{N}_1$ defined in (4.5). The Hamiltonian $\hat{H}$ can be expressed in terms of the operators acting on the space (4.5)

$$\hat{H} \;=\; \frac{1}{2m}\hat{p}^2 \;+\; \frac{1}{2}m\,\omega^2\,\hat{x}^2 \tag{4.7}$$

which is why these operators are of interest to us.
The cardinal property exploited below, beside the representation (4.7) of the Hamiltonian, is the commutation property

$$[\hat{p},\,\hat{x}] \;=\; \frac{\hbar}{i}\,\mathbb{1} \tag{4.8}$$

which holds for any position and momentum operator. This property states that $\hat{p}$ and $\hat{x}$ obey an algebra in which the two do not commute, however, the commutator has a simple form. In order

to derive (4.8) we recall that all operators act on functions in $\mathcal{N}_1$ and, hence, the action of $[\hat{p}, \hat{x}]$ on such functions must be considered. One obtains

$$[\hat{p}, \hat{x}]\, f(x) \;=\; \frac{\hbar}{i}\,\frac{d}{dx}\,x\,f(x) \;-\; x\,\frac{\hbar}{i}\,\frac{d}{dx}\,f(x) \;=\; \frac{\hbar}{i}\,f(x) \;=\; \frac{\hbar}{i}\,\mathbb{1}\,f(x)\;. \tag{4.9}$$

From this follows (4.8).
Our next step is an attempt to factorize the Hamiltonian (4.7) assuming that the factors are easier to handle than the Hamiltonian in yielding spectrum and eigenstates. Being guided by the identity for scalar numbers

$$(b - ic)(b + ic) \;=\; b^2 \;-\; i(cb - bc) \;+\; c^2 \;=\; b^2 \;+\; c^2 \tag{4.10}$$

we define

$$\hat{a}^+ \;=\; \sqrt{\frac{m\omega}{2\hbar}}\,\hat{x} \;-\; \frac{i}{\sqrt{2m\hbar\omega}}\,\hat{p}\;, \quad \hat{a}^- \;=\; \sqrt{\frac{m\omega}{2\hbar}}\,\hat{x} \;+\; \frac{i}{\sqrt{2m\hbar\omega}}\,\hat{p}\;. \tag{4.11}$$

The reader may note that we have attempted, in fact, to factor $\hat{H}/\hbar\omega$. Since $\hat{a}^+$ and $\hat{a}^-$ are operators and not scalars we cannot simply expect that the identy (4.10) holds for $\hat{a}^+$ and $\hat{a}^-$ since $[\hat{a}^-, \hat{a}^+] \;=\; \hat{a}^-\hat{a}^+ \;-\; \hat{a}^+\hat{a}^-$ does not necessarily vanish. In fact, the commutator property (4.8) implies

$$[\hat{a}^-, \hat{a}^+] \;=\; \mathbb{1}\;. \tag{4.12}$$

To prove this commutation property we determine using (4.11)

$$\hat{a}^-\hat{a}^+ \;=\; \frac{m\omega}{2\hbar}\,\hat{x}^2 \;+\; \frac{1}{2m\hbar\omega}\,\hat{p}^2 \;+\; \frac{i}{2\hbar}\,[\hat{p}, \hat{x}]\;. \tag{4.13}$$

(4.7) and (4.8) yield

$$\hat{a}^-\hat{a}^+ \;=\; \frac{1}{\hbar\omega}\,\hat{H} \;+\; \frac{1}{2}\,\mathbb{1}\;. \tag{4.14}$$

Similarly one can show

$$\hat{a}^+\hat{a}^- \;=\; \frac{1}{\hbar\omega}\,\hat{H} \;-\; \frac{1}{2}\,\mathbb{1}\;. \tag{4.15}$$

(4.14) and (4.15) together lead to the commutation property (4.12).
Before we continue we like to write (4.14, 4.15) in a form which will be useful below

$$\hat{H} \;=\; \hbar\omega\,\hat{a}^-\hat{a}^+ \;-\; \frac{\hbar\omega}{2}\,\mathbb{1} \tag{4.16}$$

$$\hat{H} \;=\; \hbar\omega\,\hat{a}^+\hat{a}^- \;+\; \frac{\hbar\omega}{2}\,\mathbb{1}\;. \tag{4.17}$$

We also express $\hat{a}^+$ and $\hat{a}^-$ directly in terms of the coordinate x and the differential operator $\frac{d}{dx}$

$$\hat{a}^+ \;=\; \sqrt{\frac{m\omega}{2\hbar}}\,x \;-\; \sqrt{\frac{\hbar}{2m\omega}}\,\frac{d}{dx}\;, \quad \hat{a}^- \;=\; \sqrt{\frac{m\omega}{2\hbar}}\,\hat{x} \;+\; \sqrt{\frac{\hbar}{2m\omega}}\,\frac{d}{dx}\;. \tag{4.18}$$

It is of interest to note that the operators $\hat{a}^+$ and $\hat{a}^-$ are real differential operators.

Relationship between a^+ and a^-

The operators $\hat{a}^+$ and $\hat{a}^-$ are related to each other by the following property which holds for all functions $f,\, g \in \mathcal{N}_1$

$$\int_{-\infty}^{+\infty} dx\, f(x)\, a^+\, g(x) \;=\; \int_{-\infty}^{+\infty} dx\, g(x)\, a^-\, f(x)\;. \tag{4.19}$$

This property states that the operators $\hat{a}^+$ and $\hat{a}^-$ are the adjoints of each other. The property follows directly from (**??**). Using (**??**) we like to state (4.19) in the form[1]

$$\langle f|\hat{a}^+ g\rangle_{\Omega_\infty} \;=\; \overline{\langle g|\hat{a}^- f\rangle_{\Omega_\infty}}\;. \tag{4.20}$$

In the following we will determine the spectrum of $\hat{H}$ and its eigenstates. The derivation will be based solely on the properties (4.12, 4.16, 4.17, 4.19).

$\hat{a}^+$ and $\hat{a}^-$ as Ladder Operators

The operators $\hat{a}^+$ and $\hat{a}^-$ allow one to generate all stationary states of a harmonic oscillator once one such state $\phi_E(x)$

$$\hat{H}\,\phi_E(x) \;=\; E\,\phi_E(x) \tag{4.21}$$

is available. In fact, one obtains using (4.16, 4.17, 4.21)

$$\begin{aligned}
\hat{H}\,\hat{a}^-\,\phi_E(x) \;&=\; (\,\hbar\omega\,\hat{a}^-\hat{a}^+ \;-\; \frac{\hbar\omega}{2}\,\mathbb{1}\,)\,\hat{a}^-\,\phi_E(x) \\
&=\; \hat{a}^-\,(\,\hbar\omega\,\hat{a}^+\hat{a}^- \;+\; \frac{\hbar\omega}{2}\,\mathbb{1} \;-\; \hbar\omega\,\mathbb{1}\,)\,\phi_E(x) \\
&=\; \hat{a}^-\,(\,\hat{H} \;-\; \hbar\omega\,)\,\phi_E(x) \\
&=\; (\,E \;-\; \hbar\omega\,)\,\hat{a}^-\,\phi_E(x)\;.
\end{aligned} \tag{4.22}$$

Similarly one can show using again (4.16, 4.17, 4.21)

$$\begin{aligned}
\hat{H}\,\hat{a}^+\,\phi_E(x) \;&=\; (\,\hbar\omega\,\hat{a}^+\hat{a}^- \;+\; \frac{\hbar\omega}{2}\,\mathbb{1}\,)\,\hat{a}^+\,\phi_E(x) \\
&=\; \hat{a}^+\,(\,\hbar\omega\,\hat{a}^-\hat{a}^+ \;-\; \frac{\hbar\omega}{2}\,\mathbb{1} \;+\; \hbar\omega\,\mathbb{1}\,)\,\phi_E(x) \\
&=\; \hat{a}^+\,(\,\hat{H} \;+\; \hbar\omega\,)\,\phi_E(x) \\
&=\; (\,E \;+\; \hbar\omega\,)\,\hat{a}^+\,\phi_E(x)\;.
\end{aligned} \tag{4.23}$$

Together, it holds that for a stationary state $\phi_E(x)$ of energy E defined through (4.21) $\hat{a}^-\phi_E(x)$ is a stationary state of energy $E - \hbar\omega$ and $\hat{a}^+\phi_E(x)$ is a stationary state of energy $E + \hbar\omega$.
The results (4.22, 4.23) can be generalized to m-fold application of the operators $\hat{a}^+$ and $\hat{a}^-$

$$\begin{aligned}
\hat{H}\;(\hat{a}^+)^m\;\phi_E(x) \;&=\; (\,E \;+\; m\,\hbar\omega\,)\;\left(\hat{a}^+\right)^m\;\phi_E(x) \\
\hat{H}\;(\hat{a}^-)^m\;\phi_E(x) \;&=\; (\,E \;-\; m\,\hbar\omega\,)\;\left(\hat{a}^-\right)^m\;\phi_E(x)\;.
\end{aligned} \tag{4.24}$$

[1] This property states that the operators in the function space $\mathcal{N}_1$ are the hermitian conjugate of each other. This property of operators is investigated more systematically in Section 5.

One can use these relationships to construct an infinite number of stationary states by stepping up and down the spectrum in steps of $\pm\hbar\omega$. For obvious reasons one refers to $\hat{a}^+$ and $\hat{a}^-$ as ladder operators. Another name is creation ($\hat{a}^+$) and annihilation ($\hat{a}^-$) operators since these operators create and annihilate vibrational quanta $\hbar\omega$. There arise, however, two difficulties: (i) one needs to know at least one stationary state to start the construction; (ii) the construction appears to yield energy eigenvalues without lower bounds when, in fact, one expects that $E = 0$ should be a lower bound. It turns out that both difficulties can be resolved together. In fact, a state $\phi_0(x)$ which obeys the property

$$\hat{a}^- \, \phi_0(x) \; = \; 0 \tag{4.25}$$

on one side would lead to termination of the sequence $E_0 + m$, $m \in \mathbb{Z}$ when m is decreased, on the other side such a state is itself an eigenstate of $\hat{H}$ as can be shown using (4.17)

$$\hat{H} \, \phi_0(x) \; = \; (\, \hbar\omega \, \hat{a}^+ \hat{a}^- \; + \; \frac{\hbar\omega}{2} \, \mathbb{1} \,) \, \phi_0(x) \; = \; \frac{1}{2} \hbar\omega \, \phi_0(x) \; . \tag{4.26}$$

Of course, the solution $\phi_0(x)$ of (4.25) needs to be normalizable in order to represent a bound state of the harmonic oscillator, i.e., $\phi_0(x)$ should be an element of the function space $\mathcal{N}_1$ defined in (4.5).

The property (**??**) has an important consequence for the stationary states $\phi_E(x)$. Let $\phi_E(x)$ and $\phi_{E'}(x)$ be two normalized stationary states corresponding to two *different* energies E, E', $E \neq E'$. For $f(x) = \phi_E(x)$ and $g(x) = \phi_{E'}(x)$ in (**??**) follows (Note that according to (**??**) the eigenvalue E is real.)

$$0 \; = \; \langle \phi_E | \hat{H} \phi_{E'} \rangle_{\Omega_\infty} \; - \; \overline{\langle \phi_{E'} | \hat{H} \phi_E \rangle_{\Omega_\infty}} \; = \; (\, E' \; - \; E \,) \, \langle \phi_E | \phi_{E'} \rangle_{\Omega_\infty} \; . \tag{4.27}$$

Since $E \neq E'$ one can conclude

$$\langle \phi_E | \phi_{E'} \rangle_{\Omega_\infty} \; = \; 0 \; . \tag{4.28}$$

4.2 Ground State of the Harmonic Oscillator

A suitable solution of (4.25) can, in fact, be found. Using (4.6, 4.11) one can rewrite (4.25)

$$\left(\frac{d}{dy} \; + \; y \right) \, \phi_0(y) \; = \; 0 \tag{4.29}$$

where

$$y \; = \; \sqrt{\frac{m\omega}{\hbar}} \, x \; . \tag{4.30}$$

Assuming that $\phi_0(y)$ does not vanish anywhere in its domain $]-\infty, +\infty[$ one can write (4.29)

$$\frac{1}{\phi_0(y)} \, \frac{d}{dy} \phi_0(y) \; = = \; \frac{d}{dy} \ln \phi_0(y) \; = \; - \, y \; , \tag{4.31}$$

the solution of which is

$$\ln \phi_0(y) \; = \; - \frac{1}{2} y^2 \; + \; c_0 \tag{4.32}$$

for some constant c_0 or

$$\phi_0(y) \; = \; c_0 \exp \left(- \frac{1}{2} y^2 \right) \; . \tag{4.33}$$

This solution is obviously normalizable. The conventional normalization condition

$$\langle \phi_0 | \phi_0 \rangle_{\Omega_\infty} \;=\; 1 \tag{4.34}$$

reads

$$\int_{-\infty}^{+\infty} dx\, |\phi_0(x)|^2 \;=\; |c_0|^2 \int_{-\infty}^{+\infty} dx \exp\left(-\frac{m\omega x^2}{2\hbar} \right) \;=\; |c_0|^2 \sqrt{\frac{\pi\hbar}{m\omega}}\,. \tag{4.35}$$

The appropriate ground state is

$$\phi_0(x) \;=\; \left[\frac{m\omega}{\pi\hbar} \right]^{\frac{1}{4}} \exp\left(-\frac{m\omega x^2}{2\hbar} \right)\,. \tag{4.36}$$

Since the first order differential equation (4.25) admits only one solution there is only one set of states with energy $E + m\,\hbar\omega$, $m \in \mathbb{Z}$ which properly terminate at some minimum value $E + m_0\,\hbar\omega \geq 0$. We recall that according to (4.26) the energy value associated with this state is $\frac{1}{2}\hbar\omega$. This state of lowest energy is called the *ground state* of the oscillator. The set of allowed energies of the oscillator according to (4.24) can then be written as follows

$$E_n \;=\; \hbar\omega\,(\, n \,+\, \frac{1}{2}\,)\,, \quad n \,=\, 0,\, 1,\, 2,\, \ldots \tag{4.37}$$

It is most remarkable that the energy $\frac{1}{2}\hbar\omega$ of the ground state is larger than the lowest classically allowed energy $E = 0$. The reason is that in the Hamiltonian (4.2) there are two competing contributions to the energy, the potential energy contribution which for a state $\delta(x)$, i.e., a state confined to the minimum corresponding to the classical state of lowest energy, would yield a vanishing contribution, and the kinetic energy contribution, which for a narrowly localized state yields a large positive value. The ground state (4.36) assumes a functional form such that both terms together assume a minimum value. We will consider this point more systematically in Section 5.

4.3 Excited States of the Harmonic Oscillator

Having obtained a suitable stationary state with lowest energy, we can now construct the stationary states corresponding to energies (4.37) above the ground state energy, i.e., we construct the states for $n = 1, 2, \ldots$, the so-called *excited states*. For this purpose we apply the operator $\hat{a}^+$ to the ground state (4.36) n times. Such states need to be suitably normalized for which purpose we introduce a normalization constant c'_n

$$\phi_n(x) \;=\; c'_n\, \left(\hat{a}^+\right)^n\, \phi_0(x)\,, \quad n \,=\, 0,\, 1,\, 2,\, \ldots \tag{4.38}$$

These states correspond to the energy eigenvalues (4.39), i.e., it holds

$$\hat{H}\,\phi_n(x) \;=\; \hbar\omega\,(\, n \,+\, \frac{1}{2}\,)\,\phi_n(x)\,, \quad n \,=\, 0,\, 1,\, 2,\, \ldots \tag{4.39}$$

We notice that the ground state wave function $\phi_0(x)$ as well as the operators $(\hat{a}^+)^n$ are real. We can, therefore, choose the normalization constants c'_n and the functions $\phi_n(x)$ real as well.

We need to determine now the normalization constants c'_n. To determine these constants we adopt a recursion scheme. For $n = 0$ holds $c'_0 = 1$. We consider then the situation that we have obtained a properly normalized state $\phi_n(x)$. A properly normalized state $\phi_{n+1}(x)$ is then of the form

$$\phi_{n+1}(x) \;=\; \alpha_n \, \hat{a}^+ \, \phi_n(x) \tag{4.40}$$

for some real constant α_n which is chosen to satisfy

$$\langle \phi_{n+1} | \phi_{n+1} \rangle_{\Omega_\infty} \;=\; \alpha_n^2 \, \langle \hat{a}^+ \phi_n | \hat{a}^+ \phi_n \rangle_{\Omega_\infty} \;=\; 1 \; . \tag{4.41}$$

Employing the adjoint property (4.20) yields

$$\alpha_n^2 \, \langle \phi_n | \hat{a}^- \hat{a}^+ \phi_n \rangle_{\Omega_\infty} \;=\; 1 \; . \tag{4.42}$$

Using (4.14) together with $\hat{H} \, \phi_n(x) \;=\; \hbar\omega(n+\frac{1}{2}) \, \phi_n(x)$ leads to the condition (note that we assumed $\phi_n(x)$ to be normalized)

$$\alpha_n^2 \, (\, n \, + \, 1 \,) \, \langle \phi_n | \hat{a}^- \hat{a}^+ \phi_n \rangle_{\Omega_\infty} \;=\; \alpha_n^2 \, (\, n \, + \, 1 \,) \;=\; 1 \tag{4.43}$$

From this follows $\alpha_n \;=\; 1/\sqrt{n+1}$ and, according to (4.40),

$$\hat{a}^+ \, \phi_n(x) \;=\; \sqrt{n+1} \, \phi_{n+1}(x) \; . \tag{4.44}$$

One can conclude then that the stationary states of the oscillator are described by the functions

$$\phi_n(x) \;=\; \frac{1}{\sqrt{n!}} \, \left(\hat{a}^+ \right)^n \, \phi_0(x) \; , \quad n \;=\; 0, 1, 2, \ldots \tag{4.45}$$

We like to note that these functions according to (4.28) and the construction (4.40–4.45) form an *orthonormal set*, i.e., they obey

$$\langle \phi_n | \phi_{n'} \rangle \;=\; \delta_{n \, n'} \; , \quad n, n' \;=\; 0, 1, 2, \ldots \; . \tag{4.46}$$

According to (4.24) holds in analogy to (4.40)

$$\phi_{n-1}(x) \;=\; \beta_n \, \hat{a}^- \, \phi_n(x) \tag{4.47}$$

for some suitable constants β_n. Since $\hat{a}^-$ is a real differential operator [see (4.18)] and since the $\phi_n(x)$ are real functions, β_n must be real as well. To determine β_n we note using (4.20)

$$1 \;=\; \langle \phi_{n-1} | \phi_{n-1} \rangle_{\Omega_\infty} \;=\; \beta_n^2 \, \langle \hat{a}^- \phi_n | \hat{a}^- \phi_n \rangle_{\Omega_\infty} \;=\; \beta_n^2 \, \langle \phi_n | \hat{a}^+ \hat{a}^- \phi_n \rangle_{\Omega_\infty} \; . \tag{4.48}$$

Equations (4.15 , 4.39) yield

$$1 \;=\; \beta_n^2 \, n \, \langle \phi_n | \phi_n \rangle_{\Omega_\infty} \;=\; \beta_n^2 \, n \; . \tag{4.49}$$

From this follows $\beta_n \;=\; 1/\sqrt{n}$ and, according to (4.47),

$$\hat{a}^- \, \phi_n(x) \;=\; \sqrt{n} \, \phi_{n-1}(x) \; . \tag{4.50}$$

Repeated application of this relationship yields

$$\phi_{n-s}(x) \;=\; \sqrt{\frac{(n-s)!}{n!}} \, \left(\hat{a}^- \right)^s \, \phi_n(x) \; . \tag{4.51}$$

Evaluating the Stationary States

We want to derive now an analytical expression for the stationary state wave functions $\phi_n(x)$ defined through (4.39). For this purpose we start from expression (4.45), simplifying the calculation, however, by introducing the variable y defined in (4.30) and employing the normalization

$$\int_{-\infty}^{+\infty} dy\, \phi_n^2(y) \;=\; 1 \tag{4.52}$$

This normalization of the wave functions differs from that postulated in (4.35) by a constant, n–independent factor, namely the square root of the Jacobian dx/dy, i.e., by

$$\sqrt{\left|\frac{dx}{dy}\right|} \;=\; \left[\frac{m\omega}{\hbar}\right]^{\frac{1}{4}} \;. \tag{4.53}$$

We will later re-introduce this factor to account for the proper normalization (4.35) rather than (4.52).
In terms of y the ground state wave function is

$$\phi_0(y) \;=\; \pi^{-\frac{1}{4}}\, e^{-\frac{y^2}{2}} \tag{4.54}$$

and $\hat{a}^+$ is

$$\hat{a}^+ \;=\; \frac{1}{\sqrt{2}} \left(y \;-\; \frac{d}{dy} \right) \;. \tag{4.55}$$

The eigenstates of the Hamiltonian are then given by

$$\phi_n(y) \;=\; \frac{1}{\sqrt{2^n\, n!\, \sqrt{\pi}}}\, e^{-\frac{y^2}{2}}\, e^{\frac{y^2}{2}} \left(y \;-\; \frac{d}{dy} \right)^n e^{-\frac{y^2}{2}} \;. \tag{4.56}$$

Relationship to Hermite Polynomials

We want to demonstrate now that the expression (4.56) can be expressed in terms of Hermite polynomials $H_n(y)$ introduced in Sect. 2.7 and given, for example, by the Rodrigues formula (2.200). We will demonstrate below the identity

$$H_n(y) \;=\; e^{\frac{y^2}{2}} \left(y \;-\; \frac{d}{dy} \right)^n e^{-\frac{y^2}{2}} \tag{4.57}$$

such that one can write the stationary state wave functions of the harmonic oscillator

$$\phi_n(y) \;=\; \frac{1}{\sqrt{2^n\, n!\, \sqrt{\pi}}}\, e^{-\frac{y^2}{2}}\, H_n(y) \;. \tag{4.58}$$

This result agrees with the expression (2.233) derived in Sect. 2.7. It should be noted that the normalization (4.28) of the ground state and the definition (4.57) which includes the factor $1/\sqrt{n!}$ according to Eqs. (4.40–4.46) yields a set of normalized states.

To verify the realtionship between the definition (4.57) of the *hermite* polynomials and the definition given by (2.200) we need to verify

$$\left(y \; - \; \frac{d}{dy} \right)^n \; e^{-\frac{y^2}{2}} \; = \; (-1)^n \, e^{\frac{y^2}{2}} \; \frac{d^n}{dy^n} \, e^{-y^2} \; . \tag{4.59}$$

which implies

$$H_n(y) \; = \; (-1)^n \, e^{y^2} \; \frac{d^n}{dy^n} \, e^{-y^2} \tag{4.60}$$

and, hence, the *Rodrigues formula* (2.200).
We prove (4.59) by induction noting first that (4.59) holds for $n \, = \, 0, 1$, and showing then that the property also holds for $n+1$ in case it holds for n, i.e.,

$$g(y) \; = \; (-1)^{n+1} \, e^{\frac{y^2}{2}} \; \frac{d^{n+1}}{dy^{n+1}} \, e^{-y^2} \; = \; \left(y \; - \; \frac{d}{dy} \right)^{n+1} \; e^{-\frac{y^2}{2}} \; . \tag{4.61}$$

One can factor $g(y)$ and employ (4.59) as follows

$$\begin{aligned} g(y) \quad &= \; -e^{\frac{y^2}{2}} \; \frac{d}{dy} \, e^{-\frac{y^2}{2}} \; (-1)^n \, e^{\frac{y^2}{2}} \; \frac{d^n}{dy^n} \, e^{-y^2} \\ &= \; -e^{\frac{y^2}{2}} \; \frac{d}{dy} \, e^{-\frac{y^2}{2}} \; \left(y \; - \; \frac{d}{dy} \right)^n \; e^{-\frac{y^2}{2}} \; . \end{aligned} \tag{4.62}$$

Denoting $f(y) \; = \; (y \; - \; d/dy) \exp(-y^2/2)$ and employing

$$\frac{d}{dy} \, e^{-\frac{y^2}{2}} \, f(y) \; = \; -y \, e^{-\frac{y^2}{2}} \, f(y) \; + \; e^{-\frac{y^2}{2}} \; \frac{d}{dy} \, f(y) \tag{4.63}$$

one obtains

$$g(y) \; = \; \left(y \; - \; \frac{d}{dy} \right) \, f(y) \tag{4.64}$$

which implies that (4.61) and, therefore, (4.59) hold.

4.4 Propagator for the Harmonic Oscillator

We consider now the solution of the time-dependent Schrödinger equation of the harmonic oscillator (4.1, 4.2) for an arbitrary initial wave function $\psi(x_0, t_0)$. Our derivation will follow closely the procedure adopted for the case of a 'particle in a box' [see Eqs. (3.106–3.114)]. For the sake of notational simplicity we employ initially the coordinate y as defined in (4.30) and return later to the coordinate x.
Starting point of our derivation is the assumption that the initial condition can be expanded in terms of the eigenstates $\phi_n(y)$ (4.39)

$$\psi(y_0, t_0) \; = \; \sum_{n=0}^{\infty} \, d_n \, \phi_n(y_0) \; . \tag{4.65}$$

Such expansion is possible for any $f(y_0) = \psi(y_0, t_0)$ which is an element of $\mathcal{N}_1$ defined in (4.5), a supposition which is not proven here[2]. Employing orthogonality condition (4.46) the expansion coefficients d_n are

$$d_n = \int_{-\infty}^{+\infty} dy_0 \, \psi(y_0, t_0) \, \phi_n(y_0) \, . \tag{4.66}$$

One can extend expansion (4.65) to times $t \geq t_0$ through insertion of time-dependent coefficients $c_n(t_1)$

$$\psi(y, t_1) = \sum_{n=0}^{\infty} d_n \, \phi_n(y) \, c_n(t) \tag{4.67}$$

for which according to (3.108–3.112) and (4.39) holds

$$c_n(t_1) = \exp\left(-i \, \omega \, (\, n + \frac{1}{2} \,) \, (\, t_1 - t_0 \,) \right) \, . \tag{4.68}$$

Altogether one can express then the solution

$$\psi(y, t_1) = \int_{-\infty}^{+\infty} dy_0 \, \phi(y, t | y_0, t_0) \, \psi(y_0, t) \tag{4.69}$$

where

$$\phi(y, t_1 | y_0, t_0) = \sum_{n=0}^{\infty} \phi_n(y) \, \phi_n(y_0) \, t^{n+\frac{1}{2}} \, , \quad t = e^{-i\omega(t_1 - t_0)} \tag{4.70}$$

is the propagator of the linear harmonic oscillator. We want to demonstrate now that this propagator is identical to the propagator (2.147) for the harmonic iscillator determined in Sect. sec:harm. In order to prove the equivalence of (4.70) and (2.147) we employ the technique of generating functions as in Sect. 2.7. For this purpose we start from the integral representation (2.225) which allows one to derive a generating function for products of *Hermite polynomials* which can be applied to the r.h.s. of (4.70). We consider for this purpose the following expression for $|t| < 1$

$$w(y, y_0, t) = \sum_{n=0}^{\infty} \frac{H_n(y) e^{-y^2/2} \, H_n(y_0) e^{-y_0^2/2}}{2^n \, n! \, \sqrt{\pi}} \, t^n \, . \tag{4.71}$$

Applying (**??**) to express $H_n(y)$ and $H_n(y_0)$ yields

$$\begin{aligned} w(y, y_0, t) = \; & \pi^{-3/2} \exp\left(\frac{y^2}{2} + \frac{y_0^2}{2} \right) \int_{-\infty}^{+\infty} du \int_{-\infty}^{+\infty} dv \overbrace{\sum_{n=0}^{\infty} \frac{1}{n!} \, (\, -2tuv \,)^n}^{\exp(-2tuv)} \times \\ & \times \exp(\, -u^2 - v^2 + 2iyu + 2iy_0 v \,) \, . \end{aligned} \tag{4.72}$$

[2] A demonstration of this property can be found in *Special Functions and their Applications* by N.N.Lebedev (Prentice Hall, Inc., Englewood Cliffs, N.J., 1965) Sect. 4.15, pp. 68; this is an excellent textbook from which we have borrowed heavily in this Section.

Carrying out the sum on the r.h.s. one obtains

$$w(y,y_0,t) \;=\; \pi^{-3/2}\exp\left(\frac{y^2}{2}\,+\,\frac{y_0^2}{2}\right)\int_{-\infty}^{+\infty} dv\exp(\,-v^2\,+\,2iy_0v\,)\;\times$$
$$\times\;\underbrace{\int_{-\infty}^{+\infty} du\exp(\,-u^2\,-2u(tv\,-\,iy)\,)}_{\sqrt{\pi}\exp(\,t^2v^2\,-\,2iytv\,-\,y^2)}\;. \tag{4.73}$$

The incomplete Gaussian integral

$$\int_{-\infty}^{+\infty} dx\, e^{-a^2x^2\,-\,2bx} \;=\; \frac{\sqrt{\pi}}{a}\,e^{b^2/a^2}\;,\quad \mathrm{Re}\,a^2\;>\;0 \tag{4.74}$$

applied once results in

$$w(y,y_0,t) \;=\; \pi^{-1}\exp\left(-\frac{y^2}{2}+\frac{y_0^2}{2}\right)\int_{-\infty}^{+\infty} dv\exp(-(1-t^2)\,v^2\,-\,2i\,(yt\,-\,y_0)\,v)\;. \tag{4.75}$$

Applying (4.74) a second time yields finally together with the definition (4.71) of $w(y,y_0,t)$

$$\tfrac{1}{\sqrt{\pi(1-t^2)}}\exp\left[\,-\tfrac{1}{2}(y^2\,+\,y_0^2)\,\tfrac{1+t^2}{1-t^2}\;+\;2\,y\,y_0\,\tfrac{t}{1-t^2}\,\right]$$
$$=\;\textstyle\sum_{n=0}^{\infty}\;\frac{H_n(y)e^{-y^2/2}\,H_n(y_0)e^{-y_0^2/2}}{2^n\,n!\,\sqrt{\pi}}\,t^n\;. \tag{4.76}$$

One can express this in terms of the stationary states (4.58) of the harmonic oscillator

$$\sum_{n=0}^{\infty}\phi_n(y)\,\phi_n(y_0)\,t^n \;=\; \frac{1}{\sqrt{\pi(1-t^2)}}\exp\left[\,-\frac{1}{2}(y^2\,+\,y_0^2)\,\frac{1\,+\,t^2}{1\,-\,t^2}\;+\;2\,y\,y_0\,\frac{t}{1\,-\,t^2}\,\right]\;. \tag{4.77}$$

The sum in (4.70) is, indeed, identical to the generating function (4.77), i.e., it holds

$$\phi(y,t_1|y_0,t_0) \;=\; \frac{1}{\sqrt{2\,i\,\pi\,F(t_1,t_0)}}\exp\left[\frac{i}{2}\,(y^2\,+\,y_0^2)\,G(t_1,t_0)\;-\;i\,y\,y_0\frac{1}{F(t_1,t_0)}\right] \tag{4.78}$$

where

$$F(t_1,t_0) \;=\; \frac{1\,-\,e^{-2i\omega(t_1-t_0)}}{2\,i\,e^{-i\omega(t_1-t_0)}} \;=\; \sin\!\omega(t_1-t_0) \tag{4.79}$$

and

$$G(t_1,t_0) \;=\; i\,\frac{1\,+\,e^{-2i\omega(t_1-t_0)}}{1\,-\,e^{-2i\omega(t_1-t_0)}} \;=\; \frac{\cos\!\omega(t_1-t_0)}{\sin\!\omega(t_1-t_0)}\;. \tag{4.80}$$

We can finally express the propagator (4.78) in terms of the coordinates x and x_0. This requires that we employ (4.30) to replace y and y_0 and that we multiply the propagator by $\sqrt{m\omega/\hbar}$, i.e., by a factor $\sqrt[4]{m\omega/\hbar}$ for both $\phi_n(y)$ and $\phi_n(y_0)$. The resulting propagator is

$$\phi(x,t|x_0,t_0) \;=\; \left[\frac{m\omega}{2i\pi\hbar\sin\!\omega(t_1-t_0)}\right]^{\frac{1}{2}}\;\times$$
$$\exp\left\{\frac{im\omega}{2\hbar\sin\!\omega(t_1-t_0)}\;\left[\,(x^2\,+\,x_0^2)\cos\!\omega(t_1-t_0)\;-\;2\,x\,x_0\,\right]\right\}\;. \tag{4.81}$$

This result agrees with the propagator (2.147) derived by means of the path integral description.

4.5 Working with Ladder Operators

In the last section we have demonstrated the use of differential equation techniques, the use of generating functions. We want to introduce now techniques based on the ladder operators $\hat{a}^+$ and $\hat{a}^-$. For the present there is actually no pressing need to apply such techniques since the techniques borrowed from the theory of differential equations serve us well in describing harmonic oscillator type quantum systems. The reason for introducing the calculus of the operators $\hat{a}^+$ and $\hat{a}^-$ is that this calculus proves to be useful for the description of vibrations in crystals, i.e., phonons, and of the modes of the quantized electromagnetic field; both quantum systems are endowed with a large number of modes, each corresponding to a single harmonic oscillator of the type studied presently by us. It is with quantum electrodynamics and solid state physics in mind that we cease the opportunity of the single quantum mechanical harmonic oscillator to develop a working knowledge for $\hat{a}^+$ and $\hat{a}^-$ in the most simple setting.
To put the following material in the proper modest perspective we may phrase it as an approach which rather than employing the coordinate y and the differential operator d/dy uses the operators

$$\hat{a}^+ \;=\; \frac{1}{\sqrt{2}}\left(\, y \;-\; \frac{d}{dy}\,\right)\,, \quad \hat{a}^- \;=\; \frac{1}{\sqrt{2}}\left(\, y \;+\; \frac{d}{dy}\,\right)\,. \tag{4.82}$$

Obviously, one can express $y \;=\; \sqrt{2}(\hat{a}^- + \hat{a}^+)$ and $d/dy \;=\; \sqrt{2}(\hat{a}^- - \hat{a}^+)$ and, hence, the approaches using y, d/dy and $\hat{a}^+$, $\hat{a}^-$ must be equivalent.

Calculus of Creation and Annihilation Operators

We summarize first the key properties of the operators $\hat{a}^+$ and $\hat{a}^-$

$$\begin{aligned}
[\hat{a}^-, \hat{a}^+] \quad &= \; \mathbb{1} \\
\hat{a}^-\,\phi_0(y) \quad &= \; 0 \\
\hat{a}^+\,\phi_n(y) \quad &= \; \sqrt{n+1}\,\phi_{n+1}(y) \\
\hat{a}^-\,\phi_n(y) \quad &= \; \sqrt{n}\,\phi_{n-1}(y) \\
\langle \hat{a}^-\, f | g \rangle_{\Omega_\infty} \quad &= \; \langle f | \hat{a}^+\, g \rangle_{\Omega_\infty}\,.
\end{aligned} \tag{4.83}$$

We note that these properties imply

$$\phi_n(y) \;=\; (\hat{a}^+)^n \phi_0(y)/\sqrt{n!}\,. \tag{4.84}$$

We will encounter below functions of $\hat{a}^+$ and $\hat{a}^-$, e.g., $f(\hat{a}^+)$. Such functions are defined for $f : \mathbb{R} \to \mathbb{R}$ in case that the Taylor expansion

$$f(y) \;=\; \sum_{\nu=0}^{\infty} \frac{f^{(\nu)}(0)}{\nu!}\, y^\nu \tag{4.85}$$

is convergent everywhere in $\mathbb{R}$. Here $f^{(\nu)}(y_0)$ denotes the ν-th derivative of $f(y)$ taken at $y = y_0$. In this case we define

$$f(\hat{a}^+) \;=\; \sum_{\nu=0}^{\infty} \frac{f^{(\nu)}(0)}{\nu!}\, \left(\hat{a}^+\right)^\nu \tag{4.86}$$

and similarly for $f(\hat{a}^-)$. The following important property holds

$$\hat{a}^- f(\hat{a}^+) \;=\; f^{(1)}(\hat{a}^+) \;+\; f(\hat{a}^+)\,\hat{a}^- \;. \tag{4.87}$$

In particular,

$$\hat{a}^- f(\hat{a}^+)\,\phi_0(y) \;=\; f^{(1)}(\hat{a}^+)\,\phi_0(y) \tag{4.88}$$

which follows from $\hat{a}^-\,\phi_o(y) \;=\; 0$. We note

$$\hat{a}^- f(\hat{a}^+) \;=\; [\hat{a}^-, f(\hat{a}^+)] \;+\; f(\hat{a}^+)\,\hat{a}^- \tag{4.89}$$

which implies that in order to prove (4.87, 4.88) we need to show actually

$$[\hat{a}^-, f(\hat{a}^+)] \;=\; f^{(1)}(\hat{a}^+) \;. \tag{4.90}$$

To prove (4.90) we show that (4.90) holds for any function $f_n(\hat{a}^+) \;=\; (\hat{a}^+)^n$ which is a power of $\hat{a}^+$. The convergence of the Taylor expansion ascertains then that (4.90) holds for $f(\hat{a}^+)$.
We proceed by induction noticing first that (4.90) holds for f_0 and for f_1. The first case is trivial, the second case follows from

$$[\hat{a}^-, f_1(\hat{a}^+)] \;=\; [\hat{a}^-, \hat{a}^+] \;=\; \mathbb{1} \;=\; f_1^{(1)}(\hat{a}^+) \;. \tag{4.91}$$

Let us assume that (4.90) holds for f_n. For f_{n+1} follows then

$$\begin{aligned} [\hat{a}^-, f_{n+1}(\hat{a}^+)] \;&=\; [\hat{a}^-, (\hat{a}^+)^n \hat{a}^+] \;=\; (\hat{a}^+)^n\,[\hat{a}^-, \hat{a}^+] \;+\; [\hat{a}^-, (\hat{a}^+)^n]\,\hat{a}^+ \\ &=\; (\hat{a}^+)^n\,\mathbb{1} \;+\; n\,(\hat{a}^+)^{n-1}\hat{a}^+ \;=\; (n+1)\,(\hat{a}^+)^n \end{aligned} \tag{4.92}$$

Since any function $f(y)$ in the proper function space $\mathcal{N}_1$ can be expanded

$$f(y) \;=\; \sum_{n=0}^{\infty} d_n\,\phi_n(y) \;=\; \sum_{n=0}^{\infty} d_n\,\frac{(\hat{a}^+)^n}{\sqrt{n!}}\,\phi_0(y) \tag{4.93}$$

one can reduce all operators acting on some proper state function by operators acting on the state $\phi_0(y)$. Hence, property (4.88) is a fundamental one and will be used in the following, i.e., we will only assume operator functions acting on $\phi_0(y)$. As long as the operators act on $\phi_0(y)$ one can state then that $\hat{a}^-$ behaves like a differential operator with respect to functions $f(\hat{a}^+)$. Note that an immediate consequence of (4.88) is

$$(\hat{a}^-)^n\,f(\hat{a}^+)\,\phi_o(y) \;=\; f^{(n)}(\hat{a}^+)\,\phi_0(y) \;. \tag{4.94}$$

We like to state the following property of the functions $f(\hat{a}^\pm)$

$$\langle f(\hat{a}^-)\,\phi|\psi\rangle \;=\; \langle\phi|f(\hat{a}^+)\,\psi\rangle \;. \tag{4.95}$$

This identity follows from (4.83) and can be proven for all powers f_n by induction and then inferred for all proper functions $f(\hat{a}^\pm)$.
An important operator function is the exponential function. An example is the so-called shift operator $\exp(u\hat{a}^-)$. It holds

$$e^{u\hat{a}^-}\,f(\hat{a}^+)\,\phi_0(y) \;=\; f(\hat{a}^+ \;+\; u)\,\phi_0(y) \;. \tag{4.96}$$

To prove this property we expand $\exp(u\hat{a}^-)$

$$\sum_{\nu=0}^{\infty} \frac{u^\nu}{\nu!} (\hat{a}^-)^\nu f(\hat{a}^+) \phi_0(y) \;=\; \sum_{\nu=0}^{\infty} \frac{u^\nu}{\nu!} f^{(\nu)}(\hat{a}^+) \phi_0(y) \;=\; f(\hat{a}^+ + u) \phi_0(y) \,. \tag{4.97}$$

An example in which (4.96) is applied is

$$e^{u\hat{a}^-} e^{v\hat{a}^+} \phi_0(y) \;=\; e^{v\,(\hat{a}^+ + u)} \phi_0(y) \;=\; e^{uv} e^{v\hat{a}^+} \phi_0(y) \,. \tag{4.98}$$

The related operators $\exp[v\hat{a}^+ \pm v^*\hat{a}^-]$ play also an important role. *We assume in the following derivation, without explicitly stating this, that the operators considered act on* $\phi_0(y)$. To express these operators as products of operators $\exp(v\hat{a}^+)$ and $\exp(v^*\hat{a}^-)$ we consider the operator $\hat{C}(u) = \exp(uz\hat{a}^+)\exp(uz^*\hat{a}^-)$ where $u \in \mathbb{R}$, $z \in \mathbb{C}$, $FF^* = \mathbb{1}$ and determine its derivative

$$\begin{aligned} \tfrac{d}{du}\hat{C}(u) \;&=\; z\hat{a}^+ e^{uz\hat{a}^+} e^{uz^*\hat{a}^-} + e^{uz\hat{a}^+} z^*\hat{a}^- e^{uz^*\hat{a}^-} \;= \\ &\left(z\hat{a}^+ + z^*\hat{a}^-\right) e^{uz\hat{a}^+} e^{uz^*\hat{a}^-} - \left[z^*\hat{a}^-, e^{uz\hat{a}^+}\right] e^{uz^*\hat{a}^-} \end{aligned} \tag{4.99}$$

Using (4.90) and $zz^* = 1$ we can write this

$$\begin{aligned} \tfrac{d}{du}\hat{C}(u) \;&=\; \left(z\hat{a}^+ + z^*\hat{a}^-\right) e^{uz\hat{a}^+} e^{uz^*\hat{a}^-} - u\, e^{uz\hat{a}^+} e^{uz^*\hat{a}^-} \;= \\ &\left(z\hat{a}^+ + z^*\hat{a}^- - u\right) \hat{C}(u) \,. \end{aligned} \tag{4.100}$$

To solve this differential equation we define $\hat{C}(u) = \hat{D}(u)\exp(-u^2/2)$ which leads to

$$\frac{d}{du}\hat{D}(u) \;=\; \left(z\hat{a}^+ + z^*\hat{a}^-\right) \hat{D}(u) \tag{4.101}$$

$\hat{C}(u)$ obviously obeys $\hat{C}(0) = \mathbb{1}$. This results in $\hat{D}(0) = \mathbb{1}$ and, hence, the solution of (4.101) is $\hat{D}(u) = \exp(uz\hat{a}^+ + uz^*\hat{a}^-)$. One can conclude then using the definition of $\hat{C}(u)$ and defining $v = uz$

$$e^{v\hat{a}^+ + v^*\hat{a}^-} \phi_0(y) \;=\; e^{\frac{1}{2}vv^*} e^{v\hat{a}^+} e^{v^*\hat{a}^-} \phi_0(y) \,. \tag{4.102}$$

Similarly, one obtains

$$e^{v\hat{a}^+ - v^*\hat{a}^-} \phi_0(y) \;=\; e^{-\frac{1}{2}vv^*} e^{v\hat{a}^+} e^{-v^*\hat{a}^-} \phi_0(y) \tag{4.103}$$

$$e^{v\hat{a}^+ + v^*\hat{a}^-} \phi_0(y) \;=\; e^{-\frac{1}{2}vv^*} e^{v^*\hat{a}^-} e^{v\hat{a}^+} \phi_0(y) \tag{4.104}$$

$$e^{v\hat{a}^+ - v^*\hat{a}^-} \phi_0(y) \;=\; e^{\frac{1}{2}vv^*} e^{-v^*\hat{a}^-} e^{v\hat{a}^+} \phi_0(y) \,. \tag{4.105}$$

Below we will use the operator identity which follows for the choice of $v = \alpha$ in (4.105)

$$e^{\alpha\hat{a}^+ - \alpha^*\hat{a}^-} \phi_0(y) \;=\; e^{\frac{\alpha\alpha^*}{2}} e^{-\alpha^*\hat{a}^-} e^{\alpha\hat{a}^+} \phi_0(y) \,. \tag{4.106}$$

Applying (4.98) for $u = -\alpha^*$ and $v = \alpha$ yields

$$e^{-\alpha^*\hat{a}^-} e^{\alpha\hat{a}^+} \phi_0(y) \;=\; e^{-\alpha\alpha^*} e^{\alpha\hat{a}^+} \phi_0(y) \,. \tag{4.107}$$

This together with (4.106) yields

$$e^{\alpha\hat{a}^+ - \alpha^*\hat{a}^-} \phi_0(y) \;=\; e^{-\frac{\alpha\alpha^*}{2}} e^{\alpha\hat{a}^+} \phi_0(y) \,. \tag{4.108}$$

Generating Function in Terms of $\hat{a}^+$

We want to demonstrate now that generating functions are as useful a tool in the calculus of the ladder operators as they are in the calculus of differential operators. We will derive the equivalent of a generating function and use it to rederive the values $H_n(0)$ and the orthonormality properties of $\phi_n(y)$.
We start from the generating function (**??**) and replace according to (4.58) the *Hermite* polynomials $H_n(y)$ by the eigenstates $\phi_n(y)$

$$e^{2yt\,-\,t^2} \;=\; \pi^{\frac{1}{4}} \sum_{n=0}^{\infty} \frac{(\sqrt{2}t)^n}{\sqrt{n!}}\, e^{y^2/2}\, \phi_n(y) \tag{4.109}$$

Using (4.84) and defining $\sqrt{2}t \;=\; u$ one can write then

$$\pi^{-\frac{1}{4}} \exp\left(-\frac{y^2}{2} \;+\; \sqrt{2}\,u\,y \;-\; \frac{u^2}{2}\right) \;=\; \sum_{n=0}^{\infty} \frac{u^n}{n!}\,(\hat{a}^+)^n\,\phi_0(y) \tag{4.110}$$

or

$$\pi^{-\frac{1}{4}} \exp\left(-\frac{y^2}{2} \;+\; \sqrt{2}\,u\,y \;-\; \frac{u^2}{2}\right) \;=\; e^{u\hat{a}^+}\,\phi_0(y)\;. \tag{4.111}$$

This expression, in the ladder operator calculus, is the equivalent of the generating function (**??**). We want to derive now the values $H_n(0)$. Setting $y \;=\; 0$ in (4.111) yields

$$\pi^{-\frac{1}{4}} \exp\left(-\frac{u^2}{2}\right) \;=\; e^{u\hat{a}^+}\,\phi_0(0)\;. \tag{4.112}$$

Expanding both sides of this equation and using (4.84, 4.58) one obtains

$$\pi^{-\frac{1}{4}} \sum_{\nu=0}^{\infty} \frac{(-1)^\nu u^{2\nu}}{2^\nu \nu!} \;=\; \sum_{\nu=0}^{\infty} \frac{u^\nu}{\sqrt{\nu!}}\,\phi_\nu(0) \;=\; \sum_{\nu=0}^{\infty} \frac{u^\nu}{\sqrt{2^\nu \sqrt{\pi}\,\nu!}}\,H_\nu(0) \tag{4.113}$$

Comparision of all terms on the l.h.s. and on the r.h.s. provide the same values for $H_n(0)$ as provided in Eq. (**??**).
Similarly, we can reproduce by means of the generating function (4.111) the orthonormality properties of the wave functions $\phi_n(y)$. For this purpose we consider

$$\langle e^{u\hat{a}^+}\,\phi_0 | e^{v\hat{a}^+}\,\phi_0\rangle \;=\; \langle \phi_0 | e^{u\hat{a}^-}\,e^{v\hat{a}^+}\,\phi_0\rangle \tag{4.114}$$

where we have employed (4.95). Using (4.98) and again (4.95) one obtains

$$\langle e^{u\hat{a}^+}\,\phi_0 | e^{v\hat{a}^+}\,\phi_0\rangle \;=\; \langle \phi_0 | e^{uv}\,e^{v\hat{a}^+}\,\phi_0\rangle \;=\; e^{uv}\,\langle e^{v\hat{a}^-}\,\phi_0 | \phi_0\rangle \;=\; e^{uv} \tag{4.115}$$

where the latter step follows after expansion of $e^{v\hat{a}^-}$ and using $\hat{a}^-\,\phi_0(y) \;=\; 0$. Expanding r.h.s. and l.h.s. of (4.115) and using (4.84) yields

$$\sum_{\nu,\mu=0}^{\infty} \frac{u^\mu\,v^\nu}{\sqrt{\mu!\nu!}}\,\langle \phi_\mu | \phi_\nu\rangle \;=\; \sum_{\mu=0}^{\infty} \frac{u^\mu\,v^\mu}{\mu!} \tag{4.116}$$

from which follows by comparision of each term on both sides $\langle \phi_\mu | \phi_\nu\rangle \;=\; \delta_{\mu\,\nu}$, i.e., the expected orthonormality property.

4.6 Momentum Representation for the Harmonic Oscillator

The description of the harmonic oscillator provided so far allows one to determine the probability density $P(x)$ of finding an oscillator at position x. For example, for an oscillator in a stationary state $\phi_n(x)$ as given by (2.235), the probability density is

$$P(x) \;=\; |\phi_n(x)|^2 \,. \tag{4.117}$$

In this section we want to provide a representation for the harmonic oscillator which is most natural if one wishes to determine for a stationary state of the system the probability density $\tilde{P}(p)$ of finding the system at momentum p. For this purpose we employ the Schrödinger equation in the momentum representation.
In the position representation the wave function is a function of x, i.e., is given by a function $\psi(x)$; the momentum and position operators are as stated in (4.6) and the Hamiltonian (4.7) is given by (4.2). In the momentum representation the wave function is a function of p, i.e., is given by the function $\tilde{\psi}(p)$; the momentum and position operators are

$$\hat{p} \;=\; p \,, \quad \hat{x} \;=\; i\hbar \frac{d}{dp} \,, \tag{4.118}$$

and the Hamiltonian (4.7) is

$$\hat{H} \;=\; \frac{1}{2m}\, p^2 \;-\; \frac{m\hbar^2\omega^2}{2}\, \frac{d^2}{dp^2} \,. \tag{4.119}$$

Accordingly, the time-independent Schrödinger equation for the oscillator can be written

$$\left(-\frac{m\omega^2\hbar^2}{2}\, \frac{d^2}{dp^2} \;+\; \frac{1}{2m}\, p^2 \right) \tilde{\phi}_E(p) \;=\; E\, \tilde{\phi}_E(p) \,. \tag{4.120}$$

Multiplying this equation by $1/m^2\omega^2$ yields

$$\left(-\frac{\hbar^2}{2m}\, \frac{d^2}{dp^2} \;+\; \frac{1}{2} m\tilde{\omega}^2\, p^2 \right) \tilde{\phi}_E(p) \;=\; \tilde{E}\, \tilde{\phi}_E(p) \tag{4.121}$$

where

$$\tilde{E} \;=\; E \,/\, m^2\omega^2 \,, \qquad \tilde{\omega} \;=\; 1 \,/\, m^2\omega \,. \tag{4.122}$$

For a solution of the Schrödinger equation in the momentum representation, i.e., of (4.121), we note that the posed eigenvalue problem is formally identical to the Schrödinger equation in the position representation as given by (4.2, 4.3). The solutions can be stated readily exploiting the earlier results. The eigenvalues, according to (4.37), are

$$\tilde{E}_n \;=\; \hbar\tilde{\omega}\, (n \,+\, \tfrac{1}{2}) \,, \quad n \;=\; 0, 1, 2, \ldots \tag{4.123}$$

or, using (4.122),

$$E_n \;=\; \hbar\omega\, (n \,+\, \tfrac{1}{2})) \,, \quad n \;=\; 0, 1, 2, \ldots \tag{4.124}$$

As to be expected, the eigenvalues are identical to those determined in the position representation. The momentum representation eigenfunctions are, using (2.235),

$$\tilde{\phi}_n(p) \;=\; \frac{(-i)^n}{\sqrt{2^n\, n!}} \left[\frac{m\tilde{\omega}}{\pi\hbar} \right]^{\frac{1}{4}} \exp\left(-\frac{m\tilde{\omega}p^2}{2\hbar} \right) H_n(\sqrt{\frac{m\tilde{\omega}}{\hbar}}\; p) \,. \tag{4.125}$$

or with (4.122)

$$\tilde{\phi}_n(p) \;=\; \frac{(-i)^n}{\sqrt{2^n\,n!}}\,\left[\frac{1}{\pi m\hbar\omega}\right]^{\frac{1}{4}}\exp\left(-\frac{p^2}{2m\hbar\omega}\right)\,H_n(\sqrt{\frac{1}{m\hbar\omega}}\,p)\;. \tag{4.126}$$

Normalized eigenfunctions are, of course, defined only up to an arbitrary phase factor $e^{i\alpha}$, $\alpha\in\mathbb{R}$; this allowed us to introduce in (4.125, 4.126) a phase factor $(-i)^n$.
We can now state the probability density $\tilde{P}_n(p)$ for an oscillator with energy E_n to assume momentum p. According to the theory of the momentum representation holds

$$\tilde{P}_n(p) \;=\; |\tilde{\phi}_n(p)|^2 \tag{4.127}$$

or

$$\tilde{P}_n(p) \;=\; \frac{1}{2^n\,n!}\,\left[\frac{1}{\pi m\hbar\omega}\right]^{\frac{1}{2}}\exp\left(-\frac{p^2}{m\hbar\omega}\right)\,H_n^2(\sqrt{\frac{1}{m\hbar\omega}}\,p)\;. \tag{4.128}$$

We may also express the eigenstates (4.126) in terms of dimensionless units following the procedure adopted for the position representation where we employed the substitution (4.30) to obtain (4.58). Defining $k \;=\; \sqrt{m\tilde{\omega}/\hbar}\,p$ or

$$k \;=\; \sqrt{\frac{1}{m\hbar\omega}}\,p \tag{4.129}$$

one obtains, instead of (4.126),

$$\phi_n(k) \;=\; \frac{(-i)^n}{\sqrt{2^n\,n!\,\sqrt{\pi}}}\,e^{-\frac{k^2}{2}}\,H_n(k)\;. \tag{4.130}$$

We want to finally apply the relationship

$$\tilde{\phi}_n(p) \;=\; \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}dx\,\exp(-ipx/\hbar)\,\phi_n(x) \tag{4.131}$$

in the present case employing dimensionless units. From (4.30) and (4.129) follows

$$k\,y \;=\; p\,x\,/\,\hbar \tag{4.132}$$

and, hence,

$$\tilde{\phi}_n(k) \;=\; \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}dy\,\exp(-iky)\,\phi_n(y) \tag{4.133}$$

where we note that a change of the normalization of $\phi_n(y)$ [c.f. (4.30) and (2.235)] absorbs the replacement $dx\;\rightarrow\;dy$, i.e., the Jacobian. This implies the property of Hermite polynomials

$$(-i)^n\,e^{-\frac{k^2}{2}}\,H_n(k) \;=\; \int_{-\infty}^{+\infty}dy\,\exp(-iky)\,e^{-\frac{y^2}{2}}\,H_n(y) \tag{4.134}$$

Exercise 4.6.1: Demonstrate the validity of (4.134), i.e., the correctness of the phase factors $(-i)^n$, through direct integration. Proceed as follows.

(a) Prove first the property

$$f(k,z) \;=\; \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} dy\, \exp(-iky)\, g(z,y) \tag{4.135}$$

where

$$f(k,z) \;=\; \exp\left(-2ikz \,+\, z^2 \,-\, k^2/2\right) \;, \qquad g(k,z) \;=\; \exp\left(2yz \,-\, z^2 \,-\, y^2/2\right) \tag{4.136}$$

(b) Employ the generating function (2.194, 2.196) of Hermite polynomials and equate coefficients of equal powers of z to prove (4.134).

4.7 Quasi-Classical States of the Harmonic Oscillator

A classical harmonic oscillator, for example, a pendulum swinging at small amplitudes, carries out a periodic motion described by

$$y(t) \;=\; \mathrm{Re}\left(\, y_0\, e^{i\omega t}\,\right) \tag{4.137}$$

where $y(t)$ describes the center of the particle and where the mass of the particle remains narrowly distributed around $y(t)$ for an arbitrary period of time. The question arises if similar states exist also for the quantum oscillator. The answer is 'yes'. Such states are referred to as *quasi-classical states*, *coherent states*, or *Glauber states* of the harmonic oscillator and they play, for example, a useful role in Quantum Electrodynamics since they allow to reproduce with a quantized field as closely as possible the properties of a classical field[3]. We want to construct and characterize such states.

The *quasi-classical states* can be obtained by generalizing the construction of the ground state of the oscillator, namely through the eigenvalue problem for normalized states

$$\hat{a}^-\, \phi^{(\alpha)}(y) \;=\; \alpha\, \phi^{(\alpha)}(y) \;, \quad \alpha \,\in\, \mathbb{R} \;, \quad \langle \phi^{(\alpha)} \,|\, \phi^{(\alpha)} \rangle \;=\; \mathbb{1} \;. \tag{4.138}$$

We will show that a harmonic oscillator prepared in such a state will remain forever in such a state, except that α changes periodically in time. We will find that α can be characterized as a displacement form the minimum of the oscillator and that the state $\phi^{(\alpha)}(y)$ displays the same spatial probability distribution as the ground state. The motion of the probability distribution of such state is presented in Fig. 4.1 showing in its top row the attributes of such state just discussed. We first construct the solution of (4.138). For this purpose we assume such state exists and expand

$$\phi^{(\alpha)}(y) \;=\; \sum_{n=0}^{\infty} d_n^{(\alpha)}\, \phi_n(y) \;, \quad \sum_{n=0}^{\infty} \left| d_n^{(\alpha)} \right|^2 \;=\; 1 \tag{4.139}$$

where we have added the condition that the states be normalized. Inserting this expansion into (4.138) yields using $\hat{a}^-\, \phi_n(y) \;=\; \sqrt{n}\, \phi_{n-1}(y)$

$$\sum_{n=0}^{\infty} \left(\sqrt{n}\, d_n^{(\alpha)}\, \phi_{n-1}(y) \;-\; \alpha\, d_n^{(\alpha)}\, \phi_n(y) \right) \;=\; 0 \tag{4.140}$$

[3] An excellent textbook is *Photons and Atoms: Introduction to Quantum Electrodynamics* by C.Chen-Tannoudji, J.Dupont-Roc, and G.Grynberg (John Wiley & Sons, Inc., New York, 1989) which discusses quasi-classical states of the free electromagnetic field in Sect.— III.C.4, pp. 192

or

$$\sum_{n=0}^{\infty} \left(\sqrt{n+1}\, d_{n+1}^{(\alpha)} \, - \, \alpha\, d_n^{(\alpha)} \right) \phi_n(y) \; = \; 0 \, . \tag{4.141}$$

Because of the linear independence of the states $\phi_n(y)$ all coefficients multiplying $\phi_n(y)$ must vanish and on obtains the recursion relationship

$$d_{n+1}^{(\alpha)} \; = \; \frac{\alpha}{\sqrt{n+1}}\, d_n^{(\alpha)} \, . \tag{4.142}$$

The solution is

$$d_n^{(\alpha)} \; = \; \frac{\alpha^n}{\sqrt{n!}}\, c \tag{4.143}$$

for a constant c which is determined through the normalization condition [see (4.139)]

$$1 \; = \; \sum_{n=0}^{\infty} \left| d_n^{(\alpha)} \right|^2 \; = \; |c|^2 \sum_{n=0}^{\infty} \frac{\alpha^2}{n!} \; = \; |c|^2\, e^{\alpha^2} \, . \tag{4.144}$$

Normalization requires $c \, = \, \exp(-\alpha^2/2)$ and, hence, the quasi-classical states are

$$\phi^{(\alpha)}(y) \; = \; \exp\left(-\frac{\alpha^2}{2} \right) \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}}\, \phi_n(y) \, , \tag{4.145}$$

Using the generating function in the form (4.109), i.e.,

$$[\pi]^{-\frac{1}{4}}\, e^{-\frac{y^2}{2} + \sqrt{2}\, \alpha\, y - \alpha^2} \; = \; \exp\left(-\frac{\alpha^2}{2} \right) \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}}\, \phi_n(y) \, , \tag{4.146}$$

one can write

$$\phi^{(\alpha)}(y) \; = \; [\pi]^{-\frac{1}{4}} \exp\left(-\frac{1}{2}(y \, - \, \sqrt{2}\, \alpha)^2 \right) \tag{4.147}$$

which identifies this state as a ground state of the oscillator displaced by $\sqrt{2}\alpha$. This interpretation justifies the choice of $\alpha \in \mathbb{R}$ in (4.138). However, we will see below that $\alpha \in \mathbb{C}$ are also admissible and will provide a corresponding interpretation.
One can also write (4.145) using (4.84)

$$\phi^{(\alpha)}(y) \; = \; \exp\left(-\frac{\alpha^2}{2} \right) \sum_{n=0}^{\infty} \frac{(\alpha \hat{a}^+)^n}{n!}\, \phi_0(y) \; = \; e^{-\frac{\alpha\alpha^*}{2}}\, e^{\alpha \hat{a}^+}\, \phi_0(y) \, . \tag{4.148}$$

The identity (4.108) allows one to write

$$\phi^{(\alpha)}(y) \; = \; T^+(\alpha)\, \phi_0(y) \; = \; T^+(\alpha)\, [\pi]^{-\frac{1}{4}} e^{-y^2/2} \tag{4.149}$$

where

$$T^+(\alpha) \; = \; e^{\alpha \hat{a}^+ \, - \, \alpha^* \hat{a}^-} \tag{4.150}$$

for α presently real. Comparing (4.149) with (4.147) allows one to interprete $T(\alpha)$ as an operator which shifts the ground state by $\sqrt{2}\alpha$.

The shift operator as defined in (4.150) for complex α obeys the unitary property

$$T^{+}(\alpha)\,T(\alpha) \;=\; \mathbb{1} \tag{4.151}$$

which follows from a generalization of (4.95) to linear combinations $\alpha\hat{a}^{+} + \beta\hat{a}^{-}$ and reads then

$$\langle f(\alpha\hat{a}^{+} + \beta\hat{a}^{-})\,\phi|\psi\rangle \;=\; \langle\phi|f(\alpha^{*}\hat{a}^{-} + \beta^{*}\hat{a}^{+})\,\psi\rangle\;. \tag{4.152}$$

Since the operator function on the r.h.s. is the adjoint of the operator function on the l.h.s. one can write in case of $T^{+}(\alpha)$

$$\big(T^{+}(\alpha)\big)^{+} \;=\; T(\alpha) \;=\; e^{\alpha^{*}\hat{a}^{-}\,-\,\alpha\hat{a}^{+}} \;=\; e^{-\,(\alpha\hat{a}^{+}\,-\,\alpha^{*}\hat{a}^{-})}\;. \tag{4.153}$$

This is obviously the inverse of $T^{+}(\alpha)$ as defined in (4.150) and one can conclude that $T^{+}(\alpha)$ is a so-called unitary operator with the property

$$T^{+}(\alpha)\,T(\alpha) \;=\; T(\alpha)\,T^{+}(\alpha) \;=\; \mathbb{1} \tag{4.154}$$

or applying this to a scalar product like (4.152)

$$\langle T^{+}(\alpha)\,f\,|T^{+}(\alpha)\,g\rangle \;=\; \langle f\,|T(\alpha)\,T^{+}(\alpha)\,g\rangle \;=\; \langle f\,|g\rangle\;. \tag{4.155}$$

In particular, it holds

$$\langle T^{+}(\alpha)\,\phi_0\,|T^{+}(\alpha)\,\phi_0\rangle \;=\; \langle f\,|T(\alpha)\,T^{+}(\alpha)\,g\rangle \;=\; \langle\phi_0\,|\phi_0\rangle\;,\quad \alpha\,\in\,\mathbb{C} \tag{4.156}$$

, i.e., the shift operator leaves the ground state normalized.
Following the procedure adopted in determining the propagator of the harmonic oscillator [see (4.67, 4.68)] one can write the time-dependent solution with (4.145) as the initial state

$$\phi^{(\alpha)}(y,t_1) \;=\; \exp\left(-\frac{\alpha^2}{2}\right)\sum_{n=0}^{\infty}\exp\left(-i\,\omega\,(\,n\,+\,\frac{1}{2}\,)\,(\,t_1\,-\,t_0\,)\right)\frac{\alpha^n}{\sqrt{n!}}\,\phi_n(y) \tag{4.157}$$

or

$$\phi^{(\alpha)}(y,t_1) \;=\; u^{\frac{1}{2}}\exp\left(-\frac{\alpha^2}{2}\right)\sum_{n=0}^{\infty}\frac{(\alpha\,u)^n}{\sqrt{n!}}\,\phi_n(y)\;,\quad u\;=\;e^{-i\omega(t_1\,-\,t_0)}\;. \tag{4.158}$$

Comparision of this expression with (4.145) shows that α in (4.145) is replaced by a complex number αu which at times $t_1 = t_0 + 2\pi n/\omega$, $n = 0, 1, \ldots$ becomes real. Relationship (4.109) serves again to simplify this sum

$$\phi^{(\alpha)}(y,t_1) \;=\; \pi^{-\frac{1}{4}}\,u^{\frac{1}{2}}\exp\Bigg(\underbrace{-\frac{y^2}{2}\;+\;\sqrt{2}\,y\,\alpha\,u\;-\;\frac{\alpha^2\,(u^2\,+\,1)}{2}}_{E}\Bigg) \tag{4.159}$$

Noticing the identity which holds for $|u|^2 = 1$

$$u^2\,+\,1 \;=\; (\mathrm{Re}\,u)^2\,-\,(\mathrm{Im}\,u)^2\,+\,1\,+\,2\,i\,\mathrm{Re}\,u\,\mathrm{Im}\,u \;=\; 2\,(\mathrm{Re}\,u)^2\,+\,2\,i\,\mathrm{Re}\,u\,\mathrm{Im}\,u \tag{4.160}$$

the exponent E can be written

$$\begin{aligned} E &= -\frac{y^2}{2} + \sqrt{2}\,\alpha y\,\mathrm{Re}u - \alpha^2\,(\mathrm{Re}\,u)^2 + i\sqrt{2}\,\alpha y\,\mathrm{Im}u - i\,\alpha^2\,\mathrm{Re}u\,\mathrm{Im}u \\ &= \underbrace{-\frac{1}{2}\left(y - \sqrt{2}\alpha\mathrm{Re}u\right)^2}_{\text{displaced Gaussian}} + \underbrace{i\sqrt{2}y\alpha\,\mathrm{Im}u}_{\text{momentum}} - \underbrace{i\,\alpha^2\mathrm{Re}u\,\mathrm{Im}u}_{\text{phase term}} \end{aligned} \tag{4.161}$$

and is found to describe a displaced Gaussian, a momentum factor and a phase factor. It is of interest to note that the displacement y_0 of the Gaussian as well as the momentum k_0 and phase ϕ_0 associated with (4.161) are time-dependent

$$\begin{aligned} y_0(t_1) &= \sqrt{2}\,\alpha\,\mathrm{cos}\omega(t_1 - t_0) \\ k_0(t_1) &= -\sqrt{2}\,\alpha\,\mathrm{sin}\omega(t_1 - t_0) \\ \phi_0(t_1) &= -\frac{1}{2}\alpha^2\mathrm{sin}2\omega(t_1 - t_0) \end{aligned} \tag{4.162}$$

with a periodic change. The oscillations of the mean position $y_0(t_1)$ and of the mean momentum $k_0(t_1)$ are out of phase by $\frac{\pi}{2}$ just as in the case of the classical oscillator. Our interpretation explains also the meaning of a complex α in (4.138): a real α corresponds to a displacement such that initially the oscillator is at rest, a complex α corresponds to a displaced oscillator which has initially a non-vanishing velocity; obviously, this characterization corresponds closely to that of the possible initial states of a classical oscillator. The time-dependent wave function (4.159) is then

$$\phi^{(\alpha)}(y, t_1) = \pi^{-\frac{1}{4}}\mathrm{exp}\left(-\frac{1}{2}[y - y_0(t_1)]^2 + iy\,k_0(t_1) - i\phi_0(t_1) - \frac{i}{2}\omega(t_1 - t_0)\right) . \tag{4.163}$$

This solution corresponds to the initial state given by (4.138, 4.147).
In Fig. 4.1 (top row) we present the probability distribution of the Glauber state for various instances in time. The diagram illustrates that the wave function retains its Gaussian shape with constant width for all times, moving solely its center of mass in an oscillator fashion around the minimum of the harmonic potential. This shows clearly that the Glauber state is a close analogue to the classical oscillator, except that it is not pointlike.
One can express (4.163) also through the shift operator. Comparing (4.148, 4.150) with (4.158) allows one to write

$$\phi^{(\alpha)}(y, t_1) = u^{\frac{1}{2}}\,T^+(\alpha\,u)\,\phi_0(y) \tag{4.164}$$

where according to (4.150)

$$T^+(\alpha\,u) = \mathrm{exp}\left(\alpha\,(\,e^{-i\omega(t_1 - t_0)}\hat{a}^+ - e^{i\omega(t_1 - t_0)}\hat{a}^-\,)\right) . \tag{4.165}$$

This provides a very compact description of the quasi-classical state.

Arbitrary Gaussian Wave Packet Moving in a Harmonic Potential

We want to demonstrate now that any initial state described by a Gaussian shows a time dependence very similar to that of the Glauber states in that such state remains Gaussian, being displaced around the center with period $2\pi/\omega$ and, in general, experiences a change of its width (relative to

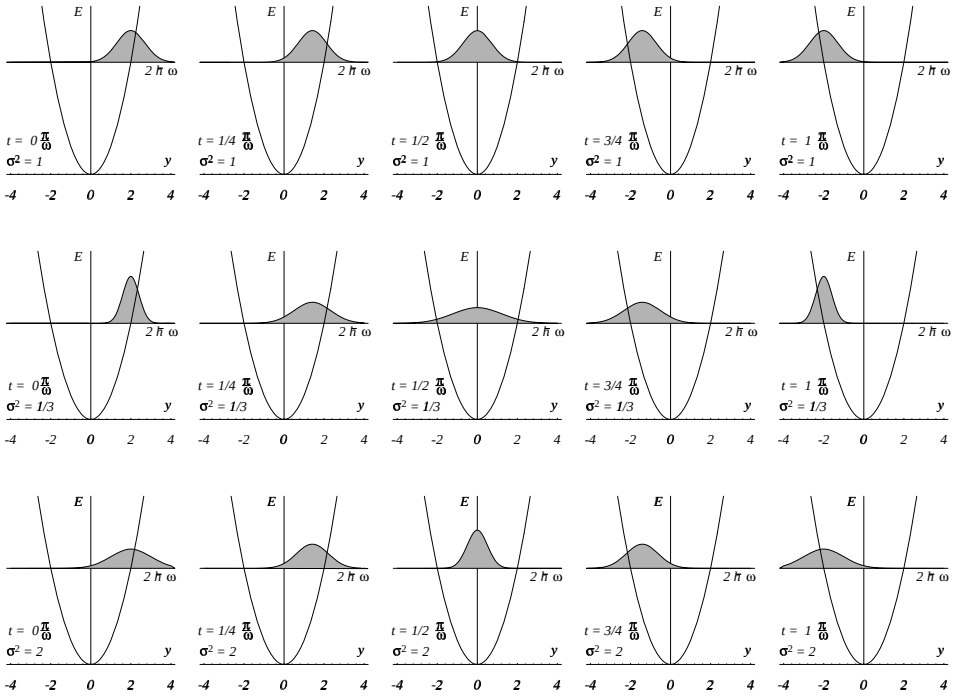


Figure 4.1: Time dependence of Gaussian wave packets in harmonic oscillator potential.

that of the Glauber states) with a period π/ω. This behaviour is illustrated in Fig. 4.1 (second and third row). One can recognize the cyclic displacement of the wave packet and the cyclic change of its width with a period of twice that of the oscillator.
To describe such state which satisfies the initial condition

$$\phi(a_o, \sigma|y, t_0) \;=\; [\pi\sigma^2]^{-\frac{1}{4}} \exp\left(-\frac{(y-a_o)^2}{2\sigma^2} \right) \tag{4.166}$$

we expand

$$\phi(a_o, \sigma|y, t_0) \;=\; \sum_{n=0}^{\infty} d_n \, \phi_n(y) \; . \tag{4.167}$$

One can derive for the expansion coefficients

$$\begin{aligned} d_n \;&=\; \int_{-\infty}^{+\infty} dy \, \phi_n(y) \, \phi(a_o, \sigma|y, t_0) \\ &=\; \frac{1}{\sqrt{2^n n! \pi\sigma}} \int_{-\infty}^{+\infty} dy \exp\left(-\frac{(y-a_o)^2}{2\sigma^2} - \frac{y^2}{2} \right) H_n(y) \\ &=\; \frac{e^{-a_o q/2}}{\sqrt{2^n n! \pi\sigma}} \int_{-\infty}^{+\infty} dy \, e^{-p(y-q)^2} \, H_n(y) \end{aligned} \tag{4.168}$$

where

$$p \;=\; \frac{1+\sigma^2}{2\sigma^2} \, , \quad q \;=\; \frac{a}{1+\sigma^2} \; . \tag{4.169}$$

The solution of integral (4.168) is[4]

$$d_n \;=\; \frac{e^{-aq/2}}{\sqrt{2^n n!\sigma}}\,\frac{(p-1)^{n/2}}{p^{(n+1)/2}}\,H_n(q\sqrt{\frac{p}{p-1}})\;. \tag{4.170}$$

In order to check the resulting coefficients we consider the limit $\sigma \to 1$. In this limit the coefficients d_n should become identical to the coefficients (4.143) of the Glauber state, i.e., a ground state shifted by $a_o = \sqrt{2}\alpha$. For this purpose we note that according to the explicit expresssion (**??**) for $H_n(y)$ the leading power of the Hermite polynomial is

$$H_n(y) \;=\; (2\,y)^n \;+\; \left\{\begin{array}{cc} O(y^{n-2}) & \text{for } n \geq 2 \\ 1 & \text{for } n = 0,1 \end{array}\right. \tag{4.171}$$

and, therefore,

$$\lim_{\sigma\to 1}\left(\frac{p-1}{p}\right)^{\frac{n}{2}} H_n(q\sqrt{\frac{p}{p-1}}) \;=\; \lim_{\sigma\to 1}(2q)^n \;=\; a^n\;. \tag{4.172}$$

One can then conclude

$$\lim_{\sigma\to 1} d_n \;=\; \frac{e^{-a^2/4}}{\sqrt{2^n n!}}\,a^n \tag{4.173}$$

which is, in fact, in agreement with the expected result (4.143).
Following the strategy adopted previously one can express the time-dependent state corresponding to the initial condition (4.166) in analogy to (4.157) as follows

$$\phi(a_o,\sigma|y,t_1) \;= \tag{4.174}$$
$$\tfrac{e^{-aq/2}}{\sqrt{p\sigma}}\,\textstyle\sum_{n=0}^{\infty} e^{-i\omega(n+\frac{1}{2})(t_1-t_o)}\,\frac{1}{\sqrt{2^n n!}}\,\left(\frac{p-1}{p}\right)^{n/2}\,H_n(q\sqrt{\frac{p}{p-1}})\,\phi_n(y)\;.$$

This expansion can be written

$$\begin{aligned}\phi(a_o,\sigma|y,t_1) \;&=\; \tfrac{1}{\sqrt{p\sigma\sqrt{\pi}}}\exp\Big(-\tfrac{1}{2}\tfrac{a^2}{1+\sigma^2} \;+\; \tfrac{1}{2}y_o^2 \;-\; \tfrac{1}{2}\omega(t_1-t_o)\Big)\;\times \\ &\times\;\textstyle\sum_{n=0}^{\infty}\frac{t^n}{2^n n!}\,H_n(y_o)\,e^{-y_o^2/2}\,H_n(y)\,e^{-y^2/2}\end{aligned} \tag{4.175}$$

where

$$y_o \;=\; q\sqrt{\frac{p}{p-1}} \;=\; \frac{a}{\sqrt{1\,-\,\sigma^4}}\;,\quad t \;=\; \sqrt{\frac{1\,-\,\sigma^2}{1\,+\,\sigma^2}}\,e^{-i\omega(t_1-t_o)} \tag{4.176}$$

The generating function (4.76) permits us to write this

$$\phi(a_o,\sigma|y,t_1) \;=\; \tfrac{1}{\sqrt{p\sigma\sqrt{\pi}}}\,\tfrac{1}{\sqrt{1-t^2}}\;\times \tag{4.177}$$
$$\times\,\exp\Big[-\tfrac{1}{2}(y^2\,+\,y_o^2)\,\tfrac{1+t^2}{1-t^2}\;-\;\tfrac{1}{2}\tfrac{a^2}{1+\sigma^2}\;+\;\tfrac{1}{2}\,y_o^2\;+\;2\,y\,y_o\,\tfrac{t}{1-t^2}\;-\;\tfrac{1}{2}\omega(t_1-t_o)\Big]\;.$$

[4]This integral can be found in *Integrals and Series, vol. 2* by A.P. Prudnikov, Yu. Brychkov, and O.I. Marichev (Wiley, New York, 1990); this 3 volume integral table is likely the most complete today, a worthy successor of the famous Gradshteyn, unfortunately very expensive, namely $750

Chapter 5

Theory of Angular Momentum and Spin

Rotational symmetry transformations, the group SO(3) of the associated rotation matrices and the corresponding transformation matrices of spin–$\frac{1}{2}$ states forming the group SU(2) occupy a very important position in physics. The reason is that these transformations and groups are closely tied to the properties of elementary particles, the building blocks of matter, but also to the properties of composite systems. Examples of the latter with particularly simple transformation properties are closed shell atoms, e.g., helium, neon, argon, the magic number nuclei like carbon, or the proton and the neutron made up of three quarks, all composite systems which appear spherical as far as their charge distribution is concerned. In this section we want to investigate how elementary and composite systems are described.

To develop a systematic description of rotational properties of composite quantum systems the consideration of rotational transformations is the best starting point. As an illustration we will consider first rotational transformations acting on vectors $\vec{r}$ in 3-dimensional space, i.e., $\vec{r} \in \mathbb{R}_{\not\Vdash}$, we will then consider transformations of wavefunctions $\psi(\vec{r})$ of single particles in $\mathbb{R}_{\not\Vdash}$, and finally transformations of products of wavefunctions like $\prod_{j=1}^{N} \psi_j(\vec{r}_j)$ which represent a system of N (spin-zero) particles in $\mathbb{R}_{\not\Vdash}$.

We will also review below the well-known fact that spin states under rotations behave essentially identical to angular momentum states, i.e., we will find that the algebraic properties of operators governing spatial and spin rotation are identical and that the results derived for products of angular momentum states can be applied to products of spin states or a combination of angular momentum and spin states.

5.1 Matrix Representation of the group SO(3)

In the following we provide a brief introduction to the group of three-dimensional rotation matrices. We will also introduce the generators of this group and their algebra as well as the representation of rotations through exponential operators. The mathematical techniques presented in this section will be used throughout the remainder of these notes.

Properties of Rotations in $\mathbb{R}_{\not\in}$

Rotational transformations of vectors $\vec{r} \in \mathbb{R}_{\not\in}$, in Cartesian coordinates $\vec{r} = (x_1, x_2, x_3)^T$, are linear and, therefore, can be represented by 3×3 matrices $R(\vec{\vartheta})$, where $\vec{\vartheta}$ denotes the rotation, namely around an axis given by the direction of $\vec{\vartheta}$ and by an angle $|\vec{\vartheta}|$. We assume the convention that rotations are right-handed, i.e., if the thumb in your right hand fist points in the direction of $\vec{\vartheta}$, then the fingers of your fist point in the direction of the rotation. $\vec{\vartheta}$ parametrizes the rotations uniquely as long as $|\vec{\vartheta}| < 2\pi$.
Let us define the rotated vector as

$$\vec{r}\,' \;=\; R(\vec{\vartheta})\,\vec{r} \quad . \tag{5.1}$$

In Cartesian coordinates this reads

$$x'_k \;=\; \sum_{j=1}^{3} [R(\vec{\vartheta})]_{kj}\, x_j \quad k \;=\; 1,2,3 \qquad . \tag{5.2}$$

Rotations conserve the scalar product between any pair of vectors $\vec{a}$ and $\vec{b}$, i.e., they conserve $\vec{a} \cdot \vec{b} = \sum_{j=1}^{3} a_j b_j$. It holds then

$$\vec{a}\,' \cdot \vec{b}\,' \;=\; \sum_{j,k,\ell=1}^{3} [R(\vec{\vartheta})]_{jk}\, [R(\vec{\vartheta})]_{j\ell}\, a_k\, b_\ell \;=\; \sum_{j=1}^{3} a_j b_j \quad . \tag{5.3}$$

Since this holds for any $\vec{a}$ and $\vec{b}$, it follows

$$\sum_{j=1}^{3} [R(\vec{\vartheta})]_{jk}\, [R(\vec{\vartheta})]_{j\ell} \;=\; \delta_{k\ell} \quad . \tag{5.4}$$

With the definition of the transposed matrix R^T

$$[R^T]_{jk} \;\equiv\; R_{kj} \tag{5.5}$$

this property can be written

$$R(\vec{\vartheta})\, R^T(\vec{\vartheta}) \;=\; R^T(\vec{\vartheta})\, R(\vec{\vartheta}) \;=\; \mathbb{1} \quad . \tag{5.6}$$

This equation states the key characteristic of rotation matrices. From (5.6) follows immediatety for the determinant of $R(\vec{\vartheta})$ using $\det A\, B \;=\; \det A \det B$ and $\det A^T \;=\; \det A$

$$\det R(\vec{\vartheta}) \;=\; \pm 1 \quad . \tag{5.7}$$

Let us consider briefly an example to illustrate rotational transformations and to interpret the sign of $\det R(\vec{\vartheta})$. A rotation around the x_3-axis by an angle φ is described by the matrix

$$R(\vec{\vartheta} \;=\; (0,0,\varphi)^T) \;=\; \pm \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{5.8}$$

In case of a prefactor $+1$, the matrix represents a proper rotation, in case of a prefactor -1 a rotation and an inversion at the origin. In the latter case the determinant of $R(\vec{\vartheta} \;=\; (0,0,\varphi)^T)$ is negative, i.e. a minus sign in Eq.(5.7) implies that a rotation is associated with an inversion. In the following we want to exclude inversions and consider only rotation matrices with $\det R \;=\; 1$.

Definition of the Group SO(3)

We will consider then the following set

$$SO(3) \;=\; \{\, \text{real}\, 3\times 3 \text{ matrices}\, R;\; R\,R^T \;=\; R^T R \;=\; \mathbb{1},\; \det R \;=\; 1\,\} \tag{5.9}$$

This set of matrices is closed under matrix multiplication and, in fact, forms a group $\mathcal{G}$ satisfying the axioms

(i) For any pair $a,\, b \in \mathcal{G}$ the product $c \;=\; a \circ b$ is also in $\mathcal{G}$.

(ii) There exists an element e in $\mathcal{G}$ with the property $\forall a \in \mathcal{G} \rightarrow e \circ a \;=\; a \circ e \;=\; a$.

(iii) $\forall a,\, a \in \mathcal{G} \rightarrow \exists a^{-1} \in \mathcal{G}$ with $a \circ a^{-1} \;=\; a^{-1} \circ a \;=\; e$.

(iv) For products of three elements holds the *associative law*
$a \circ (b \circ c) \;=\; (a \circ b) \circ c$.

We want to prove now that SO(3) as defined in (5.9) forms a group. For this purpose we must demonstrate that the group properties (i–iv) hold. In the following we will assume $R_1, R_2 \in \text{SO}(3)$ and do not write out "$\circ$" explicitly since it represents the matrix product.

(i) Obviously, $R_3 \;=\; R_1 R_2$ is a real, 3×3–matrix. It holds

$$R_3^T\, R_3 \;=\; (\, R_1\, R_2\,)^T\, R_1\, R_2 = \; R_2^T\, R_1^T\; R_1\, R_2 \;=\; R_2^T\; R_2 \;=\; \mathbb{1}\;. \tag{5.10}$$

Similarly, one can show $R_3 R_3^T \;=\; \mathbb{1}$. Furthermore, it holds

$$\det R_3 \;=\; \det(\, R_1\, R_2\,) \;=\; \det R_1 \det R_2 \;=\; 1\;. \tag{5.11}$$

It follows that R_3 is also an element of SO(3).

(ii) The group element 'e' is played by the 3×3 identity matrix $\mathbb{1}$.

(iii) From $\det R_1 \neq 0$ follows that R_1 is non-singular and, hence, there exists a real 3×3–matrix R_1^{-1} which is the inverse of R_1. We need to demonstrate that this inverse belongs also to SO(3). Since $(R_1^{-1})^T \;=\; (R_1^T)^{-1}$ it follows

$$(R_1^{-1})^T\, R_1^{-1} \;=\; (R_1^T)^{-1}\, R_1^{-1} \;=\; \left(R_1\, R_1^T\right)^{-1} \;=\; \mathbb{1}^{-1} \;=\; \mathbb{1} \tag{5.12}$$

which implies $R_1^{-1} \in \text{SO}(3)$.

(iv) Since the associative law holds for multiplication of any square matrices this property holds for the elements of SO(3).

We have shown altogether that the elements of SO(3) form a group.
We would like to point out the obvious property that for all elements R of SO(3) holds

$$R^{-1} \;=\; R^T \tag{5.13}$$

Exercise 5.1.1: Test if the following sets together with the stated binary operation '$\circ$' form groups; identify subgroups; establish if 1-1 homomorphic mappings exist, i.e. mappings between the sets which conserve the group properties:
(a) the set of real and imaginary numbers $\{1, i, -1, -i\}$ together with multiplication as the binary operation '$\circ$';
(b) the set of matrices $\{M_1,\, M_2,\, M_3,\, M_4\}$

$$= \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right\} \tag{5.14}$$

together with matrix multiplication as the binary operation '$\circ$';
(c) the set of four rotations in the x_1, x_2–plane {rotation by O^o, rotation by 90^o, rotation by 180^o, rotation by 270^o} and consecutive execution of rotation as the binary operation '$\circ$'.

Generators, Lie Group, Lie Algebra

We want to consider now a most convenient, compact representation of $R(\vec{\vartheta})$ which will be of general use and, in fact, will play a central role in the representation of many other symmetry transformations in Physics. This representation expresses rotation matrices in terms of three 3×3–matrices $J_1,\, J_2,\, J_3$ as follows

$$R(\vec{\vartheta}) \;=\; \exp\left(-\frac{i}{\hbar} \vec{\vartheta} \cdot \vec{J} \right) \tag{5.15}$$

Below we will construct appropriate matrices J_k, presently, we want to assume that they exist. We want to show that for the property $R\, R^T \;=\; 1$ to hold, the matrices

$$L_k \;=\; -\frac{i}{\hbar} J_k \tag{5.16}$$

must be antisymmetric. To demonstrate this property consider rotations with $\vec{\vartheta} \;=\; \vartheta_k \hat{e}_k$ where $\hat{e}_k$ is a unit vector in the direction of the Cartesian k–axis. Geometric intuition tells us

$$R(\vartheta_k \hat{e}_k)^{-1} \;=\; R(-\vartheta_k \hat{e}_k) \;=\; \exp\left(-\vartheta_k\, L_k \right) \tag{5.17}$$

Using the property which holds for matrix functions

$$[f(R)]^T \;=\; f(R^T) \tag{5.18}$$

we can employ (5.13)

$$R(\vartheta_k \hat{e}_k)^{-1} \;=\; R(\vartheta_k \hat{e}_k)^T \;=\; \left[\exp\left(\vartheta_k\, L_k \right)\right]^T \;=\; \exp\left(\vartheta_k\, L_k^T \right) \tag{5.19}$$

and, due to the uniqueness of the inverse, we arrive at the property

$$\exp\left(\vartheta_k\, L_k^T \right) \;=\; \exp\left(-\vartheta_k\, L_k \right) \tag{5.20}$$

from which follows $L_k^T \;=\; -L_k$.

We can conclude then that if elements of SO(3) can be expressed as $R = e^A$, the real matrix A is anti-symmetric. This property gives A then only three independent matrix elements, i.e., A must have the form

$$A = \begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix} . \tag{5.21}$$

It is this the reason why one can expect that three-dimensional rotation matrices can be expressed through (5.15) with three real parameters $\vartheta_1, \vartheta_2, \vartheta_3$ and three matrices L_1, L_2, L_3 which are independent of the rotation angles.
We assume now that we parameterize rotations through a three-dimensional rotation vector $\vec{\vartheta} = (\vartheta_1, \vartheta_2, \vartheta_3)^T$ such that for $\vec{\vartheta} = (0,0,0)^T$ the rotation is the identity. One can determine then the matrices L_k defined in (5.15) as the derivatives of $R(\vec{\vartheta})$ with respect to the Cartesian components ϑ_k taken at $\vec{\vartheta} = (0,0,0)^T$. Using the definition of partial derivatives we can state

$$L_1 = \lim_{\vartheta_1 \to 0} \vartheta_1^{-1} \left(R(\vec{\vartheta} = (\vartheta_1, 0, 0)^T) - \mathbb{1} \right) \tag{5.22}$$

and similar for L_2 and L_3.
Let us use this definition to evaluate L_3. Using (5.8) we can state

$$\lim_{\vartheta_3 \to 0} R[(0,0,\vartheta_3)^T] = \lim_{\vartheta_3 \to 0} \begin{pmatrix} \cos\vartheta_3 & -\sin\vartheta_3 & 0 \\ \sin\vartheta_3 & \cos\vartheta_3 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -\vartheta_3 & 0 \\ \vartheta_3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} . \tag{5.23}$$

Insertion into (5.22) for $k = 3$ yields

$$L_3 = \vartheta_3^{-1} \begin{pmatrix} 0 & -\vartheta_3 & 0 \\ \vartheta_3 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} . \tag{5.24}$$

One obtains in this way

$$-\frac{i}{\hbar} J_1 = L_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \tag{5.25}$$

$$-\frac{i}{\hbar} J_2 = L_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \tag{5.26}$$

$$-\frac{i}{\hbar} J_3 = L_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{5.27}$$

The matrices L_k are called the *generators* of the group SO(3).

Exercise 5.1.2: Determine the generator L_1 according to Eq.(5.22).

Sofar it is by no means obvious that the generators L_k allow one to represent the rotation matrices for any finite rotation, i.e., that the operators (5.15) obey the group property. The latter property implies that for any $\vec{\vartheta}_1$ and $\vec{\vartheta}_2$ there exists a $\vec{\vartheta}_3$ such that

$$\exp\left(-\frac{i}{\hbar}\vec{\vartheta}_1\cdot\vec{J}\right)\,\exp\left(-\frac{i}{\hbar}\vec{\vartheta}_2\cdot\vec{J}\right)\;=\;\exp\left(-\frac{i}{\hbar}\vec{\vartheta}_3\cdot\vec{J}\right) \tag{5.28}$$

The construction of the generators through the limit taken in Eq. (5.22) implies, however, that the generators represent infinitesimal rotations with $1 >> |\vec{\vartheta}|$ around the x_1-, x_2-, and x_3-axes. Finite rotations can be obtained by applications of many infinitesimal rotations of the type $R(\vec{\vartheta} = (\epsilon_1, 0, 0)^T) = \exp(\epsilon_1 L_1)$, $R(\vec{\vartheta} = (0, \epsilon_2, 0)^T) = \exp(\epsilon_2 L_2)$ and $R(\vec{\vartheta} = (0, 0, \epsilon_3)^T) = \exp(\epsilon_3 L_3)$. The question is then, however, if the resulting products of exponential operators can be expressed in the form (5.15). An answer can be found considering the Baker-Campbell-Hausdorf expansion

$$\exp(\,\lambda A\,)\exp(\,\lambda B\,)\;=\;\exp(\sum_{n=1}^{\infty}\lambda^n\,Z_n\,) \tag{5.29}$$

where $Z_1 = A + B$ and where the remaining operators Z_n are commutators of A and B, commutators of commutators of A and B, etc. The expression for Z_n are actually rather complicated, e.g. $Z_2 = \frac{1}{2}[A,\,B]$, $Z_3 = \frac{1}{12}(\,[A,\,[A,\,B]] + [[A,\,B],\,B]\,)$. From this result one can conclude that expressions of the type (5.15) will be closed under matrix multiplication only in the case that commutators of L_k can be expressed in terms of linear combinations of generators, i.e., it must hold

$$[L_k,\,L_\ell]\;=\;\sum_{m=1}^{3}f_{k\ell m}\,L_m\quad. \tag{5.30}$$

Groups with the property that their elements can be expressed like (5.15) and their generators obey the property (5.30) are called *Lie groups*, the property (5.30) is called the *Lie algebra*, and the constants $f_{k\ell m}$ are called the *structure constants*. Of course, Lie groups can have any number of generators, three being a special case.
In case of the group SO(3) the structure constants are particularly simple. In fact, the Lie algebra of SO(3) can be written

$$[L_k,\,L_\ell]\;=\;\epsilon_{k\ell m}\,L_m \tag{5.31}$$

where $\epsilon_{k\ell m}$ are the elements of the totally antisymmetric 3-dimensional tensor, the elements of which are

$$\epsilon_{k\ell m}\;=\;\begin{cases}0 & \text{if any two indices are identical}\\ 1 & \text{for } k=1,\ell=2,m=3\\ 1 & \text{for any even permutation of } k=1,\ell=2,m=3\\ -1 & \text{for any odd permutation of } k=1,\ell=2,m=3\end{cases}\quad. \tag{5.32}$$

For the matrices $J_k = i\hbar L_k$ which, as we see later, are related to angular momentum operators, holds the algebra

$$[J_k,\,J_\ell]\;=\;i\,\hbar\,\epsilon_{k\ell m}\,J_m\quad. \tag{5.33}$$

We want to demonstrate now that $\exp(\vec{\vartheta}\cdot\vec{L})$ yields the known rotational transformations, e.g., the matrix (5.8) in case of $\vartheta = (0, 0, \varphi)^T$. We want to consider, in fact, only the latter example and

determine $\exp(\varphi L_3)$. We note, that the matrix φL_3, according to (5.27), can be written

$$\varphi L_3 \;=\; \begin{pmatrix} A & & 0 \\ & & 0 \\ 0 \;\; 0 & & 0 \end{pmatrix}, \quad A \;=\; \varphi \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{5.34}$$

One obtains then for the rotational transformation

$$\begin{aligned} \exp\begin{bmatrix} A & & 0 \\ & & 0 \\ 0 \;\; 0 & & 0 \end{bmatrix} \;&=\; \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \sum_{\nu=1}^{\infty} \frac{1}{\nu!} \begin{pmatrix} A & & 0 \\ & & 0 \\ 0 \;\; 0 & & 0 \end{pmatrix}^{\nu} \\ &=\; \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \sum_{\nu=1}^{\infty} \frac{1}{\nu!} \begin{pmatrix} A^{\nu} & & 0 \\ & & 0 \\ 0 \;\; 0 & & 0 \end{pmatrix} \\ &=\; \begin{pmatrix} \sum_{\nu=0}^{\infty} \frac{1}{\nu!} A^{\nu} & & 0 \\ & & 0 \\ 0 \;\; 0 & & 1 \end{pmatrix} = \begin{pmatrix} e^{A} & & 0 \\ & & 0 \\ 0 \;\; 0 & & 1 \end{pmatrix}. \end{aligned} \tag{5.35}$$

To determine $expA = \sum_{\nu=1}^{\infty} A^{\nu}/\nu!$ we split its Taylor expansion into even and odd powers

$$e^{A} \;=\; \sum_{n=0}^{\infty} \frac{1}{(2n)!} A^{2n} \;+\; \sum_{n=0}^{\infty} \frac{1}{(2n+1)!} A^{2n+1}. \tag{5.36}$$

The property

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^{2} \;=\; -\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{5.37}$$

allows one to write

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^{2n} \;=\; (-1)^{n} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^{2n+1} \;=\; (-1)^{n} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{5.38}$$

and, accordingly,

$$e^{A} \;=\; \sum_{n=0}^{\infty} \frac{(-1)^{n}}{(2n)!} \varphi^{2n} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \;+\; \sum_{n=0}^{\infty} \frac{(-1)^{n}}{(2n+1)!} \varphi^{2n+1} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{5.39}$$

Recognizing the Taylor expansions of $\cos\varphi$ and $\sin\varphi$ one obtains

$$e^{A} \;=\; \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \tag{5.40}$$

and, using (5.34, 5.35),

$$e^{\varphi L_3} \;=\; \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{5.41}$$

which agrees with (5.8).

Exercise 5.1.3: Show that the generators L_k of SO(3) obey the commutation relationship (5.31).

5.2 Function space representation of the group SO(3)

In this section we will investigate how rotational transformations act on single particle wavefunctions $\psi(\vec{r})$. In particular, we will learn that transformation patterns and invariances are connected with the angular momentum states of quantum mechanics.

Definition

We first define how a rotational transformation acts on a wavefunction $\psi(\vec{r})$. For this purpose we require stringent continuity properties of the wavefunction: the wavefunctions under consideration must be elements of the set $\mathbb{C}_\infty(\mathbb{R})$, i.e., complex-valued functions over the 3-dimensional space $\mathbb{R}_\mathbb{R}$ which can be differentiated infinitely often. In analogy to $R(\vec{\vartheta})$ being a linear map $\mathbb{R}_\mathbb{R} \rightarrow \mathbb{R}_\mathbb{R}$, we define rotations $\mathcal{R}(\vec{\vartheta})$ as linear maps $\mathbb{C}_\infty(\mathbb{R}) \rightarrow \mathbb{C}_\infty(\mathbb{R})$, namely

$$\mathcal{R}(\vec{\vartheta})\psi(\vec{r}) \;=\; \psi(R^{-1}(\vec{\vartheta})\,\vec{r}) \quad . \tag{5.42}$$

Obviously, the transformation $\mathcal{R}(\vec{\vartheta})$: $\mathbb{C}_\infty(\mathbb{R}) \rightarrow \mathbb{C}_\infty(\mathbb{R})$ is related to the transformation $R(\vec{\vartheta})$: $\mathbb{R}_\mathbb{R} \rightarrow \mathbb{R}_\mathbb{R}$, a relationship which we like to express as follows

$$\mathcal{R}(\vec{\vartheta}) \;=\; \rho\left(R(\vec{\vartheta})\right) \quad . \tag{5.43}$$

$\rho(\ )$ conserves the group property of SO(3), i.e. for A, $B \in$ SO(3) holds

$$\rho(A\,B) \;=\; \rho(A)\,\rho(B) \quad . \tag{5.44}$$

This important property which makes $\rho(\ \mathrm{SO}(3)\ )$ also into a group can be proven by considering

$$\rho\left(R(\vec{\vartheta}_1)\,R(\vec{\vartheta}_2)\right)\,\psi(\vec{r}) \qquad = \;\psi([\,R(\vec{\vartheta}_1)\,R(\vec{\vartheta}_2)\,]^{-1}\vec{r}) \tag{5.45}$$

$$=\;\psi([R(\vec{\vartheta}_2)]^{-1}\,[R(\vec{\vartheta}_1)]^{-1}\vec{r}) \quad =\;\rho\left(\,R(\vec{\vartheta}_2)\right)\,\psi([R(\vec{\vartheta}_1)]^{-1}\vec{r}) \tag{5.46}$$

$$=\;\rho\left(\,R(\vec{\vartheta}_1)\right)\,\rho\left(\,R(\vec{\vartheta}_2)\right)\,\psi(\vec{r}) \quad . \tag{5.47}$$

Since this holds for any $\psi(\vec{r})$ one can conclude $\rho\left(\,R(\vec{\vartheta}_1)\right)\,\rho\left(\,R(\vec{\vartheta}_2)\right) \;=\; \rho\left(R(\vec{\vartheta}_1)\,R(\vec{\vartheta}_2)\right)$, i.e. the group property (5.44) holds.

Exercise 5.2.1: Assume the definition $\rho'\left(R(\vec{\vartheta})\right)\psi(\vec{r}) \;=\; \psi(R(\vec{\vartheta})\,\vec{r})$. Show that in this case holds $\rho'(A\,B) \;=\; \rho'(B)\,\rho'(A)$.

Generators

One can assume then that $\mathcal{R}(\vec{\vartheta})$ should also form a Lie group, in fact, one isomorphic to SO(3) and, hence, its elements can be represented through an exponential

$$\mathcal{R}(\vec{\vartheta}) \;=\; \exp\left(\,-\frac{i}{\hbar}\,\vec{\vartheta}\cdot\vec{\mathcal{J}}\,\right) \tag{5.48}$$

where the generators can be determined in analogy to Eq. (5.22) according to

$$-\frac{i}{\hbar}\mathcal{J}_1 \;=\; \lim_{\vartheta_1\to 0}\vartheta_1^{-1}\,\Big(\mathcal{R}(\vec{\vartheta}\,=\,(-\vartheta_1,0,0)^T)\;-\;1\Big) \tag{5.49}$$

(note the minus sign of $-\vartheta_1$ which originates from the inverse of the rotation in Eq.(5.42)) and similarly for $\mathcal{J}_2$ and $\mathcal{J}_3$.
The generators $\mathcal{J}_k$ can be determined by applying (5.49) to a function $f(\vec{r})$, i.e., in case of $\mathcal{J}_3$,

$$\begin{aligned} -\frac{i}{\hbar}\mathcal{J}_3\, f(\vec{r}) \;&=\; \lim_{\vartheta_3\to 0}\vartheta_3^{-1}\,\Big(\mathcal{R}(\vec{\vartheta}\,=\,(0,0,\vartheta_3)^T)\, f(\vec{r})\;-\;f(\vec{r})\Big) \\ &=\; \lim_{\vartheta_3\to 0}\vartheta_3^{-1}\,\big(\,f(\,R(0,0,-\vartheta_3)\,\vec{r}\,)\;-\;f(\vec{r})\,\big)\;. \end{aligned} \tag{5.50}$$

Using (5.23, 5.24) one can expand

$$R(0,0,-\vartheta_3)\,\vec{r}\;\approx\;\begin{pmatrix}1 & \vartheta_3 & 0\\ -\vartheta_3 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}\;=\;\begin{pmatrix}x_1+\vartheta_3 x_2\\ -\vartheta_3 x_1+x_2\;.\\ x_3\end{pmatrix} \tag{5.51}$$

Inserting this into (5.50) and Taylor expansion yields

$$\begin{aligned} -\frac{i}{\hbar}\mathcal{J}_3\, f(\vec{r}) \;&=\; \lim_{\vartheta_3\to 0}\vartheta_3^{-1}\,\big(\,f(\vec{r})\;+\;\vartheta_3\, x_2\,\partial_1 f(\vec{r})\;-\;\vartheta_3\, x_1\,\partial_2 f(\vec{r})\;-\;f(\vec{r})\,\big) \\ &=\; \big(\,x_2\,\partial_1\;-\;x_1\,\partial_2\,\big)\,f(\vec{r})\;. \end{aligned} \tag{5.52}$$

Carrying out similar calculations for $\mathcal{J}_1$ and $\mathcal{J}_2$ one can derive

$$-\frac{i}{\hbar}\mathcal{J}_1 \;=\; x_3\partial_2\;-\;x_2\partial_3 \tag{5.53}$$

$$-\frac{i}{\hbar}\mathcal{J}_2 \;=\; x_1\partial_3\;-\;x_3\partial_1 \tag{5.54}$$

$$-\frac{i}{\hbar}\mathcal{J}_3 \;=\; x_2\partial_1\;-\;x_1\partial_2\quad. \tag{5.55}$$

Not only is the group property of SO(3) conserved by $\rho(\;)$, but also the Lie algebra of the generators. In fact, it holds

$$[\mathcal{J}_k,\,\mathcal{J}_\ell]\;=\;i\,\hbar\,\epsilon_{k\ell m}\,\mathcal{J}_m\quad. \tag{5.56}$$

We like to verify this property for $[\mathcal{J}_1,\,\mathcal{J}_2]$. One obtains using (5.53–5.55) and $f(\vec{r})\,=\,f$

$$\begin{aligned} &[\mathcal{J}_1,\mathcal{J}_2]\,f(\vec{r})\;=\;-\hbar^2\,[-\tfrac{i}{\hbar}\mathcal{J}_1,-\tfrac{i}{\hbar}\mathcal{J}_2]\,f(\vec{r})\;=-\hbar^2\times \\ &[\;(x_3\partial_2\;-\;x_2\partial_3)\,(x_1\partial_3 f\;-\;x_3\partial_1 f)\;-\;(x_1\partial_3\;-\;x_3\partial_1)\,(x_3\partial_2 f\;-\;x_2\partial_3 f)\;] \\ &=\;-\hbar^2\,[\,+x_1x_3\partial_2\partial_3 f\;-x_1x_2\partial_3^2 f\;-x_3^2\partial_1\partial_2 f\;+x_2x_3\partial_1\partial_3 f\;+\;x_2\partial_1 f \\ &\quad -x_1\partial_2 f\;-x_1x_3\partial_2\partial_3 f\;+x_3^2\partial_1\partial_2 f\;+x_1x_2\partial_3^2 f\;-x_2x_3\partial_1\partial_3 f\;] \\ &\quad =\;-\hbar^2\,(\,x_2\partial_1\;-\;x_1\partial_2\,)\,f(\vec{r})\;=\;i\hbar\mathcal{J}_3\,f(\vec{r})\;. \end{aligned} \tag{5.57}$$

Exercise 5.2.2: (a) Derive the generators $\mathcal{J}_k$, $k=1,2$ by means of limits as suggested in Eq. (5.49); show that (5.56) holds for $[\mathcal{J}_2,\,\mathcal{J}_3]$ and $[\mathcal{J}_3,\,\mathcal{J}_1]$.

Exercise 5.2.3: Consider a wave function $\psi(\phi)$ which depends only on the azimutal angle ϕ of the spherical coordinate system (r, θ, ϕ), i.e. the system related to the Cartesian coordinates through $x_1 = r\sin\theta\cos\phi$, $x_2 = r\sin\theta\sin\phi$, and $x_3 = r\cos\theta$. Show that for $\mathcal{J}_3$ as defined above holds $\exp(\frac{i\alpha}{\hbar}\mathcal{J}_3)\,\psi(\phi) = \psi(\phi+\alpha)$.

5.3 Angular Momentum Operators

The generators $\mathcal{J}_k$ can be readily recognized as the angular momentum operators of quantum mechanics. To show this we first note that (5.53–5.55) can be written as a vector product of $\vec{r}$ and ∇

$$-\frac{i}{\hbar}\,\vec{\mathcal{J}} \;=\; -\,\vec{r}\,\times\,\nabla\;. \tag{5.58}$$

From this follows, using the momentum operator $\vec{p} = \frac{\hbar}{i}\nabla$,

$$\vec{\mathcal{J}} \;=\; \vec{r}\,\times\,\frac{\hbar}{i}\,\nabla \;=\; \vec{r}\times\vec{p} \tag{5.59}$$

which, according to the correspondence principle between classical and quantum mechanics, identifies $\vec{\mathcal{J}}$ as the quantum mechanical angular momentum operator.
Equation (5.56) are the famous commutation relationships of the quantium mechanical angular momentum operators. The commutation relationships imply that no simultaneous eigenstates of all three $\mathcal{J}_k$ exist. However, the operator

$$\mathcal{J}^2 \;=\; \mathcal{J}_1^2 \,+\mathcal{J}_2^2 \,+\mathcal{J}_3^2 \tag{5.60}$$

commutes with all three $\mathcal{J}_k$, $k = 1, 2, 3$, i.e.,

$$[\mathcal{J}^2, \mathcal{J}_k] \;=\; 0\;, \quad k = 1, 2, 3\;. \tag{5.61}$$

We demonstrate this property for $k = 3$. Using $[AB, C] = A[B, C] + [A, C]B$ and (5.56) one obtains

$$\begin{aligned}
[\mathcal{J}^2, \mathcal{J}_3] \;&=\; [\mathcal{J}_1^2 \,+\, \mathcal{J}_2^2, \mathcal{J}_3] \qquad (5.62)\\
&=\; \mathcal{J}_1\,[\mathcal{J}_1, \mathcal{J}_3] \,+\, [\mathcal{J}_1, \mathcal{J}_3]\,\mathcal{J}_1 \,+\, \mathcal{J}_2\,[\mathcal{J}_2, \mathcal{J}_3] \,+\, [\mathcal{J}_2, \mathcal{J}_3]\,\mathcal{J}_2\\
&=\; -i\hbar\,\mathcal{J}_1\,\mathcal{J}_2 \,-\, i\hbar\,\mathcal{J}_2\,\mathcal{J}_1 \,+\, i\hbar\,\mathcal{J}_2\,\mathcal{J}_1 \,+\, i\hbar\,\mathcal{J}_1\,\mathcal{J}_2 \;=\; 0
\end{aligned}$$

According to (5.61) simultaneous eigenstates of $\mathcal{J}^2$ and of one of the $\mathcal{J}_k$, usually chosen as $\mathcal{J}_3$, can be found. These eigenstates are the well-known spherical harmonics $Y_{\ell m}(\theta, \phi)$. We will show now that the properties of spherical harmonics

$$\begin{aligned}
\mathcal{J}^2\,Y_{\ell m}(\theta,\phi) \;&=\; \hbar^2\ell(\ell+1)\,Y_{\ell m}(\theta,\phi) \qquad (5.63)\\
\mathcal{J}_3\,Y_{\ell m}(\theta,\phi) \;&=\; \hbar\,m\,Y_{\ell m}(\theta,\phi) \qquad (5.64)
\end{aligned}$$

as well as the effect of the so-called raising and lowering operators

$$\mathcal{J}_\pm \;=\; \mathcal{J}_1 \,\pm\, i\,\mathcal{J}_2 \tag{5.65}$$

on the spherical harmonics, namely,

$$\begin{aligned}
\mathcal{J}_+ Y_{\ell m}(\theta,\phi) &= \hbar\sqrt{(\ell+m+1)(\ell-m)}\, Y_{\ell m+1}(\theta,\phi) && (5.66)\\
\mathcal{J}_+ Y_{\ell\ell}(\theta,\phi) &= 0 && (5.67)\\
\mathcal{J}_- Y_{\ell m+1}(\theta,\phi) &= \hbar\sqrt{(\ell+m+1)(\ell-m)}\, Y_{\ell m}(\theta,\phi) && (5.68)\\
\mathcal{J}_- Y_{\ell -\ell}(\theta,\phi) &= 0 && (5.69)
\end{aligned}$$

are a consequence of the Lie algebra (5.56). For this purpose we prove the following theorem.

An Important Theorem

Theorem 1.1 *Let L_k, $k = 1,2,3$ be operators acting on a Hilbert space $\mathbb{H}$ which obey the algebra $[L_k, L_\ell] = \epsilon_{k\ell m} L_m$ and let $L_\pm = L_1 \pm i\, L_2$, $L^2 = L_1^2 + L_2^2 + L_3^2$. Let there exist states $|\ell\, \ell\rangle$ in $\mathbb{H}$ with the property*

$$\begin{aligned}
L_+ |\ell\, \ell\rangle &= 0 && (5.70)\\
L_3 |\ell\, \ell\rangle &= -i\, \ell\, |\ell\, \ell\rangle && (5.71)
\end{aligned}$$

then the states defined through

$$L_- |\ell\, m+1\rangle = -i\, \beta_m\, |\ell\, m\rangle \qquad (5.72)$$

have the following properties

$$\begin{aligned}
(i)\quad L_+ |\ell\, m\rangle &= -i\, \alpha_m\, |\ell\, m+1\rangle && (5.73)\\
(ii)\quad L_3 |\ell\, m\rangle &= -i\, m\, |\ell\, m\rangle && (5.74)\\
(iii)\quad L^2 |\ell\, m\rangle &= -\,\ell(\ell+1)\, |\ell\, m\rangle && (5.75)
\end{aligned}$$

To prove this theorem we first show by induction that (i) holds. For this purpose we first demonstrate that the property holds for $m = \ell - 1$ by considering (i) for the state $L_-|\ell\ell\rangle \sim |\ell\ell - 1\rangle$. It holds $L_+L_-|\ell\ell\rangle = (L_+L_- - L_-L_+ + L_-L_+)|\ell\, \ell\rangle = ([L_+, L_-] + L_-L_+)|\ell\, \ell\rangle$. Noting $[L_+, L_-] = [L_1 + iL_2, L_1 - iL_2] = -2iL_3$ and $L_+|\ell\, \ell\rangle = 0$ one obtains $L_+L_-|\ell\ell\rangle = -2iL_3|\ell\ell\rangle = -2\ell|\ell\ell\rangle$, i.e., in fact, L_+ raises the m-value of $|\ell\ell - 1\rangle$ from $m = \ell - 1$ to $m = \ell$. The definitions of the coefficients $\alpha_{\ell-1}$ and $\beta_{\ell-1}$ yield $L_+L_-|\ell\ell\rangle = L_+(\, i\beta_{\ell-1}|\ell\ell - 1\rangle = -\,\alpha_{\ell-1}\beta_{\ell-1}\, |\ell\ell\rangle$, i.e.

$$\alpha_{\ell-1}\beta_{\ell-1} = 2\ell \quad . \qquad (*) \qquad (5.76)$$

We now show that property (ii) also holds for $m = \ell - 1$ proceeding in a similar way. We note $L_3L_-|\ell\ell\rangle = (\,[L_3, L_-] + L_-L_3\,)|\ell\ell\rangle$. Using $[L_3, L_-] = [L_3, L_1 - iL_2] = L_2 + iL_1 = i\,(L_1 - iL_2) = iL_-$ and $L_3|\ell\ell\rangle = -i\ell|\ell\ell\rangle$ we obtain $L_3L_-|\ell\ell\rangle = (\, iL_- + L_-L_3\,)|\ell\ell\rangle = i(-\ell + 1)\, L_-|\ell\ell\rangle$. From this follows that $L_-|\ell\ell\rangle$ is an eigenstate of L_3, and since we have already shown $|\ell\ell - 1\rangle = (i\beta_{\ell-1})^{-1}L_-|\ell\ell\rangle$ we can state $L_3|\ell\ell - 1\rangle = -i(\ell - 1)|\ell\ell - 1\rangle$.
Continuing our proof through induction we assume now that property (i) holds for m. We will show that this property holds then also for $m - 1$. The arguments are very similar to the ones used above and we can be brief. We consider $L_+L_-|\ell m\rangle = (\,[L_+, L_-] + L_-L_+\,)|\ell m\rangle = (\,-2iL_3 +$

$L_-L_+)|\ell m\rangle = (-2m - \alpha_m\beta_m)|\ell m\rangle$. This implies that (i) holds for $m-1$, in particular, from $L_+L_-|\ell m\rangle = -\alpha_{m-1}\beta_{m-1}|\ell m\rangle$ follows the recursion relationship

$$\alpha_{m-1}\beta_{m-1} = 2m + \alpha_m\beta_m \quad . \qquad (**) \tag{5.77}$$

We will show now that if (ii) holds for m, it also holds for $m-1$. Again the argumemts are similar to the ones used above. We note $L_3L_-|\ell m\rangle = ([L_3, L_-] + L_-L_3)|\ell m\rangle = (iL_- - i\,m\,L_-)|\ell m\rangle = -i\,(m-1)\,L_-|\ell m\rangle$ from which we can deduce $L_3|\ell m-1\rangle = -i(m-1)|\ell m-1\rangle$.
Before we can finally show that property (iii) holds as well for all $-\ell \le m \le \ell$ we need to determine the coefficients α_m and β_m. We can deduce from $(*)$ and $(**)$

$$\begin{aligned}
\alpha_{\ell-1}\beta_{\ell-1} &= 2\ell \\
\alpha_{\ell-2}\beta_{\ell-2} &= 2\,(\ell - 1 + \ell) \\
\alpha_{\ell-3}\beta_{\ell-3} &= 2\,(\ell - 2 + \ell - 1 + \ell) \\
&\cdot \\
&\cdot
\end{aligned}$$

Obviously, it holds $\alpha_m\beta_m = 2\left(\sum_{k=m+1}^{\ell} k\right)$. Using the familiar formula $\sum_{k=0}^{n} k = n\,(n+1)/2$ one obtains

$$\begin{aligned}
\alpha_m\beta_m &= (\ell+1)\,\ell - (m+1)\,m = (\ell+1)\,\ell + m\,\ell - \ell\,m - (m+1)\,m \\
&= (\ell + m + 1)(\ell - m)
\end{aligned} \tag{5.78}$$

One can normalize the states $|\ell m\rangle$ such that $\alpha_m = \beta_m$, i.e., finally

$$\alpha_m = \beta_m = \sqrt{(\ell + m + 1)(\ell - m)}\,. \tag{5.79}$$

We can now show that (iii) holds for all proper m–values. We note that we can write $L^2 = \frac{1}{2}L_+L_- + \frac{1}{2}L_-L_+ + L_3^2$. It follows then $L^2|\ell m\rangle = -\left(\frac{1}{2}\alpha_{m-1}\beta_{m-1} + \frac{1}{2}\alpha_m\beta_m + m^2\right)|\ell m\rangle = -\left(\frac{1}{2}(\ell+m+1)(\ell-m)\right.$
$\left. + \frac{1}{2}(\ell+m)(\ell-m+1) + m^2\right) = -\ell(\ell+1)|\ell m\rangle$. This completes the proof of the theorem.
We note that Eqs. (5.70, 5.72, 5.73, 5.79) read

$$L_+|\ell m\rangle = -i\sqrt{(\ell+m+1)(\ell-m)}\,|\ell m+1\rangle \tag{5.80}$$

$$L_-|\ell m+1\rangle = -i\sqrt{(\ell+m+1)(\ell-m)}\,|\ell m\rangle\,, \tag{5.81}$$

i.e., yield properties (5.66, 5.69).
We will have various opportunities to employ the Theorem just derived. A first application will involve the construction of the eigenstates of the angular momentum operators as defined in (5.63, 5.64). For this purpose we need to express the angular momentum operators in terms of spherical coordinates (r, θ, ϕ.

Angular Momentum Operators in Spherical Coordinates

We want to express now the generators $\mathcal{J}_k$, given in (5.53–5.55), in terms of spherical coordinates defined through

$$x_1 = r\sin\theta\cos\phi \tag{5.82}$$

$$x_2 = r\sin\theta\sin\phi \tag{5.83}$$

$$x_3 = r\cos\theta \tag{5.84}$$

We first like to demonstrate that the generators $\mathcal{J}_k$, actually, involve only the angular variables θ and ϕ and not the radius r. In fact, we will prove

$$-\frac{i}{\hbar}\mathcal{J}_1 \;=\; \mathcal{L}_1 \;=\; \sin\phi\frac{\partial}{\partial\theta} \;+\; \cot\theta\cos\phi\frac{\partial}{\partial\phi} \tag{5.85}$$

$$-\frac{i}{\hbar}\mathcal{J}_2 \;=\; \mathcal{L}_2 \;=\; -\cos\phi\frac{\partial}{\partial\theta} \;+\; \cot\theta\sin\phi\frac{\partial}{\partial\phi} \tag{5.86}$$

$$-\frac{i}{\hbar}\mathcal{J}_3 \;=\; \mathcal{L}_3 \;=\; -\frac{\partial}{\partial\phi} \quad . \tag{5.87}$$

To derive these properties we consider, for example, $\mathcal{L}_1 f(\theta,\phi) = (x_3\partial_2 - x_2\partial_3)f(\theta,\phi)$. Using (5.82–5.84) one obtains, applying repeatedly the chain rule,

$$\mathcal{L}_1 f(\theta,\phi) \;=\; (\,x_3\partial_2 \;-\; x_2\partial_3\,)\,f(\theta,\phi) \tag{5.88}$$

$$= \Big(\,x_3\frac{\partial x_2/x_1}{\partial x_2}\,\frac{\partial}{\partial\tan\phi} \;+\; x_3\frac{\partial x_3/r}{\partial x_2}\,\frac{\partial}{\partial\cos\theta} \tag{5.89}$$

$$- \;x_2\underbrace{\frac{\partial x_2/x_1}{\partial x_3}}_{=0}\;\tfrac{\partial}{\partial\tan\phi} \;-\; x_2\tfrac{\partial x_3/r}{\partial x_3}\,\tfrac{\partial}{\partial\cos\theta}\Big)\,f(\theta,\phi)$$

$$= \left(\frac{x_3}{x_1}\,\frac{\partial}{\partial\tan\phi} \;-\; \frac{x_3^2x_2}{r^3}\,\frac{\partial}{\partial\cos\theta} \;-\; x_2\frac{x_1^2\,+\,x_2^2}{r^3}\,\frac{\partial}{\partial\cos\theta}\right)\,f(\theta,\phi) \tag{5.90}$$

This yields, using $\partial/\partial\tan\phi = \cos^2\phi\,\partial/\partial\phi$ and $\partial/\partial\cos\theta = -(1/\sin\theta)\,\partial/\partial\theta$,

$$\mathcal{L}_1 f(\theta,\phi) \;=\; \left(\frac{\cos\theta}{\sin\theta\cos\phi}\cos^2\phi\frac{\partial}{\partial\phi} + \right. \tag{5.91}$$

$$\left. +\;\frac{\cos^2\theta\sin\theta\sin\phi}{\sin\theta}\frac{\partial}{\partial\theta} \;+\; \frac{\sin\theta\sin\phi}{\sin\theta}\sin^2\theta\frac{\partial}{\partial\theta}\right)\,f(\theta,\phi)$$

which agrees with expression (5.85).
We like to express now the operators $\mathcal{J}_3$ and $\mathcal{J}^2$, the latter defined in (5.60), in terms of spherical coordinates. According to (5.87) holds

$$\mathcal{J}_3 \;=\; \frac{\hbar}{i}\,\partial_\phi\,. \tag{5.92}$$

To determine $\mathcal{J}^2$ we note, using (5.85),

$$\begin{aligned}(\mathcal{L}_1)^2 \;&=\; \left(\sin\phi\,\frac{\partial}{\partial\theta} \;+\; \cot\theta\,\cos\phi\,\frac{\partial}{\partial\phi}\right)^2\\ &=\; \sin^2\phi\,\frac{\partial^2}{\partial\theta^2} \;-\; \frac{1}{\sin^2\theta}\sin\phi\cos\phi\,\frac{\partial}{\partial\phi} \;+\; 2\cot\theta\sin\phi\cos\phi\,\frac{\partial^2}{\partial\theta\partial\phi}\\ &\qquad +\cot\theta\cos^2\phi\,\frac{\partial}{\partial\theta} \;+\; \cot^2\theta\cos^2\phi\,\frac{\partial^2}{\partial\phi^2}\end{aligned} \tag{5.93}$$

and similarly, using (5.86),

$$(\mathcal{L}_2)^2 \;=\; \left(-\cos\phi\,\frac{\partial}{\partial\theta} \;+\; \cot\theta\,\sin\phi\,\frac{\partial}{\partial\phi}\right)^2$$

$$= \cos^2\phi \frac{\partial^2}{\partial\theta^2} + \frac{1}{\sin^2\theta}\sin\phi\cos\phi\frac{\partial}{\partial\phi} - 2\cot\theta\sin\phi\cos\phi\frac{\partial^2}{\partial\theta\partial\phi} + \cot\theta\sin^2\phi\frac{\partial}{\partial\theta} + \cot^2\theta\sin^2\phi\frac{\partial^2}{\partial\phi^2} . \quad (5.94)$$

It follows

$$(\mathcal{L}_1)^2 + (\mathcal{L}_2)^2 + (\mathcal{L}_3)^2 = \frac{\partial^2}{\partial\theta^2} + \cot\theta\frac{\partial}{\partial\theta} + (\cot^2\theta + 1)\frac{\partial^2}{\partial\phi^2} . \quad (5.95)$$

With $\cot^2\theta + 1 = 1/\sin^2\theta$,

$$\frac{1}{\sin^2\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)\left(\sin\theta\frac{\partial}{\partial\theta}\right) = \frac{\partial^2}{\partial\theta^2} + \cot\theta\frac{\partial}{\partial\theta} , \quad (5.96)$$

and the relationship between $\mathcal{L}_k$ and $\mathcal{J}_k$ as given in (5.85, 5.86, 5.87), one can conclude

$$\mathcal{J}^2 = -\frac{\hbar^2}{\sin^2\theta}\left[\left(\sin\theta\frac{\partial}{\partial\theta}\right)^2 + \frac{\partial^2}{\partial\phi^2}\right] . \quad (5.97)$$

Kinetic Energy Operator The kinetic energy of a classical particle can be expressed

$$\frac{\vec{p}^2}{2m} = \frac{p_r^2}{2m} + \frac{J^2_{\text{class}}}{2m\,r^2} \quad (5.98)$$

where $\vec{J}_{\text{class}} = \vec{r}\times m\vec{v}$ is the angular momentum and $p_r = m\dot{r}$ is the radial momentum. The corresponding expression for the quantum mechanical kinetic energy operator is

$$\hat{T} = -\frac{\hbar^2}{2m}\nabla^2 = -\frac{\hbar^2}{2m}\frac{1}{r}\frac{\partial^2}{\partial r^2}r + \frac{\mathcal{J}^2}{2m\,r^2} , \quad (5.99)$$

which follows from the identity

$$\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{1}{r^2\sin^2\theta}\left[\left(\sin\theta\frac{\partial}{\partial\theta}\right)^2 + \frac{\partial^2}{\partial\phi^2}\right] \quad (5.100)$$

$$\frac{1}{r}\frac{\partial^2}{\partial r^2}r = \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} \quad (5.101)$$

and comparision with (5.97).

5.4 Angular Momentum Eigenstates

According to the theorem on page 107 above the eigenfunctions (5.63, 5.64) can be constructed as follows. One first identifies the Hilbert space $\mathbb{H}$ on which the generators $\mathcal{L}_k$ operate. The operators in the present case are $\mathcal{L}_k = -\frac{i}{\hbar}\mathcal{J}_k$ where the $\mathcal{J}_k$ are differential operators in θ and ϕ given in (5.85, 5.86, 5.87). These generators act on a subspace of $\mathbb{C}_\infty(\mathbb{\nVdash})$, namely on the space of complex-valued functions on the unit 3-dim. sphere S_2, i.e. $\mathbb{C}_\infty(\mathbb{S}_{\nVdash})$. The functions in $\mathbb{C}_\infty(\mathbb{S}_{\nVdash})$ have real variables

θ, ϕ, $0 \leq \theta < \pi$, $0 \leq \phi < 2\pi$, and to be admissable for description of quantum states, must be cyclic in ϕ with period 2π. The norm in $\mathbb{H}$ is defined through

$$|f|^2 \;=\; \int_{S_2} d\Omega\, f^*(\theta,\phi) f(\theta,\phi) \tag{5.102}$$

where $d\Omega \;=\; \sin\theta\, d\theta\, d\phi$ is the volume element of S_2. The function space endowed with this norm and including only functions for which the integral (5.102) exists, is indeed a Hilbert space.
The eigenfunctions (5.63, 5.64) can be constructed then by seeking first functions $Y_{\ell\ell}(\theta,\phi)$ in $\mathbb{H}$ which satisfy $\mathcal{L}_+ Y_{\ell\ell}(\theta,\phi) \;=\; 0$ as well as $\mathcal{L}_3 Y_{\ell\ell}(\theta,\phi) \;=\; -i\ell\, Y_{\ell\ell}(\theta,\phi)$, normalizing these functions, and then determining the family $\{Y_{\ell m}(\theta,\phi), m = -\ell, -\ell+1, \ldots, \ell\}$ applying

$$\mathcal{L}_- Y_{\ell m+1}(\theta,\phi) \;=\; -i\sqrt{(\ell+m+1)(\ell-m)}\, Y_{\ell m}(\theta,\phi) \tag{5.103}$$

iteratively for $m = \ell-1, \ell-2, \ldots, -\ell$. One obtains in this way

$$Y_{\ell m}(\theta,\phi) \;=\; \Delta(\ell,m) \left(\frac{1}{\hbar}\mathcal{J}_-\right)^{\ell-m} Y_{\ell\ell}(\theta,\phi) \tag{5.104}$$

$$\Delta(\ell,m) \;=\; \left[\frac{(\ell\,+\,m)!}{(2\ell)!\,(\ell\,-\,m)!}\right]^{\frac{1}{2}} \tag{5.105}$$

Constructing $Y_{\ell\ell} \;\to\; Y_{\ell\ell-1} \;\to\!\to\; \cdots\, Y_{\ell-\ell}$

We like to characterize the resulting eigenfunctions $Y_{\ell m}(\theta,\phi)$, the so-called *spherical harmonics*, by carrying out the construction according to (5.104, 5.105) explicitly. For this purpose we split off a suitable normalization factor as well as introduce the assumption that the dependence on ϕ is described by a factor $\exp(im\phi)$

$$Y_{\ell m}(\theta,\phi) \;=\; \sqrt{\frac{2\ell+1}{4\pi}\frac{(\ell-m)!}{(\ell+m)!}}\, P_{\ell m}(\cos\theta)\, e^{im\phi}\;. \tag{5.106}$$

It must hold

$$\mathcal{L}_+\, Y_{\ell\ell}(\theta,\phi) \;=\; 0 \tag{5.107}$$

$$\mathcal{L}_3\, Y_{\ell\ell}(\theta,\phi) \;=\; -i\,\ell\, Y_{\ell\ell}(\theta,\phi)\;. \tag{5.108}$$

The latter property is obviously satisfied for the chosen ϕ-dependence. The raising and lowering operators

$$\mathcal{L}_\pm \;=\; \mathcal{L}_1 \;\pm\; i\,\mathcal{L}_2\;, \tag{5.109}$$

using (5.85–5.87), can be expressed

$$\mathcal{L}_\pm \;=\; \mp i\, e^{\pm i\phi} \left(\frac{\partial}{\partial\theta} \;\pm\; i\cot\theta\,\frac{\partial}{\partial\phi}\right)\;. \tag{5.110}$$

Accordingly, (5.107) reads

$$\left(\frac{\partial}{\partial\theta} \;+\; i\cot\theta\,\frac{\partial}{\partial\phi}\right) P_{\ell\ell}(\cos\theta)\, e^{i\ell\phi} \;=\; 0 \tag{5.111}$$

or

$$\left(\frac{\partial}{\partial\theta} \; - \; \ell \cot\theta \right) P_{\ell\ell}(\cos\theta) \; = \; 0 \tag{5.112}$$

where we employed the definition of $P_{\ell\ell}(\cos\theta)$ in (5.107). A proper solution of this equation, i.e., one for which

$$\int_0^\pi d\theta \sin\theta \, |P_{\ell\ell}(\cos\theta)|^2 \tag{5.113}$$

is finite, is

$$P_{\ell\ell}(\cos\theta) \; = \; N_\ell \sin^\ell\theta \tag{5.114}$$

as one can readily verify. The normalization factor N_ℓ is chosen such that

$$\int_0^\pi d\theta \sin\theta \, |Y_{\ell\ell}(\cos\theta)|^2 \; = \; 1 \tag{5.115}$$

holds. (5.106, 5.114) yield

$$|N_\ell|^2 \, \frac{2\ell+1}{4\pi} \, \frac{1}{(2\ell)!} \, 2\pi \int_0^\pi d\theta \sin^{2\ell+1}\theta \; = \; 1 \, . \tag{5.116}$$

The integral appearing in the last expression can be evaluated by repeated integration by parts. One obtains, using the new integration variable $x = \cos\theta$,

$$\begin{aligned}
\int_0^\pi d\theta \sin^{2\ell+1}\theta \; &= \; \int_{-1}^{+1} dx \, (1 \, - \, x^2)^\ell \\
&= \quad \Big[x \, (1 \, - \, x^2)^\ell \Big]_{-1}^{+1} \, + \, 2\ell \int_{-1}^{+1} dx \, x^2 \, (1 \, - \, x^2)^{\ell-1} \\
&= \quad 2\ell \left[\frac{x^3}{3} \, (1 \, - \, x^2)^{\ell-1} \right]_{-1}^{+1} \, + \, \frac{2\ell \cdot (2\ell-2)}{1 \cdot 3} \int_{-1}^{+1} dx \, x^4 \, (1 \, - \, x^2)^{\ell-2} \\
&\; \vdots \\
&= \quad \frac{2\ell \cdot (2\ell-2) \cdots 2}{1 \cdot 3 \cdot 5 \cdots (2\ell-1)} \int_{-1}^{+1} dx \, x^{2\ell} \\
&= \quad \frac{2\ell \cdot (2\ell-2) \cdots 2}{1 \cdot 3 \cdot 5 \cdots (2\ell-1)} \, \frac{2}{2\ell+1} \; = \; \frac{(2\ell)!}{[1 \cdot 3 \cdot 5 \cdots (2\ell-1)]^2} \, \frac{2}{2\ell+1}
\end{aligned} \tag{5.117}$$

Accordingly, (5.116) implies

$$|N_\ell|^2 \, \frac{1}{[1 \cdot 3 \cdot 5 \cdots (2\ell-1)]^2} \; = \; 1 \tag{5.118}$$

from which we conclude

$$N_\ell \; = \; (-1)^\ell \, 1 \cdot 3 \cdot 5 \cdots (2\ell-1) \tag{5.119}$$

where the factor $(-1)^\ell$ has been included to agree with convention[1]. We note that (5.106, 5.114, 5.119) provide the following expression for $Y_{\ell\ell}$

$$Y_{\ell\ell}(\theta,\phi) \; = \; (-1)^\ell \sqrt{\frac{2\ell+1}{4\pi} \, \frac{1}{(2\ell)!}} \, \frac{1}{2^\ell \ell!} \, \sin^\ell\theta \, e^{i\ell\phi} \, . \tag{5.120}$$

[1]See, for example, "Classical Electrodynamics, 2nd Ed." by J.D. Jackson (John Wiley, New York, 1975)

We have constructed a normalized solution of (5.106), namely $Y_{\ell\ell}(\theta,\phi)$, and can obtain now, through repeated application of (5.103, 5.110), the eigenfunctions $Y_{\ell m}(\theta,\phi)$, $m = \ell-1, \ell-2, \ldots -\ell$ defined in (5.63, 5.64). Actually, we seek to determine the functions $P_{\ell m}(\cos\theta)$ and, therefore, use (5.103, 5.110) together with (5.106)

$$\begin{aligned} & i\,e^{-i\phi}\left(\frac{\partial}{\partial\theta} - i\cot\theta\,\frac{\partial}{\partial\phi}\right)\sqrt{\frac{2\ell+1}{4\pi}\frac{(\ell-m-1)!}{(\ell+m+1)!}}\,P_{\ell m+1}(\cos\theta)\,e^{i(m+1)\phi} \\ &= -i\sqrt{(\ell+m+1)(\ell-m)}\sqrt{\frac{2\ell+1}{4\pi}\frac{(\ell-m)!}{(\ell+m)!}}\,P_{\ell m}(\cos\theta)\,e^{im\phi}. \end{aligned} \tag{5.121}$$

This expressions shows that the factor $e^{im\phi}$, indeed, describes the ϕ-dependence of $Y_{\ell m}(\theta,\phi)$; every application of $\mathcal{L}_-$ reduces the power of $e^{i\phi}$ by one. (5.121) states then for the functions $P_{\ell m}(\cos\theta)$

$$\left(\frac{\partial}{\partial\theta} + (m+1)\cot\theta\right)P_{\ell m+1}(\cos\theta) = -(\ell+m+1)(\ell-m)\,P_{\ell m}(\cos\theta)\,. \tag{5.122}$$

The latter identity can be written, employing again the variable $x = \cos\theta$,

$$\begin{aligned} P_{\ell m}(x) &= \frac{-1}{(\ell+m+1)(\ell-m)}\times \\ &\quad\times\left[-(1-x^2)^{\frac{1}{2}}\frac{\partial}{\partial x} + (m+1)\frac{x}{(1-x^2)^{\frac{1}{2}}}\right]P_{\ell m+1}(x)\,. \end{aligned} \tag{5.123}$$

We want to demonstrate now that the recursion equation (5.123) leads to the expression

$$P_{\ell m}(x) = \frac{1}{2^\ell \ell!}\frac{(\ell+m)!}{(\ell-m)!}(1-x^2)^{-\frac{m}{2}}\frac{\partial^{\ell-m}}{\partial x^{\ell-m}}(x^2-1)^\ell \tag{5.124}$$

called the *associated Legendre polynomials*. The reader should note that the associated Legendre polynomials, as specified in (5.124), are real. To prove (5.124) we proceed by induction. For $m = \ell$ (5.124) reads

$$\begin{aligned} P_{\ell\ell}(x) &= \frac{(2\ell)!}{2^\ell\ell!}(1-x^2)^{-\frac{\ell}{2}}(x^2-1)^\ell \\ &= (-1)^\ell\,1\cdot3\cdot5\cdots(2\ell-1)\,\sin^\ell\theta \end{aligned} \tag{5.125}$$

which agrees with (5.114, 5.119). Let us then assume that (5.124) holds for $m+1$, i.e.,

$$P_{\ell m+1}(x) = \frac{1}{2^\ell\ell!}\frac{(\ell+m+1)!}{(\ell-m-1)!}(1-x^2)^{-\frac{m+1}{2}}\frac{\partial^{\ell-m-1}}{\partial x^{\ell-m-1}}(x^2-1)^\ell\,. \tag{5.126}$$

Then holds, according to (5.123),

$$\begin{aligned} P_{\ell m}(x) &= \frac{-1}{2^\ell\ell!}\frac{(\ell+m)!}{(\ell-m)!}\left[-(1-x^2)^{\frac{1}{2}}\frac{\partial}{\partial x} + (m+1)\frac{x}{(1-x^2)^{\frac{1}{2}}}\right]\times \\ &\quad\times(1-x^2)^{-\frac{m+1}{2}}\frac{\partial^{\ell-m-1}}{\partial x^{\ell-m-1}}(x^2-1)^\ell\,. \end{aligned} \tag{5.127}$$

By means of

$$\begin{aligned} -(1-x^2)^{\frac{1}{2}}\frac{\partial}{\partial x}(1-x^2)^{-\frac{m+1}{2}}\frac{\partial^{\ell-m-1}}{\partial x^{\ell-m-1}}(x^2-1)^\ell \\ = \quad -(1-x^2)^{-\frac{m}{2}}\frac{\partial^{\ell-m}}{\partial x^{\ell-m}}(x^2-1)^\ell \\ -(m+1)\frac{x}{(1-x^2)^{\frac{1}{2}}}(1-x^2)^{-\frac{m+1}{2}}\frac{\partial^{\ell-m-1}}{\partial x^{\ell-m-1}} \end{aligned} \tag{5.128}$$

one can show that (5.127) reproduces (5.124) and, therefore, that expression (5.124) holds for m if it holds for $m+1$. Since (5.124) holds for $m=\ell$, it holds then for all m.
We want to test if the recursion (5.123) terminates for $m=-\ell$, i.e., if

$$\mathcal{L}_-\, Y_{\ell\ell}(\theta,\phi) \;=\; 0 \tag{5.129}$$

holds. In fact, expression (5.124), which is equivalent to recursive application of $\mathcal{L}_-$, yields a vanishing expression for $m=-\ell-1$ as long as ℓ is a non-negative integer. In that case $(1-x^2)^\ell$ is a polynomial of degree 2ℓ and, hence, the derivative $\partial^{2\ell+1}/\partial x^{2\ell+1}$ of this expression vanishes.

Constructing $Y_{\ell-\ell} \to Y_{\ell-\ell+1} \to \cdots Y_{\ell\ell}$

An alternative route to construct the eigenfunctions $Y_{\ell m}(\theta,\phi)$ determines first a normalized solution of

$$\mathcal{L}_-\, f(\theta,\phi) \;=\; 0 \tag{5.130}$$

$$\mathcal{L}_3\, f(\theta,\phi) \;=\; i\,\ell\, f(\theta,\phi)\,, \tag{5.131}$$

identifies $f(\theta,\phi) = Y_{\ell-\ell}(\theta,\phi)$ choosing the proper sign, and constructs then the eigenfunctions $Y_{\ell-\ell+1}$, $Y_{\ell-\ell+2}$, etc. by repeated application of the operator $\mathcal{L}_+$. Such construction reproduces the eigenfunctions $Y_{\ell m}(\theta,\phi)$ as given in (5.106, 5.124) and, therefore, appears not very interesting. However, from such construction emerges an important symmetry property of $Y_{\ell m}(\theta,\phi)$, namely,

$$Y^*_{\ell m}(\theta,\phi) \;=\; (-1)^m\, Y_{\ell-m}(\theta,\phi) \tag{5.132}$$

which reduces the number of spherical harmonics which need to be evaluated independently roughly by half ; therefore, we embark on this construction in order to prove (5.132).
We first determine $Y_{\ell-\ell}(\theta,\phi)$ using (5.106, 5.124). It holds, using $x=\cos\theta$,

$$P_{\ell-\ell}(x) \;=\; \frac{1}{2^\ell \ell!}\,\frac{1}{(2\ell)!}\,(1-x^2)^{\frac{\ell}{2}}\,\frac{\partial^{2\ell}}{\partial x^{2\ell}}\,(x^2-1)^\ell\,. \tag{5.133}$$

Since the term with the highest power of $(x^2-1)^\ell$ is $x^{2\ell}$, it holds

$$\frac{\partial^{2\ell}}{\partial x^{2\ell}}\,(x^2-1)^\ell \;=\; (2\ell)! \tag{5.134}$$

and, therefore,

$$P_{\ell-\ell}(x) \;=\; \frac{1}{2^\ell \ell!}\,(1-x^2)^{\frac{\ell}{2}}\,. \tag{5.135}$$

Due to (5.106, 5.124) one arrives at

$$Y_{\ell -\ell}(\theta,\phi) \;=\; \sqrt{\frac{2\ell+1}{4\pi}\,\frac{1}{(2\ell)!}}\,\frac{1}{2^\ell \ell!}\,\sin^\ell\theta\, e^{-i\ell\phi}\;. \tag{5.136}$$

We note in passing that this expression and the expression (5.120) for $Y_{\ell\ell}$ obey the postulated relationship (5.132).
Obviously, the expression (5.136) is normalized and has the proper sign, i.e., a sign consistent with the family of functions $Y_{\ell m}$ constructed above. One can readily verify that this expression provides a solution of (5.129). This follows from the identity

$$\left(\frac{\partial}{\partial\theta} \;-\; i\cot\theta\,\frac{\partial}{\partial\phi}\right)\sin^\ell\theta\, e^{-i\ell\phi} \;=\; 0\;. \tag{5.137}$$

One can use (5.136) to construct all other $Y_{\ell m}(\theta,\phi)$. According to (5.80) and (5.110) applies the recursion

$$e^{i\phi}\left(\frac{\partial}{\partial\theta} \;+\; i\cot\theta\,\frac{\partial}{\partial\phi}\right) Y_{\ell m}(\theta,\phi) \;=\; \sqrt{(\ell+m+1)(\ell-m)}\,Y_{\ell m+1}(\theta,\phi)\;. \tag{5.138}$$

Employing (5.106), one can conclude for the associated Legendre polynomials

$$\begin{aligned}\sqrt{\frac{(\ell-m)!}{(\ell+m)!}}\left(\frac{\partial}{\partial\theta} \;-\; m\cot\theta\right) P_{\ell m}(\cos\theta)& \\ \;=\; \sqrt{(\ell+m+1)(\ell-m)}\,\sqrt{\frac{(\ell-m-1)!}{(\ell+m+1)!}}\,P_{\ell m+1}(\cos\theta)&\end{aligned} \tag{5.139}$$

or

$$P_{\ell m+1}(\cos\theta) \;=\; \left(\frac{\partial}{\partial\theta} \;-\; m\cot\theta\right) P_{\ell m}(\cos\theta)\;. \tag{5.140}$$

Introducing again the variable $x = \cos\theta$ leads to the recursion equation [c.f. (5.122, 5.123)]

$$P_{\ell m+1}(x) \;=\; \left(-(1-x^2)^{\frac{1}{2}}\frac{\partial}{\partial x} \;-\; m\,\frac{x}{(1-x^2)^{\frac{1}{2}}}\right) P_{\ell m}(x)\;. \tag{5.141}$$

We want to demonstrate now that this recursion equation leads to the expression

$$P_{\ell m}(x) \;=\; \frac{(-1)^m}{2^\ell \ell!}\,(1-x^2)^{\frac{m}{2}}\,\frac{\partial^{\ell+m}}{\partial x^{\ell+m}}\,(x^2-1)^\ell \tag{5.142}$$

For this purpose we proceed by induction, following closely the proof of eq. (5.124). We first note that for $m = -\ell$ (5.142) yields

$$P_{\ell -\ell}(x) \;=\; \frac{(-1)^\ell}{2^\ell\ell!}\,(1-x^2)^{-\frac{\ell}{2}}\,(x^2-1)^\ell \;=\; \frac{1}{2^\ell\ell!}\,(1-x^2)^{\frac{\ell}{2}} \tag{5.143}$$

wich agrees with (5.135). We then assume that (5.142) holds for m. (5.141) reads then

$$P_{\ell m+1}(x) \; = \; \left(-(1-x^2)^{\frac{1}{2}} \frac{\partial}{\partial x} \; - \; m \, \frac{x}{(1-x^2)^{\frac{1}{2}}} \right) \frac{(-1)^m}{2^\ell \ell!} \, (1-x^2)^{\frac{m}{2}} \, \frac{\partial^{\ell+m}}{\partial x^{\ell+m}} \, (x^2-1)^\ell \; . \tag{5.144}$$

Replacing in (5.128) $m+1 \to -m$ or, equivalently, $m \to -m-1$ yields

$$\begin{aligned} & -(1-x^2)^{\frac{1}{2}} \frac{\partial}{\partial x} \, (1-x^2)^{\frac{m}{2}} \, \frac{\partial^{\ell+m}}{\partial x^{\ell+m}} \, (x^2-1)^\ell \\ = \; & -(1-x^2)^{\frac{m+1}{2}} \, \frac{\partial^{\ell+m+1}}{\partial x^{\ell+m+1}} \, (x^2-1)^\ell \\ & + m \, \frac{x}{(1-x^2)^{\frac{1}{2}}} \, (1-x^2)^{\frac{m}{2}} \, \frac{\partial^{\ell+m}}{\partial x^{\ell+m}} \; . \end{aligned} \tag{5.145}$$

Hence, (5.144) reads

$$P_{\ell m+1}(x) \; = \; \frac{(-1)^{m+1}}{2^\ell \ell!} \, (1-x^2)^{\frac{m+1}{2}} \, \frac{\partial^{\ell+m+1}}{\partial x^{\ell+m+1}} \, (x^2-1)^\ell \, , \tag{5.146}$$

i.e., (5.142) holds for $m+1$ if it holds for m. Since (5.142) holds for $m = -\ell$ we verified that it holds for all m.

The construction beginning with $Y_{\ell -\ell}$ and continuing with $Y_{\ell -\ell+1}$, $Y_{\ell -\ell+2}$, etc. yields the same eigenfunctions as the previous construction beginning with $Y_{\ell\ell}$ and stepping down the series of functions $Y_{\ell\ell-1}$, $Y_{\ell\ell-2}$, etc. Accordingly, also the associated Legendre polynomials determined this way, i.e., given by (5.142) and by (5.124), are identical. However, one notes that application of (5.124) yields

$$(-1)^m \frac{(\ell-m)!}{(\ell+m)!} \, P_{\ell -m}(x) \; = \; \frac{(-1)^m}{2^\ell \ell!} \, (1-x^2)^{\frac{m}{2}} \, \frac{\partial^{\ell+m}}{\partial x^{\ell+m}} \, (x^2-1)^\ell \, , \tag{5.147}$$

the r.h.s. of which agrees with the r.h.s. of (5.142). Hence, one can conclude

$$P_{\ell m}(x) \; = \; (-1)^m \frac{(\ell-m)!}{(\ell+m)!} \, P_{\ell -m}(x) \; . \tag{5.148}$$

According to (5.106) this implies for $Y_{\ell m}(\theta, \phi)$ the identity (5.132).

The Legendre Polynomials

The functions

$$P_\ell(x) \; = \; P_{\ell 0}(x) \tag{5.149}$$

are called *Legendre polynomials*. According to both (5.124) and (5.142) one can state

$$P_\ell(x) \; = \; \frac{1}{2^\ell \ell!} \, \frac{\partial^\ell}{\partial x^\ell} \, (x^2-1)^\ell \; . \tag{5.150}$$

The first few polynomials are

$$P_0(x) \;=\; 1\,, \qquad P_1(x) \;=\; x\,, \qquad P_2(x) \;=\; \tfrac{1}{2}(3x^2 \,-\, 1)$$
$$P_3(x) \;=\; \tfrac{1}{2}\,(5x^3 \,-\, 3x) \qquad . \tag{5.151}$$

Comparision of (5.150) and (5.142) allows one to express the associated Legendre polynomials in terms of Legendre polynomials

$$P_{\ell m}(x) \;=\; (-1)^m\,(1-x^2)^{\frac{m}{2}}\,\frac{\partial^m}{\partial x^m}\,P_\ell(x)\;,\; m \,\geq\, 0 \tag{5.152}$$

and, accordingly, the spherical harmonics for $m \,\geq\, 0$

$$Y_{\ell m}(\theta,\phi) \quad = \quad \sqrt{\frac{2\ell+1}{4\pi}\,\frac{(\ell-m)!}{(\ell+m)!}}\,(-1)^m\,\sin^m\theta\;\times \tag{5.153}$$
$$\times\;\left(\frac{\partial}{\partial\cos\theta}\right)^m\,P_\ell(\cos\theta)\,e^{im\phi}\;.$$

The spherical harmonics for $m \,<\, 0$ can be obtained using (5.132).
The Legendre polynomials arise in classical electrodynamics as the expansion coefficients of the electrostatic potential around a point charge $q/|\vec{r}_1 \,-\, \vec{r}_2|$ where q is the charge, $\vec{r}_1$ is the point where the potential is measured, and $\vec{r}_2$ denotes the location of the charge. In case $q = 1$ and $|\vec{r}_2| \,<\, |\vec{r}_1|$ holds the identity[2]

$$\frac{1}{|\vec{r}_1 \,-\, \vec{r}_2|} \;=\; \sum_{\ell=0}^{\infty}\,\frac{r_2^\ell}{r_1^{\ell+1}}\,P_\ell(\cos\gamma) \tag{5.154}$$

where γ is the angle between $\vec{r}_1$ and $\vec{r}_2$. Using $x \,=\, \cos\gamma$, $t \,=\, r_2/r_1$, and

$$|\vec{r}_1 \,-\, \vec{r}_2| \;=\; r_1\,\sqrt{1 \,-\, 2\,x\,t \,+\, t^2}\;, \tag{5.155}$$

(5.154) can be written

$$w(x,t) \;=\; \frac{1}{\sqrt{1 \,-\, 2xt \,+\, t^2}} \;=\; \sum_{\ell=0}^{\infty}\,P_\ell(x)\,t^\ell\;. \tag{5.156}$$

$w(x,t)$ is called a *generating function* of the Legendre polynomials[3].
The generating function allows one to derive useful properties of the Legendre polynomials. For example, in case of $x \,=\, 1$ holds

$$w(1,t) \;=\; \frac{1}{\sqrt{1 \,-\, 2t \,+\, t^2}} \;=\; \frac{1}{1 \,-\, t} \;=\; \sum_{\ell=0}^{\infty}\,t^\ell\;. \tag{5.157}$$

Comparision with (5.156) yields

$$P_\ell(1) \;=\; 1\,, \qquad \ell \,=\, 0,\,1,\,2,\,\ldots \tag{5.158}$$

[2] see, e.g., "Classical Electrodynamics, 2nd Ed." by J.D. Jackson (John Wiley, New York, 1975), pp. 92

[3] To prove this property see, for example, pp. 45 of "Special Functions and their Applications" by N.N. Lebedev (Prentice-Hall, Englewood Cliffs, N.J., 1965) which is an excellent compendium on special functions employed in physics.

For $w(x,t)$ holds

$$\frac{\partial \ln w(x,t)}{\partial t} \;=\; \frac{x \,-\, t}{1 \,-\, 2xt \,+\, t^2}\,. \tag{5.159}$$

Using $\partial \ln w/\partial t \,=\, (1/w)\,\partial w/\partial t$ and multiplying (5.159) by $w(1 \,-\, 2xt \,+\, t^2)$ leads to the differential equation obeyed by $w(x,t)$

$$(1 \,-\, 2xt \,+\, t^2)\,\frac{\partial w}{\partial t} \;+\; (t \,-\, x)\,w \;=\; 0\,. \tag{5.160}$$

Employing (5.156) and $(\partial/\partial t)\sum_\ell P_\ell(x)t^\ell \,=\, \sum_\ell \ell P_\ell(x)t^{\ell-1}$ this differential equation is equivalent to

$$(1 \,-\, 2xt \,+\, t^2)\sum_{\ell=0}^{\infty} \ell P_\ell(x)\,t^{\ell-1} \;+\; (t \,-\, x)\sum_{\ell=0}^{\infty} P_\ell(x)\,t^\ell \;=\; 0 \tag{5.161}$$

Collecting coefficients with equal powers t^ℓ yields

$$\begin{aligned} 0 \;=\; & P_1(x) \;-\; x\,P_0(x) \\ & +\sum_{\ell=1}^{\infty}\left[(\ell+1)\,P_{\ell+1}(x) \;-\; (2\ell+1)x\,P_\ell(x) \;+\; \ell\,P_{\ell-1}(x)\right]t^\ell \end{aligned} \tag{5.162}$$

The coefficients for all powers t^ℓ must vanish individually. Accordingly, holds

$$P_1(x) \;=\; x\,P_0(x) \tag{5.163}$$

$$(\ell \,+\, 1)\,P_{\ell+1}(x) \;=\; (2\ell \,+\, 1)\,x\,P_\ell(x) \;-\; \ell\,P_{\ell-1}(x) \tag{5.164}$$

which, using $P_0(x) \,=\, 1$ [c.f. (5.151)] allows one to determine $P_\ell(x)$ for $\ell \,=\, 1,\, 2,\, \ldots$.

Inversion Symmetry of $Y_{\ell m}(\theta,\phi)$ Under inversion at the origin vectors $\vec{r}$ are replaced by $-\vec{r}$. If the spherical coordinates of $\vec{r}$ are (r,θ,ϕ), then the coordinates of $-\vec{r}$ are $(r,\pi-\theta,\pi+\phi)$. Accordingly, under inversion $Y_{\ell m}(\theta,\phi)$ goes over to $Y_{\ell m}(\pi-\theta,\pi+\phi)$. Due to $\cos(\pi-\theta) \,=\, -\cos\theta$ the replacement $\vec{r} \,\to\, -\vec{r}$ alters $P_{\ell m}(x)$ into $P_{\ell m}(-x)$. Inspection of (5.142) allows one to conclude

$$P_{\ell m}(-x) \;=\; (-1)^{\ell+m}\,P_{\ell m}(x) \tag{5.165}$$

since $\partial^n/\partial(-x)^n \,=\, (-1)^n\partial^n/\partial x^n$. Noting $\exp[im(\pi+\phi)] \,=\, (-1)^m\exp(im\phi)$ we determine

$$Y_{\ell m}(\pi-\theta,\pi+\phi) \;=\; (-1)^\ell\,Y_{\ell m}(\theta,\phi)\,. \tag{5.166}$$

Properties of $Y_{\ell m}(\theta,\phi)$

We want to summarize the properties of the spherical harmonics $Y_{\ell m}(\theta,\phi)$ derived above.

1. The spherical harmonics are eigenfunctions of the angular momentum operators

$$\mathcal{J}^2\,Y_{\ell m}(\theta,\phi) \;=\; \hbar^2\,\ell(\ell+1)\,Y_{\ell m}(\theta,\phi) \tag{5.167}$$

$$\mathcal{J}_3\,Y_{\ell m}(\theta,\phi) \;=\; \hbar\,m\,Y_{\ell m}(\theta,\phi)\,. \tag{5.168}$$

2. The spherical harmonics form an orthonormal basis of the space $\mathbb{C}_\infty(S_2)$ of normalizable, uniquely defined functions over the unit sphere S_2 which are infinitely often differentiable

$$\int_0^\pi \sin\theta d\theta \int_0^{2\pi} d\phi Y^*_{\ell' m'}(\theta,\phi)\, Y_{\ell m}(\theta,\phi) \;=\; \delta_{\ell'\ell}\delta_{m'm}\;. \tag{5.169}$$

3. The spherical harmonics form, in fact, a *complete* basis of $\mathbb{C}_\infty(S_2)$, i.e., for any $f(\theta,\phi) \in \mathbb{C}_\infty(S_2)$ holds

$$f(\theta,\phi) \;=\; \sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell} C_{\ell m}\, Y_{\ell m}(\theta,\phi) \tag{5.170}$$

$$C_{\ell m} \;=\; \int_0^\pi \sin\theta d\theta \int_0^{2\pi} d\phi\, Y^*_{\ell m}(\theta,\phi) f(\theta,\phi)\;. \tag{5.171}$$

4. The spherical harmonics obey the recursion relationships

$$\mathcal{J}_+\, Y_{\ell m}(\theta,\phi) \;=\; \hbar\sqrt{(\ell+m+1)(\ell-m)}\, Y_{\ell m+1}(\theta,\phi) \tag{5.172}$$

$$\mathcal{J}_-\, Y_{\ell m+1}(\theta,\phi) \;=\; \hbar\sqrt{(\ell+m+1)(\ell-m)}\, Y_{\ell m}(\theta,\phi) \tag{5.173}$$

where $\mathcal{J}_\pm \;=\; \mathcal{J}_1 \pm i\mathcal{J}_2$.

5. The spherical harmonics are given by the formula

$$Y_{\ell m}(\theta,\phi) \;=\; \sqrt{\frac{2\ell+1}{4\pi}\,\frac{(\ell-m)!}{(\ell+m)!}}\,\frac{(-1)^m}{2^\ell \ell!}\,\sin^m\theta\;\times \tag{5.174}$$
$$\times\left(\frac{\partial}{\partial\cos\theta}\right)^m P_\ell(\cos\theta)\, e^{im\phi}\;,\quad m\;\geq\;0$$

where $(\ell = 1, 2, \ldots)$

$$P_0(x) \;=\; 1\;,\qquad P_1(x) \;=\; x\;, \tag{5.175}$$

$$P_{\ell+1}(x) \;=\; \tfrac{1}{\ell+1}\left[\,(2\ell\;+\;1)\,x\,P_\ell(x)\;-\;\ell\,P_{\ell-1}(x)\,\right] \tag{5.176}$$

are the *Legendre polynomials*. The spherical harmonics for $m < 0$ are given by

$$Y_{\ell -m}(\theta,\phi) \;=\; (-1)^m\, Y^*_{\ell m}(\theta,\phi)\;. \tag{5.177}$$

Note that the spherical harmonics are real, except for the factor $\exp(im\phi)$.

6. The spherical harmonics $Y_{\ell 0}$ are

$$Y_{\ell 0}(\theta,\phi) \;=\; \sqrt{\frac{2\ell+1}{4\pi}}\, P_\ell(\cos\theta)\;. \tag{5.178}$$

7. For the Legendre polynomials holds the orthogonality property

$$\int_{-1}^{+1} dx\; P_\ell(x)\, P(\ell'(x) \;=\; \frac{2}{2\ell\;+\;1}\,\delta_{\ell\ell'}\;. \tag{5.179}$$

8. The spherical harmonics for $\theta = 0$ are

$$Y_{\ell m}(\theta = 0, \phi) \;=\; \delta_{m0} \sqrt{\frac{2\ell + 1}{4\pi}} \; . \tag{5.180}$$

9. The spherical harmonics obey the inversion symmetry

$$Y_{\ell m}(\theta, \phi) \;=\; (-1)^{\ell} \, Y_{\ell m}(\pi - \theta, \pi + \phi) \; . \tag{5.181}$$

10. The spherical harmonics for $\ell = 0$, 1, 2 are given by

$$Y_{00} \;=\; \frac{1}{\sqrt{4\pi}} \tag{5.182}$$

$$Y_{11} \;=\; -\sqrt{\frac{3}{8\pi}} \, \sin\theta \, e^{i\phi} \tag{5.183}$$

$$Y_{10} \;=\; \sqrt{\frac{3}{4\pi}} \, \cos\theta \tag{5.184}$$

$$Y_{22} \;=\; \frac{1}{4} \sqrt{\frac{15}{2\pi}} \, \sin^2\theta \, e^{2i\phi} \tag{5.185}$$

$$Y_{21} \;=\; -\sqrt{\frac{15}{8\pi}} \, \sin\theta \cos\theta \, e^{i\phi} \tag{5.186}$$

$$Y_{20} \;=\; \sqrt{\frac{5}{4\pi}} \left(\frac{3}{2} \, \cos^2\theta \, - \, \frac{1}{2} \right) \tag{5.187}$$

together with (5.177).

11. For the Laplacian holds

$$\nabla^2 \, h(r) \, Y_{\ell m}(\theta, \phi) \;=\; \left[\frac{1}{r} \frac{\partial^2}{\partial r^2} \, r \; - \; \frac{\ell(\ell + 1)}{r^2} \right] \, h(r) \, Y_{\ell m}(\theta, \phi) \; . \tag{5.188}$$

Exercise 5.4.1: Derive expressions (5.86), (5.87).
Exercise 5.4.2: Derive the orthogonality property for the Legendre polynomials (5.179) using the generating function $w(x, t)$ stated in (5.156). For this purpose start from the identity

$$\int_{-1}^{+1} dx \; w(x, t)^2 \;=\; \sum_{\ell, \ell' = 0}^{\infty} t^{\ell + \ell'} \; \int_{-1}^{+1} dx P_{\ell}(x) \, P_{\ell'}(x) \; , \tag{5.189}$$

evaluate the integral on the l.h.s., expand the result in powers of t and equate the resulting powers to those arising on the r.h.s. of (5.189).
Exercise 5.4.3: Construct all spherical harmonics $Y_{3m}(\theta, \phi)$.

5.5 Irreducible Representations

We will consider now the effect of the rotational transformations introduced in (5.42), $\exp\left(-\frac{i}{\hbar}\,\vec{\vartheta}\cdot\vec{\mathcal{J}}\right)$, on functions $f(\theta,\phi)\ \in\ \mathbb{C}_\infty(\mathbb{S}_{\not\in})$. We denote the image of a function $f(\theta,\phi)$ under such rotational transformation by $\tilde{f}(\theta,\phi)$, i.e.,

$$\tilde{f}(\theta,\phi)\ =\ \exp\left(-\frac{i}{\hbar}\,\vec{\vartheta}\cdot\vec{\mathcal{J}}\right)f(\theta,\phi)\ . \tag{5.190}$$

Since the spherical harmonics $Y_{\ell m}(\theta,\phi)$ provide a complete, orthonormal basis for the function space $\mathbb{C}_\infty(\mathbb{S}_{\not\in})$ one can expand

$$\tilde{f}(\theta,\phi)\ \ =\ \ \sum_{\ell,m} c_{\ell m}\ Y_{\ell m}(\theta,\phi) \tag{5.191}$$

$$c_{\ell m}\ \ =\ \ \int_{S_2} d\Omega'\,Y^*_{\ell m}(\theta',\phi')\exp\left(-\frac{i}{\hbar}\,\vec{\vartheta}\cdot\vec{\mathcal{J}}\right)f(\theta',\phi') \tag{5.192}$$

One can also represent $\mathbb{C}_\infty(\mathbb{S}_{\not\in})$

$$\mathbb{C}_\infty(\mathbb{S}_{\not\in})\ =\ [\,\{\mathbb{Y}_{\ell\gtrdot},\ \ell\ =\ \not\prec,\not\perp,\ldots\,\infty,\ \gtrdot\ =\ -\ell,-\ell+\not\perp,\ \ldots,\ell\}\,] \tag{5.193}$$

where $[\,\{\quad\}\,]$ denotes closure of a set by taking all possible linear combinations of the elements of the set. The transformations $\exp\left(-\frac{i}{\hbar}\,\vec{\vartheta}\cdot\vec{\mathcal{J}}\right)$, therefore, are characterized completely if one specifies the transformation of any $Y_{\ell m}(\theta,\phi)$

$$\tilde{Y}_{\ell m}(\theta,\phi)\ \ =\ \ \sum_{\ell',m'}\Big[\mathcal{D}(\vec{\vartheta})\Big]_{\ell' m';\,\ell m}\ Y_{\ell m}(\theta,\phi) \tag{5.194}$$

$$\Big[\mathcal{D}(\vec{\vartheta})\Big]_{\ell' m';\,\ell m}\ \ =\ \ \int_{S_2} d\Omega\,Y^*_{\ell' m'}(\theta,\phi)\,\exp\!\left(-\frac{i}{\hbar}\,\vec{\vartheta}\cdot\vec{\mathcal{J}}\right)Y_{\ell m}(\theta,\phi)\ , \tag{5.195}$$

i.e., if one specifies the functional form of the coefficients $[\mathcal{D}(\vartheta)]_{\ell' m';\,\ell m}$. These coefficients can be considered the elements of an infinite-dimensional matrix which provides the representation of $\exp\left(-\frac{i}{\hbar}\,\vec{\vartheta}\cdot\vec{\mathcal{J}}\right)$ in the basis $\{Y_{\ell m},\ \ell\ =\ 0,1,\ldots\,\infty,\ m\ =\ -\ell,-\ell+1,\ \ldots,\ell\}$.
We like to argue that the matrix representing the rotational transformations of the function space $\mathbb{C}_\infty(\mathbb{S}_{\not\in})$, with elements given by (5.195), assumes a particularly simple form, called the irreducible representation. For this purpose we consider the subspaces of $\mathbb{C}_\infty(\mathbb{S}_{\not\in})$

$$\mathbb{X}_\ell\ =\ [\,\{\mathbb{Y}_{\ell\gtrdot},\ \gtrdot\ =\ -\ell,-\ell+\not\perp,\ \ldots,\ell\}\,]\quad,\ \ell\ =\not\prec,\not\perp,\not\vDash,\ldots \tag{5.196}$$

Comparision with (5.193) shows

$$\mathbb{C}_\infty(\mathbb{S}_{\not\in})\ =\ \bigcup_{\ell=\not\prec}^{\infty}\mathbb{X}_\ell\ . \tag{5.197}$$

The subspaces $\mathbb{X}_\ell$ have the important property that (i) they are invariant under rotations $\exp\left(-\frac{i}{\hbar}\,\vec{\vartheta}\cdot\vec{\mathcal{J}}\right)$, i.e., $\exp\left(-\frac{i}{\hbar}\,\vec{\vartheta}\cdot\vec{\mathcal{J}}\right)(\mathbb{X}_\ell)\ =\ \mathbb{X}_\ell$, and (ii) they form the lowest dimensional sets $\mathbb{X}_\ell$ obeying $\mathbb{C}_\infty(\mathbb{S}_{\not\in})\ =\ \bigcup_\ell\mathbb{X}_\ell$ which have this invariance property.

The invariance of the subspaces $\mathbb{X}_\ell$ under rotations follows from a Taylor expansion of the exponential operator $\exp\left(-\frac{i}{\hbar}\,\vec{\vartheta}\cdot\vec{\mathcal{J}}\right)$. One proceeds by expanding first $\vec{\vartheta}\cdot\vec{\mathcal{J}}$ in terms of $\mathcal{J}_+$, $\mathcal{J}_-$, and $\mathcal{J}_3$

$$\vartheta\cdot\vec{\mathcal{J}} \;=\; \frac{1}{2}\vartheta_-\mathcal{J}_+ \;+\; \frac{1}{2}\vartheta_+\mathcal{J}_- \;+\; \vartheta_3\mathcal{J}_3 \tag{5.198}$$

where $\vartheta_\pm \;=\; \vartheta_1 \mp \vartheta_2$. The exponential operator $\exp\left(-\frac{i}{\hbar}\,\vec{\vartheta}\cdot\vec{\mathcal{J}}\right)$ can be expanded in powers $(\frac{1}{2}\vartheta_-\mathcal{J}_+ \;+\; \frac{1}{2}\vartheta_+\mathcal{J}_- \;+\; \vartheta_3\mathcal{J}_3)^n$. One employs then the commutation properties

$$[\mathcal{J}_+,\,\mathcal{J}_-] \;=\; 2\hbar\mathcal{J}_3\;, \quad [\mathcal{J}_3,\,\mathcal{J}_\pm] \;=\; \pm\hbar\mathcal{J}_\pm \tag{5.199}$$

which follow readily from (5.56, 5.65). Application of (5.199) allows one to express

$$\exp\left(-\frac{i}{\hbar}\,\vec{\vartheta}\cdot\vec{\mathcal{J}}\right) \;=\; \sum_{n_1,n_2,n_3} c_{n_1,n_2,n_3}\,\mathcal{J}_+^{n_1}\mathcal{J}_-^{n_2}\mathcal{J}_3^{n_3}\;. \tag{5.200}$$

Accordingly, one needs to consider the action of $\mathcal{J}_+^{n_1}\mathcal{J}_-^{n_2}\mathcal{J}_3^{n_3}$ on the subspaces $\mathbb{X}_\ell$. Since the $Y_{\ell m}$ are eigenfunctions of $\mathcal{J}_3$ the action of $\mathcal{J}_3^{n_3}$ reproduces the basis functions of $\mathbb{X}_\ell$, i.e., leaves $\mathbb{X}_\ell$ invariant. The action of $\mathcal{J}_-$ on $Y_{\ell m}$ produces $Y_{\ell m-1}$ or, in case $m \;=\; -\ell$, produces 0, i.e., $\mathcal{J}_-^{n_2}$ also leaves $\mathbb{X}_\ell$ invariant. Similarly, one can conclude that $\mathcal{J}_+^{n_1}$ leaves $\mathbb{X}_\ell$ invariant. Since any additive term in (5.200) leaves $\mathbb{X}_\ell$ invariant, the transformation $\exp\left(-\frac{i}{\hbar}\,\vec{\vartheta}\cdot\vec{\mathcal{J}}\right)$ leaves $\mathbb{X}_\ell$ invariant as well. One can conclude then that in the expansion (5.194) only spherical harmonics with $\ell' \;=\; \ell$ contribute or, equivalently, one concludes for the matrix elements (5.195)

$$[\mathcal{D}(\vartheta)]_{\ell' m';\,\ell m} \;=\; \delta_{\ell'\ell}\int_{S_2} d\Omega\, Y^*_{\ell m'}(\theta,\phi)\exp\left(-\frac{i}{\hbar}\,\vec{\vartheta}\cdot\vec{\mathcal{J}}\right)Y_{\ell m}(\theta,\phi)\;. \tag{5.201}$$

From expression (5.200) of the rotational transformations one can also conclude that the $\mathbb{X}_\ell$ are the smallest invariant subspaces of $\mathbb{C}_\infty(\mathbb{S}_{\not\in})$.
If one orders the basis according to the partitoning (5.197) of $\mathbb{C}_\infty(\mathbb{S}_{\not\in})$ the matrix representation of $\mathcal{D}(\vartheta)$ assumes a block-diagonal form, with blocks of dimension $1, 3, 5, \ldots$, i.e.,

$$\mathcal{D}(\vartheta) \;=\; \begin{pmatrix} \boxed{1\times 1} & & & & \\ & \boxed{3\times 3} & & & \\ & & \ddots & & \\ & & & \boxed{(2\ell+1)\times(2\ell+1)} & \\ & & & & \ddots \end{pmatrix} \tag{5.202}$$

This representation of $\mathcal{D}(\vartheta)$ is the one wich has blocks of the lowest dimensions possible. The representation is referred to as the *irreducible representation*, other representations are termed *reducible representations*.

Exercise 5.5.1: Representations of the group SO(2)
SO(2) is the set of all 2×2 matrices R which are orthogonal, i.e. for which holds $R^T R = R\,R^T = \mathbb{1}$, and for which $\det R = 1$.
(a) Show that SO(2) with the group operation $\circ$ defined as matrix multiplication, is a group.
(b) Prove that the elements of SO(2) can be completely characterized through a single parameter.
(c) Show that the map $R(\varphi)$

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = R(\varphi) \begin{pmatrix} x \\ y \end{pmatrix}$$

defines a representation of SO(2).
(d) Show that the similarity transformation $T(\varphi)$ defined in the space of non-singular, real 2×2–matrices O through

$$O' = T(\varphi)O = R(\varphi)OR^{-1}(\varphi)$$

with $R(\varphi)$ as in (c) is also a representation of SO(2).
(e) In the space $\mathbb{C}^\infty(\mathbb{K})$ of infinitely often differentiable, and **periodic** functions $f(\alpha)$, i.e. $f(\alpha + 2\pi) = f(\alpha)$, the map $\rho(\varphi)$ defined through

$$f(\alpha) \xrightarrow{\rho} g(\varphi) = f(\alpha - \varphi) = \rho(\varphi) f(\alpha)$$

also defines a representation of SO(2). Determine the generator of $\rho(\varphi)$ in analogy to (5.22, 5.49).
(f) The transformation $\rho(\varphi)$ as defined in (e) leaves the following subspace of functions considered in (e)

$$X_m = \{\, f(\alpha) = A\, e^{-im\alpha}, A \in \mathbb{C}, \gtrless \in \mathbb{N}\}$$

invariant. Give the corresponding expression of $\rho(\varphi)$.

5.6 Wigner Rotation Matrices

We will now take an important first step to determine the functional form of the matrix elements of $\mathcal{D}(\vartheta)$. This step reconsiders the parametrization of rotations by the vector $\vec{\vartheta}$ assumed sofar. The three components $\vartheta_k, k = 1, 2, 3$ certainly allow one to describe any rotation around the origin. However, this parametrization, though seemingly natural, does not provide the simplest mathematical description of rotations. A more suitable parametrization had been suggested by Euler: every rotation $\exp\left(-\frac{i}{\hbar}\,\vec{\vartheta}\cdot\vec{\mathcal{J}}\right)$ can be represented also uniquely by three consecutive rotations:

(i) a first rotation around the original x_3–axis by an angle α,

(ii) a second rotation around the new x_2'–axis by an angle β,

(iii) a third rotation around the new x_3''–axis by an angle γ.

The angles α, β, γ will be referred to as Euler angles. The axis x_2' is defined in the coordinate frame which is related to the original frame by rotation (i), the axis x_3'' is defined in the coordinate frame which is related to the original frame by the consecutive rotations (i) and (ii).
The Euler rotation replaces $\exp\left(-\frac{i}{\hbar}\,\vec{\vartheta}\cdot\vec{\mathcal{J}}\right)$ by

$$e^{-\frac{i}{\hbar}\gamma J_3''}e^{-\frac{i}{\hbar}\beta J_2'}\,e^{-\frac{i}{\hbar}\alpha J_3} \quad . \tag{5.203}$$

For any $\vartheta \in \mathbb{R}^{\not\in}$ one can find Euler angles α, β, $\gamma \in \mathbb{R}$ such that

$$\exp\left(-\frac{i}{\hbar}\,\vec{\vartheta}\cdot\vec{\mathcal{J}}\right) \;=\; e^{-\frac{i}{\hbar}\gamma J_3''}e^{-\frac{i}{\hbar}\beta J_2'}\,e^{-\frac{i}{\hbar}\alpha J_3} \tag{5.204}$$

is satisfied. Accordingly, one can replace (5.194, 5.195) by

$$\tilde{Y}_{\ell m}(\theta,\phi) \;=\; \sum_{\ell',m'}\left[\mathcal{D}(\alpha,\beta,\gamma)\right]_{\ell' m';\,\ell m}\; Y_{\ell m}(\theta,\phi) \tag{5.205}$$

$$\left[\mathcal{D}(\alpha,\beta,\gamma)\right]_{\ell' m';\,\ell m} \;=\; \int_{S_2} d\Omega\, Y^*_{\ell' m'}(\theta,\phi)\; e^{-\frac{i}{\hbar}\gamma J_3''}e^{-\frac{i}{\hbar}\beta J_2'}\,e^{-\frac{i}{\hbar}\alpha J_3}Y_{\ell m}(\theta,\phi)\;, \tag{5.206}$$

The expression (5.203) has the disadvantage that it employs rotations defined with respect to three different frames of reference. We will demonstrate that (5.203) can be expressed, however, in terms of rotations defined with respect to the original frame. For this purpose we notice that J_2' can be expressed through the similarity transformation

$$J_2' \;=\; e^{-\frac{i}{\hbar}\alpha J_3}J_2\,e^{\frac{i}{\hbar}\alpha J_3} \tag{5.207}$$

which replaces J_2' by the inverse transformation (i), i.e., the transformation from the rotated frame to the original frame, followed by J_2 in the original frame, followed by transformation (i), i.e., the transformation from the original frame to the rotated frame. Obviously, the l.h.s. and the r.h.s. of (5.207) are equivalent. For any similarity transformation involving operators A and S holds

$$e^{S\,A\,S^{-1}} \;=\; S\,e^{A}\,S^{-1}\;. \tag{5.208}$$

Accordingly, we can write

$$e^{-\frac{i}{\hbar}\beta J_2'} \;=\; e^{-\frac{i}{\hbar}\alpha J_3}\,e^{-\frac{i}{\hbar}\beta J_2}\,e^{\frac{i}{\hbar}\alpha J_3} \tag{5.209}$$

The first two rotations in (5.203), i.e., (i) and (ii), can then be written

$$e^{-\frac{i}{\hbar}\beta J_2'}\,e^{-\frac{i}{\hbar}\alpha J_3} \;=\; e^{-\frac{i}{\hbar}\alpha J_3}\,e^{-\frac{i}{\hbar}\beta J_2} \tag{5.210}$$

The third rotation in (5.203), in analogy to (5.209), is

$$e^{-\frac{i}{\hbar}\gamma J_3''} \;=\; e^{-\frac{i}{\hbar}\beta J_2'}\,e^{-\frac{i}{\hbar}\alpha J_3}\,e^{-\frac{i}{\hbar}\gamma J_3}\,e^{\frac{i}{\hbar}\alpha J_3}\,e^{\frac{i}{\hbar}\beta J_2'} \tag{5.211}$$

Using (5.209) in this expression one obtains

$$\begin{aligned} e^{-\frac{i}{\hbar}\gamma J_3''} \;&=\; e^{-\frac{i}{\hbar}\alpha J_3}\,e^{-\frac{i}{\hbar}\beta J_2}\,e^{\frac{i}{\hbar}\alpha J_3}\,e^{-\frac{i}{\hbar}\alpha J_3}\,e^{-\frac{i}{\hbar}\gamma J_3}\,e^{\frac{i}{\hbar}\alpha J_3}\,e^{-\frac{i}{\hbar}\alpha J_3}\,e^{\frac{i}{\hbar}\beta J_2}\,e^{\frac{i}{\hbar}\alpha J_3} \\ &=\; e^{-\frac{i}{\hbar}\gamma J_3''} \;=\; e^{-\frac{i}{\hbar}\alpha J_3}\,e^{-\frac{i}{\hbar}\beta J_2}\,e^{-\frac{i}{\hbar}\gamma J_3}\,e^{\frac{i}{\hbar}\beta J_2}\,e^{\frac{i}{\hbar}\alpha J_3} \end{aligned} \tag{5.212}$$

Multiplication from the left with (5.210) yields the simple result

$$e^{-\frac{i}{\hbar}\gamma J_3''}\, e^{-\frac{i}{\hbar}\beta J_2'}\, e^{-\frac{i}{\hbar}\alpha J_3} \;=\; e^{-\frac{i}{\hbar}\alpha J_3}\; e^{-\frac{i}{\hbar}\beta J_2}\; e^{-\frac{i}{\hbar}\gamma J_3}\;, \tag{5.213}$$

i.e., to redefine the three rotations in (5.203) with respect to the original (unprimed) frame one simply needs to reverse the order of the rotations. This allows one to express the rotational trasnformation (5.205, 5.206) by

$$\tilde{Y}_{\ell m}(\theta,\phi) \;=\; \sum_{\ell',m'} [\mathcal{D}(\alpha,\beta,\gamma)]_{\ell' m';\,\ell m}\; Y_{\ell m}(\theta,\phi) \tag{5.214}$$

$$[\mathcal{D}(\alpha,\beta,\gamma)]_{\ell' m';\,\ell m} \;=\; \int_{S_2} d\Omega\, Y^*_{\ell' m'}(\theta,\phi)\; e^{-\frac{i}{\hbar}\alpha J_3}\; e^{-\frac{i}{\hbar}\beta J_2}\; e^{-\frac{i}{\hbar}\gamma J_3} Y_{\ell m}(\theta,\phi)\;. \tag{5.215}$$

The evaluation of the matrix elements (5.215) benefits from the choice of Euler angles for the parametrization of rotational transformations. The eigenvalue property (5.64) yields

$$\int_{S_2} d\Omega\, f(\theta,\phi)\; e^{-\frac{i}{\hbar}\gamma J_3} Y_{\ell m}(\theta,\phi) \;=\; \int_{S_2} d\Omega\, f(\theta,\phi)\; Y_{\ell m}(\theta,\phi)\; e^{-im\gamma}\;. \tag{5.216}$$

Using, in addition, the self-adjointness of the operator $\mathcal{J}_3$ one can state

$$\int_{S_2} d\Omega\, Y^*_{\ell m}(\theta,\phi)\; e^{-\frac{i}{\hbar}\alpha J_3} f(\theta,\phi) \;=\; e^{-im'\alpha} \int_{S_2} d\Omega\, Y^*_{\ell m'}(\theta,\phi)\; f(\theta,\phi)\;. \tag{5.217}$$

Accordingly, one can write

$$[\mathcal{D}(\alpha,\beta,\gamma)]_{\ell m;\,\ell' m'} \;=\; e^{-i\alpha m'} \left(\int_{S_2} d\Omega\, Y^*_{\ell m'}(\theta,\phi)\, e^{-\frac{i}{\hbar}\beta \mathcal{J}_2}\; Y_{\ell m}(\theta,\phi) \right) e^{-i\gamma m}\, \delta_{\ell\ell'} \tag{5.218}$$

Defining the so-called *Wigner rotation matrix*

$$d^{(\ell)}_{m m'}(\beta) \;=\; \int_{S_2} d\Omega\, Y^*_{\ell m'}(\theta,\phi)\, e^{-\frac{i}{\hbar}\beta \mathcal{J}_2}\; Y_{\ell m}(\theta,\phi) \tag{5.219}$$

one can express the rotation matrices

$$[\mathcal{D}(\alpha,\beta,\gamma)]_{\ell m;\,\ell' m'} \;=\; e^{-i\alpha m'}\; d^{(\ell)}_{m' m}(\beta)\; e^{-i\gamma m}\;\; \delta_{\ell\ell'}\;. \tag{5.220}$$

We will derive below [see Eqs. (5.309, 5.310)] an explicit expression for the Wigner rotation matrix (5.219).

5.7 Spin $\frac{1}{2}$ and the group SU(2)

The spin describes a basic and fascinating property of matter. Best known is the spin of the electron, but many other elementary components of matter are endowed with spin-like properties. Examples are the other five members of the lepton family to which the electron belongs, the electron e and the electron neutrino ν_e of the first generation, the muon μ and its neutrino ν_μ of the second

generation, the tau τ and its neutrino ν_τ of the third generation and their antiparticles carry a spin-$\frac{1}{2}$. So do the three generations of six quarks, two of which (in certain linear combinations) make up the vector mesons which carry spin 1, and three of which make up the baryons which carry spin-$\frac{1}{2}$ and spin-$\frac{3}{2}$. There are also the mediators, the gluon, the photon, the two $W^{\pm}$ and the Z^o, particles which mediate the strong and the electro-weak interactions, and which carry spin 1.
The particles mentioned, e.g. the quarks, carry other spin-like properties which together, however, have properties beyond those of single spins. In any case, there is nothing more elementary to matter than the spin property. The presence of this property permeates matter also at larger scales than those of the elementary particles mentioned, leaving its imprint on the properties of nuclei, atoms and molecules; in fact, the spin of the electron is likely the most important property in Chemistry. We may finally mention that the spin is at the heart of many properties of condensed matter systems, like superconductivity and magnetism. It appears to be rather impossible for a Physicist not to be enamored with the spin property. We will find that the spin in its transformation behaviour is closely related to angular momentum states, a relationship, which might be the reason why consideration of rotational symmetry is so often fruitful in the study of matter.
We will consider first only the so-called spin–$\frac{1}{2}$, generalizing then further below. Spin-$\frac{1}{2}$ systems can be related to two states which we denote by χ_+ and χ_-. Such systems can also assume any linear combination $c_+\chi_+ + c_-\chi_-$, $c_\pm \in \mathbb{C}$, as long as $|c_+|^2 + |c_-|^2 = 1$. If we identify the states χ_+ and χ_- with the basis of a Hilbert space, in which we define the scalar product between any state $|1\rangle = a_+\chi_+ + a_-\chi_-$ and $|2\rangle = b_+\chi_+ + b_-\chi_-$ as $\langle 1|2\rangle = a_+^* b_+ + a_-^* b_-$, then allowed symmetry transformations of spin states are described by 2×2-matrices U with complex-valued matrix elements U_{jk}. Conservation of the scalar product under symmetry transformations requires the property $U\,U^\dagger = U^\dagger U = \mathbb{1}$ where $U^\dagger$ denotes the adjoint matrix with elements $\left[U^\dagger\right]_{jk} = U^*_{kj}$. We will specify for the transformations considered $\det U = 1$. This specification implies that we consider transformations save for overall factors $e^{i\phi}$ since such factors are known not to affect any observable properties.
The transformation matrices are then elements of the set

$$SU(2) \;=\; \{\,\text{complex}\,2\times 2\ \text{matrices}\,U;\; U\,U^\dagger = U^\dagger U = \mathbb{1},\; \det U = 1\,\} \tag{5.221}$$

One can show readily that this set forms a group with the groups binary operation being matrix multiplication.
How can the elements of SU(2) be parametrized. As complex 2×2 matrices one needs, in principle, eight real numbers to specify the matrix elements, four real and four imaginary parts of $U_{jk}, j, k = 1, 2$. Because of the unitarity condition $U^\dagger U = \mathbb{1}$ which are really four equations in terms of real quantities, one for each matrix element of $U^\dagger U$, and because of $\det U = 1$, there are together five conditions in terms of real quantities to be met by the matrix elements and, hence, the degrees of freedom of the matrices U are three real quantities. The important feature is that all $U \in$ SU(2) can be parametrized by an exponential operator

$$U \;=\; \exp\Big(-i\,\vec{\vartheta}\cdot\vec{S}\Big) \tag{5.222}$$

where the vector $\vec{S}$ has three components, each component $S_k, k = 1, 2, 3$ representing a 2×2 matrix. One can show that the unitarity condition requires these matrices to be hermitian, i.e. $[S_k]_{mn} = ([S_k]_{nm})^*$, and the condition $\det U = 1$ requires the S_k to have vanishing trace. There

exist three such linear independent matrices, the simplest choice being

$$S_1 \;=\; \frac{1}{2}\,\sigma_1, \quad S_2 \;=\; \frac{1}{2}\,\sigma_2, \quad S_3 \;=\; \frac{1}{2}\,\sigma_3 \tag{5.223}$$

where σ_k, $k \,=\, 1, 2, 3$ are the Pauli matrices

$$\sigma_1 \;=\; \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 \;=\; \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 \;=\; \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \;. \tag{5.224}$$

Algebraic Properties of the Pauli Matrices

The Pauli matrices (5.224) provide a basis in terms of which all traceless, hermitian 2×2–matrices A can be expandend. Any such A can be expressed in terms of three real parameters x, y, z

$$A \;=\; \begin{pmatrix} z & x \,-\, iy \\ x \,+\, iy & -z \end{pmatrix}, \quad x,\, y,\, z, \in \mathbb{R} \;. \tag{5.225}$$

In fact, it holds,

$$A \;=\; x\,\sigma_1 \;+\; y\,\sigma_2 \;+\; z\,\sigma_3 \;. \tag{5.226}$$

The Pauli matrices satisfy very special commutation and anti-commutation relationships. One can readily verify the *commutation property*

$$\sigma_j\,\sigma_k \;-\; \sigma_k\,\sigma_j \;=\; [\sigma_j,\, \sigma_k]_- \;=\; 2i\epsilon_{jk\ell}\,\sigma_\ell \tag{5.227}$$

which is essentially identical to the Lie algebra (5.31) of the group SO(3). We will show below that a 2-1–homomorphic mapping exists between SU(2) and SO(3) which establishes the close relationship between the two groups.
The Pauli matrices obey the following *anti-commutation properties*

$$\sigma_j\,\sigma_k \;+\; \sigma_k\,\sigma_j \;=\; [\sigma_j,\, \sigma_k]_+ \;=\; 2\delta_{jk}\mathbb{1} \;, \tag{5.228}$$

i.e.

$$\left(\sigma^j\right)^2 \;=\; \mathbb{1} \;; \quad \sigma^j\sigma^k \;=\; -\,\sigma^k\sigma^j \quad \text{for } j \,\neq\, k \tag{5.229}$$

which can also be readily verified. According to this property the Pauli matrices generate a 3-dimensional *Clifford algebra* C_3. Clifford algebras play an important role in the mathematical structure of physics, e.g. they are associated with the important fermion property of matter. At this point we will state a useful property, namely,

$$(\vec{\sigma}\cdot\vec{a})\,(\vec{\sigma}\cdot\vec{b}) \;=\; \vec{a}\cdot\vec{b}\,\mathbb{1} \;+\; i\vec{\sigma}\,\cdot\,(\vec{a}\times\vec{b}) \tag{5.230}$$

where $\vec{a}$, $\vec{b}$ are vectors commuting with $\vec{\sigma}$, but their comonents must not necessarily commute with each other, i.e., it might hold $a_jb_k \,-\, b_ka_j \,\neq\, 0$. The proof of this relation rests on the commutation relationship and the anti-commutation relationship (5.227, 5.229) and avoids commuting the components a_j and b_k. In fact, one obtains

$$\begin{gathered}(\vec{\sigma}\cdot\vec{a})\,(\vec{\sigma}\cdot\vec{b}) \;=\; \textstyle\sum_{j,k=1}^{3}\sigma^j\sigma^k\,a_jb_k \;= \\ \textstyle\sum_{j=1}^{3}(\sigma^j)^2\,a_jb_j \;+\; \sum_{\substack{j,k=1\\ j>k}}^{3}\sigma^j\sigma^k\,a_jb_k \;+\; \sum_{\substack{j,k=1\\ j<k}}^{3}\sigma^j\sigma^k\,a_jb_k \;.\end{gathered} \tag{5.231}$$

Using $(\sigma^j)^2 = 1$ the first term on the r.h.s. yields $\vec{a}\cdot\vec{b}$. The two remaining terms yield using $\sigma^j\sigma^k = -\sigma^k\sigma^j$ for $j \neq k$ and altering 'dummy' summation indices

$$\sum_{\substack{j,k=1\\ j>k}}^{3} \sigma^j\sigma^k\,(a_jb_k - a_kb_j) =$$
$$\tfrac{1}{2}\sum_{\substack{j,k=1\\ j>k}}^{3} \sigma^j\sigma^k\,(a_jb_k - a_kb_j) - \tfrac{1}{2}\sum_{\substack{j,k=1\\ j>k}}^{3} \sigma^k\sigma^j\,(a_jb_k - a_kb_j) =$$
$$\tfrac{1}{2}\sum_{\substack{j,k=1\\ j>k}}^{3} (\sigma^j\sigma^k - \sigma^k\sigma^j)\,(a_jb_k - a_kb_j)\,. \tag{5.232}$$

The commutation property (5.227) leads to

$$i\sum_{\substack{j,k=1\\ j>k}}^{3} \epsilon_{jk\ell}\sigma_l\,(a_jb_k - a_kb_j) = i\sum_{\ell=1}^{3}\sigma_\ell\left(\vec{a}\times\vec{b}\right)_\ell\,. \tag{5.233}$$

This result together with (5.231) proves (5.230).
In the special case $\vec{a} = \vec{b} \in \mathbb{R}^{\not 4}$ holds

$$(\vec{\sigma}\cdot\vec{a})^2 = \vec{a}^{\,2}\mathbb{1}\,. \tag{5.234}$$

5.8 Generators and Rotation Matrices of SU(2)

Sofar there is nothing which relates the transformations U, i.e. (5.222), to rotations in space. This relationship emerges, however, through the algebra obeyed by the operators S_k

$$[S_k,\,S_\ell] = i\,\epsilon_{k\ell m}\,S_m \tag{5.235}$$

which is identical to that of the generators of SO(3). The algebra of the generators of a transformation is such a basic property that the 'accident' that the generators of spin transformations behave in this respect like the generators of 3-dimensional rotations makes spins appear in their physical behaviour like rotations. Well almost like it, since there is a slight difference: when you interprete the 3-component $\vec{\vartheta}$ in (5.222) as a rotation vector and you rotate once, let say around the x_2–axis by 360^o, spin changes sign; only a 720^o rotation leaves the spin invariant. We like to derive this result now by evaluating the transformations of SU(2) explicitly. For this purpose we choose to replace (5.222) by the Euler form

$$\exp(-i\gamma S_3)\exp(-i\beta S_2)\exp(-i\alpha S_3) \quad . \tag{5.236}$$

The matrix elements of this 2×2 operator can be labelled by the basis states χ_+ and χ_-, however, we like to draw in this respect also on a close analogy to angular momentum states which developes if one considers the operators $S^2 = S_1^2 + S_2^2 + S_3^2$ and S_3. Noticing the idempotence of the S_k's

$$S_k^2 = \frac{1}{4}\begin{pmatrix} 1 & 0\\ 0 & 1\end{pmatrix} \tag{5.237}$$

one obtains $S^2 = \frac{3}{4}\mathbb{1} = \frac{1}{2}(\frac{1}{2}+1)\mathbb{1}$ and, hence,

$$S^2\,\chi_\pm = \frac{1}{2}(\frac{1}{2}+1)\,\chi_\pm \quad ; \quad S_3\,\chi_\pm = \pm\frac{1}{2}\,\chi_\pm \quad . \tag{5.238}$$

This result implies that the states $\chi_\pm$ behave in relation to the generators of SU(2) like an angular momentum state $|\frac{1}{2} \pm \frac{1}{2}\rangle$ in relation to the generators of SO(3). We may, therefore, use the label $|\frac{1}{2} \pm \frac{1}{2}\rangle$ for the states $\chi_\pm$.
We obtain with this notation for the transformations (5.236)

$$\begin{aligned} &\langle \tfrac{1}{2} m \mid \exp(-i\gamma S_3) \exp(-i\beta S_2) \exp(-i\alpha S_3) \mid \tfrac{1}{2} m' \rangle \\ &\quad = \; e^{-i\gamma m} \langle \tfrac{1}{2} m \mid \exp(-i\beta S_2) \mid \tfrac{1}{2} m' \rangle \, e^{-i\alpha m'} \\ &\quad = \; e^{-i\gamma m} \, d^{(\frac{1}{2})}_{mm'}(\beta) e^{-i\alpha m'} \end{aligned} \tag{5.239}$$

where we made use of the notation (5.219).
In order to determine the Wigner matrix in (5.219,5.239) one expands the exponential operator $\exp(-i\beta S_2)$. For this purpose one needs to determine the powers of $-i\beta S_2$. The idempotence property of S_k yields particularly simple expressions for these powers, namely,

$$(-\beta S_2)^{2n} \;=\; (-1)^n \left(\frac{\beta}{2}\right)^{2n} \mathbb{1} \tag{5.240}$$

$$(-\beta S_2)^{2n+1} \;=\; (-1)^n \left(\frac{\beta}{2}\right)^{2n+1} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad . \tag{5.241}$$

Taylor expansion of the exponential operator yields then

$$\exp(-i\beta S_2) \;=\; \left[\sum_{n=0}^{\infty} (-1)^n \left(\frac{\beta}{2}\right)^{2n}\right] \mathbb{1} \;+\; \left[\sum_{n=0}^{\infty} (-1)^n \left(\frac{\beta}{2}\right)^{2n+1}\right] \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{5.242}$$

The expressions in brackets [...] can be identified with the Taylor expansion of the cos- and sin-functions and one obtains for the rotation matrices

$$\left(d^{(\frac{1}{2})}_{mm'}(\beta) \right) \;=\; \begin{pmatrix} \cos\frac{\beta}{2} & -\sin\frac{\beta}{2} \\ \sin\frac{\beta}{2} & \cos\frac{\beta}{2} \end{pmatrix}. \tag{5.243}$$

We note the property of this rotation matrix

$$\left(d^{(\frac{1}{2})}_{mm'}(2\pi) \right) \;=\; -\,\mathbb{1} \quad ; \tag{5.244}$$

i.e. rotation by 360^o changes the sign of the spin state.
The complete matrix elements (5.239) in the notation (5.220) are

$$[\mathcal{D}(\alpha,\beta,\gamma)]_{\ell m;\,\ell' m'} \;=\; \begin{pmatrix} \cos\frac{\beta}{2}\, e^{-i(\frac{\alpha}{2}+\frac{\gamma}{2}} & -\sin\frac{\beta}{2}\, e^{i(\frac{\alpha}{2}-\frac{\gamma}{2}} \\ \sin\frac{\beta}{2}\, e^{i(-\frac{\alpha}{2}+\frac{\gamma}{2}} & \cos\frac{\beta}{2}\, e^{i(\frac{\alpha}{2}+\frac{\gamma}{2}} \end{pmatrix} . \tag{5.245}$$

5.9 Constructing Spin States with Larger Quantum Numbers Through Spinor Operators

In this section we like to demonstrate, following Jourdan and Schwinger, that states $|\ell m\rangle$ for higher quantum numbers $\ell = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots$, $m = -\ell, -\ell+1, \ldots \ell$ can be constructed formally from spin states $\chi_\pm$ if one considers the two properties $\chi_\pm$ to be carried by two kinds of **bosons**, i.e. identical particles any number of which can exist in the same state χ_+ and χ_-. One cannot consider the entities carrying the spin–$\frac{1}{2}$ to be particles in the ordinary sense for it can be shown that spin–$\frac{1}{2}$ particles have fermion character, i.e. no two such particles can exist in the same state.

Definition of Spinor Creation and Annihilation Operators

We present the states $\chi_\pm$ through creation operators $b_\pm^\dagger$ which when applied to a formal vacuum state $|\Psi_0\rangle$ generate $\chi_\pm$, i.e.

$$b_+^\dagger|\Psi_0\rangle \;=\; \chi_+ \quad ; \quad b_-^\dagger|\Psi_0\rangle \;=\; \chi_- \quad . \tag{5.246}$$

The corresponding adjoint operators are denoted by b_+ and b_-. The boson character of the operators $b^\dagger = (b_+^\dagger, b_-^\dagger)^T$ and $b = (b_+, b_-)^T$ is expressed through the commutation relationships $(\zeta, \zeta' = +, -)$

$$\left[b_\zeta, b_{\zeta'}\right] = \left[b_\zeta^\dagger, b_{\zeta'}^\dagger\right] = 0 \tag{5.247}$$

$$\left[b_\zeta, b_{\zeta'}^\dagger\right] = \delta_{\zeta\zeta'} \; . \tag{5.248}$$

For the vacuum state $|\Psi_0\rangle$ holds

$$b_\pm\,|\Psi_0\rangle \;=\; 0 \tag{5.249}$$

In the following will refer to $b^\dagger = (b_+^\dagger, b_-^\dagger)^T$ and $b = (b_+, b_-)^T$ as spinor creation and annihilation operators. These operators are associated with a given spatial reference system. We consider, therefore, also operators of the type

$$x\,b^\dagger \;=\; x_+ b_+^\dagger \;+\; x_- b_-^\dagger \quad ; \, x^* b \;=\; x_+^* b_+ \;+\; x_-^* b_- \tag{5.250}$$

which for $x^*x \;=\; x_+^* x_+ \;+\; x_-^* x_- \;=\; 1$ represent spinor operators in an arbitrary reference system. For example, using (5.243) the creation operators in a coordinate system rotated by an angle β around the y–axis are

$$\left(b'_+\right)^\dagger \;=\; \cos\frac{\beta}{2}\, b_+^\dagger \;+\; \sin\frac{\beta}{2}\, b_-^\dagger \quad ; \quad \left(b'_-\right)^\dagger \;=\; -\sin\frac{\beta}{2}\, b_+^\dagger \;+\; \cos\frac{\beta}{2}\, b_-^\dagger \; . \tag{5.251}$$

One can show using this property and (5.247,5.248) that $\left(b'_\zeta\right)^\dagger$ and b'_ζ obey the commutation relationships

$$\left[b'_\zeta, b'_{\zeta'}\right] = \left[b'^\dagger_\zeta, b'^\dagger_{\zeta'}\right] = 0 \tag{5.252}$$

$$\left[b'_\zeta, b'^\dagger_{\zeta'}\right] = \delta_{\zeta\zeta'} \; . \tag{5.253}$$

The States $|\Psi(j,m)\rangle$

The operators $b_\pm^\dagger$ allow one to construct a set of states which represent $j+m$–fold and $j-m$–fold χ_+ and χ_- states as follows

$$|\Psi\,(j,m)\rangle = \frac{\left(b_+^\dagger\right)^{j+m}}{\sqrt{(j+m)!}} \frac{\left(b_-^\dagger\right)^{j-m}}{\sqrt{(j-m)!}} |\Psi_0\rangle \; . \tag{5.254}$$

We will show below that these states are orthonormal and form spin states with higher quantum numbers, i.e. $j \;=\; 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots \; ; \; m \;=\; -j, -j+1, \ldots j$.

5.10 Algebraic Properties of Spinor Operators

We want to establish first a few important and useful algebraic properties which result from the commutation relationships (5.247, 5.248) and from (5.249).

The Spinor Derivative Operator

A most important operation is the action of the annihilation operators b_+ and b_- on operators which can be expressed as polynomials, possibly infinite power series, in $b_+^\dagger$ and $b_-^\dagger$. An example is the monomial $\left(b_+^\dagger\right)^{j+m}\left(b_-^\dagger\right)^{j-m}$ which generates the states (5.254). We will derive the result, well-known for the quantum mechanical harmonic oscillator, that the annihilation operators play a role similar to the differential operator in calculus.
For this purpose we consider the following polynomial of creation operators

$$f(b_+^\dagger) \;=\; \sum_{n=0}^{N} c_n \left(b_+^\dagger\right)^n \;. \tag{5.255}$$

In order to determine how this operator is modified by multiplication from the left by b_+ we note

$$b_+ f(b_+^\dagger)\,|\text{state}\rangle \;=\; \left[\,b_+,\, f(b_+^\dagger)\right]\,|\text{state}\rangle \;+\; f(b_+^\dagger)\,b_+\,|\text{state}\rangle\;. \tag{5.256}$$

In the special case $|\text{state}\rangle = |\Psi_0\rangle$ this reads using (5.249)

$$b_+ f(b_+^\dagger)\,|\Psi_0\rangle \;=\; \left[\,b_+,\, f(b_+^\dagger)\right]\,|\Psi_0\rangle\;. \tag{5.257}$$

Equations (5.256) and (5.257) motivate us to determine the commutator $\left[\,b_+,\, f(b_+^\dagger)\right]$ which for (5.255) is

$$\left[b_+, f\left(b^\dagger\right)\right] = \sum_n c_n \left[b_+, \left(b_+^\dagger\right)^n\right]. \tag{5.258}$$

Obviously, we need to evaluate $\left[b_+, \left(b_+^\dagger\right)^n\right]$. We will show that for this commutator holds

$$\left[b_+, \left(b_+^\dagger\right)^n\right] \;=\; n\left(b_+^\dagger\right)^{n-1}\;. \tag{5.259}$$

The property is obviously true for $n = 0$. If it is true for n then using (5.248) we obtain

$$\begin{aligned}
\left[b_+, \left(b_+^\dagger\right)^{n+1}\right] &= \left[b_+, \left(b_+^\dagger\right)^n b_+^\dagger\right] \\
&= \left[b_+, \left(b_+^\dagger\right)^n\right] b_+^\dagger + \left(b_+^\dagger\right)^n \left[b_+, b_+^\dagger\right] \\
&= n\left(b_+^\dagger\right)^{n-1} b_+^\dagger + \left(b_+^\dagger\right)^n \\
&= (n+1)\left(b_+^\dagger\right)^n\;,
\end{aligned} \tag{5.260}$$

i.e. the property holds also for $n+1$. By induction we can conclude that (5.259) holds for all $n \in \mathbb{N}$. One can finally conclude for polynomials (5.255)

$$\left[b_+, f(b_+^\dagger)\right] \;=\; f'(b_+^\dagger) \tag{5.261}$$

where $f'(x) = \frac{df}{dx}$. Similarly, one can prove

$$\left[b_-, f(b_-^\dagger)\right] \;=\; f'(b_-^\dagger) \tag{5.262}$$

Because of the commutation relationships (5.247) the more general property for polynomials $f(b_+^\dagger, b_-^\dagger)$ in $b_+^\dagger$ and $b_-^\dagger$ holds ($\zeta = +, -$)

$$\left[b_\zeta, f(b_+^\dagger, b_-^\dagger)\right] \;=\; \frac{\partial}{\partial b_\zeta^\dagger} f(b_+^\dagger, b_-^\dagger)\,. \tag{5.263}$$

According to (5.257) we can conclude in particular

$$b_\zeta\, f(b_+^\dagger, b_-^\dagger)\,|\Psi_0\rangle \;=\; \frac{\partial}{\partial b_\zeta^\dagger}\, f(b_+^\dagger, b_-^\dagger)\,|\Psi_0\rangle\,. \tag{5.264}$$

This demonstrates the equivalence of the spinor operators b_ζ and the derivative operation.
Equation (5.263) corresponds to the product rule of calculus $\partial_k f(\vec{x})g(\vec{x}) = (\partial_k f(\vec{x}))g(\vec{x}) + f(\vec{x})(\partial_k g(\vec{x}))$ which can be written $(\,\partial_k f(\vec{x}) - f(\vec{x})\partial_k\,)\,g(\vec{x}) = [\partial_k, f(\vec{x})]\;g(\vec{x}) = (\partial_k f(\vec{x}))g(\vec{x})$ or

$$[\partial_k, f(\vec{x})] \;=\; \partial_k f(\vec{x})\,. \tag{5.265}$$

Generating Function of the States $|\Psi(j,m)\rangle$

We want to prove now the property

$$\exp\left(xb^\dagger\right)\,|\Psi_0\rangle \;=\; \sum_{j=0,\frac{1}{2},1,\ldots}^{\infty}\;\sum_{m=-j}^{j} \phi_{jm}\,(x)\,\Psi\,(j,m) \tag{5.266}$$

where $xb^\dagger$ has been defined in (5.250) and where $\phi_{jm}\,(x)$ represents the function of the two variables x_+ and x_- closely related to $|\Psi(j,m)\rangle$

$$\phi_{jm}\,(x) = \frac{x_+^{j+m} x_-^{j-m}}{\sqrt{(j+m)!\,(j-m)!}}\,. \tag{5.267}$$

$\exp\left(xb^\dagger\right)\,|\Psi_0\rangle$ is called a *generating function* of $|\Psi(j,m)\rangle$.
In order to derive (5.266) we compare the terms

$$\phi_{jm}\,(x)\,\Psi\,(j,m) = \frac{\left(x_+ b_+^\dagger\right)^{j+m}}{(j+m)!}\,\frac{\left(x_- b_-^\dagger\right)^{j-m}}{(j-m)!}|\Psi_0\rangle. \tag{5.268}$$

with the s–th term in the binomial expansion of $(a+b)^n$

$$\frac{n!}{s!\,(n-s)!}a^{n-s}b^s\ . \tag{5.269}$$

Defining

$$\begin{aligned} j+m &= N-s \\ j-m &= s\ . \end{aligned} \tag{5.270}$$

one obtains

$$\begin{aligned} \phi_{jm}(x)\,\Psi(j,m) &= \frac{\left(x_+ b_+^\dagger\right)^{N-s}}{(N-s)!}\frac{\left(x_- b_-^\dagger\right)^s}{s!}|\Psi_0\rangle \\ &= \frac{1}{N!}\left[\frac{N!}{(N-s)!s!}\right]\left(x_+ b_+^\dagger\right)^{N-s}\left(x_- b_-^\dagger\right)^s|\Psi_0\rangle. \end{aligned} \tag{5.271}$$

Summation over s from $s=0$ to $s=N$ yields

$$\begin{aligned} &\frac{1}{N!}\sum_{s=0}^{N}\left[\frac{N!}{(N-s)!s!}\right]\left(x_+ b_+^\dagger\right)^{N-s}\left(x_- b_-^\dagger\right)^s|\Psi_0\rangle \\ =\ &\frac{1}{N!}\left(x_+ b_+^\dagger + x_- b_-^\dagger\right)^N|\Psi_0\rangle \\ =\ &\frac{1}{(2j)!}\left(x_+ b_+^\dagger + x_- b_-^\dagger\right)^{2j}|\Psi_0\rangle \end{aligned} \tag{5.272}$$

The summation over s can be written in terms of j and m using (5.270)

$$\sum_{s=0}^{N} \to \sum_{j-m=0}^{2j} \to \sum_{m=-j}^{j}\ . \tag{5.273}$$

Change of the summation indices allows one to conclude

$$\begin{aligned} &\sum_{m=-j}^{j}\phi_{jm}(x)\,\Psi(j,m) \\ =\ &\frac{1}{(2j)!}\left(x_+ b_+^\dagger + x_- b_-^\dagger\right)^{2j}|\Psi_0\rangle \\ =\ &\frac{1}{(2j)!}\left(x b^\dagger\right)^{2j}|\Psi_0\rangle. \end{aligned} \tag{5.274}$$

Summing this expression over $2j = 0,1,2,\ldots$, i.e. chosing the summation index $j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots$, leads to

$$\begin{aligned} \sum_{j=0,\frac{1}{2},1,\ldots}\ \sum_{m=-j}^{j}\phi_{jm}(x)\,|\Psi(j,m)\rangle &= \sum_{j=0,\frac{1}{2},1,\ldots}\frac{1}{2j!}\left(x b^\dagger\right)^{2j}|\Psi_0\rangle \\ &= \sum_{u=0,1,\ldots}^{\infty}\frac{1}{u!}\left(x b^\dagger\right)^u|\Psi_0\rangle \\ &= exp\left(x b^\dagger\right)|\Psi_0\rangle \end{aligned} \tag{5.275}$$

which concludes our derivation. The generating function allows one to derive various properties of the states $|\Psi(j,m)\rangle$ and will be used for this purpose below.

Orthonormality of the States $|\Psi(j,m)\rangle$

We want to show now that the states (5.254) are orthonormal, i.e. that

$$\langle \Psi(j,m) | \Psi\left(j',m'\right)\rangle \;=\; \delta_{jj'}\delta_{mm'} \tag{5.276}$$

holds. For this purpose we consider the inner product

$$\begin{aligned} & \langle e^{\left(xb^\dagger\right)}\Psi_0 | e^{\left(yb^\dagger\right)}\Psi_0\rangle \\ = & \sum_{j,m,j',m'} \phi_{jm}(x^*)\,\phi_{j'm'}(y)\,\langle \Psi(j,m)|\Psi\left(j',m'\right)\rangle\,. \end{aligned} \tag{5.277}$$

To evaluate this expression we first notice

$$\left[e^{\left(xb^\dagger\right)}\right]^\dagger \;=\; e^{(x^*b)}\,. \tag{5.278}$$

which allows us to replace the l.h.s. of (5.277) by $\langle\Psi_0|e^{(x^*b)}e^{\left(yb^\dagger\right)}\Psi_0\rangle$. In order to evaluate the operator $e^{(x^*b)}e^{(yb^\dagger)}$ we notice that the derivative property (5.264) implies

$$b_\zeta e^{\left(yb^\dagger\right)}|\Psi_0\rangle = y_\zeta e^{\left(yb^\dagger\right)}|\Psi_0\rangle\,. \tag{5.279}$$

One can generalize this to

$$(b_\zeta)^s\, e^{\left(yb^\dagger\right)}|\Psi_0\rangle = y_\zeta^s e^{\left(yb^\dagger\right)}|\Psi_0\rangle\,. \tag{5.280}$$

The commutation properties $[b_+,b_-]=0$, $[b_+^\dagger,b_-^\dagger]=0$ and $[b_-,b_+^\dagger]=0$ allow one to state that for any polynomial $f(b_+,b_-)$ holds

$$f(b_+,b_-)e^{\left(yb^\dagger\right)}|\Psi_0\rangle = f(y_+,y_-)e^{\left(yb^\dagger\right)}|\Psi_0\rangle \tag{5.281}$$

and, hence, one can write

$$\begin{aligned} & \langle e^{\left(xb^\dagger\right)}\Psi_0|e^{\left(yb^\dagger\right)}\Psi_0\rangle \\ = \;& \langle\Psi_0|e^{(x^*y)}e^{\left(yb^\dagger\right)}\Psi_0\rangle \\ = \;& e^{(x^*y)}\,\langle\Psi_0|e^{\left(yb^\dagger\right)}\Psi_0\rangle \end{aligned} \tag{5.282}$$

where we defined $x^*y = x_+^*y_+ + x_-^*y_-$. According to (5.249) for any non-vanishing integer s holds

$$\langle\Psi_0|b^s\Psi_0\rangle \;=\; 0 \tag{5.283}$$

and, therefore, we can conclude

$$\langle e^{\left(xb^\dagger\right)}\Psi_0|e^{\left(yb^\dagger\right)}\Psi_0\rangle \;=\; e^{(x^*y)} \tag{5.284}$$

and comparing (5.282) and (5.277)

$$\sum_{j,m,j',m'} \phi_{jm}\left(x^*\right)\phi_{j'm'}\left(y\right)\left\langle \Psi\left(j,m\right)\left|\Psi\left(j',m'\right)\right.\right\rangle \;=\; e^{(x^*y)}\;. \tag{5.285}$$

Following steps similar to those in Eqs. (5.268–5.275) one can show

$$e^{(x^*y)} \;=\; \sum_{j,m} \phi_{jm}\left(x^*\right)\phi_{jm}\left(y\right). \tag{5.286}$$

from which follows immediately the orthonormality property (5.276).

New Representation of Spin $0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$ States

We want to demonstrate now that the states $|\Psi\left(j,m\right)\rangle$ as given in (5.254) are eigenstates of operators J^2 and J_3 with eigenvalues $j(j+1)$ and m where J^2 and J_3 are representations of the spin operators $S^2 = S_1^2 + S_2^2 + S_3^2$ and S_3 defined in 5.223, 5.224). One defines corresponding operators J_k through

$$J_k = \frac{1}{2}\sum_{\zeta,\zeta'} b_\zeta^\dagger < \zeta\left|\sigma_k\right|\zeta' > b_{\zeta'}\;, \tag{5.287}$$

where σ_k, $k = 1,2,3$ denote the Pauli spin matrices (5.224). In our present notation the matrix elements $< \zeta\left|\sigma_k\right|\zeta' >$ for $\zeta = \pm$ correspond to the matrix elements $< \zeta\left|\sigma_k\right|\zeta' >$ for $\sigma = \pm\frac{1}{2}$. The operators are explicitly, using $< +\left|\sigma_1\right|+ > = < -\left|\sigma_1\right|- > = 0$, $< +\left|\sigma_1\right|- > = < -\left|\sigma_1\right|+ > = 1$, etc.,

$$\begin{aligned} J_1 &= \frac{1}{2}\left(b_+^\dagger b_- + b_-^\dagger b_+\right) \\ J_2 &= \frac{1}{2i}\left(b_+^\dagger b_- - b_-^\dagger b_+\right) \\ J_3 &= \frac{1}{2}\left(b_+^\dagger b_+ - b_-^\dagger b_-\right)\;. \end{aligned} \tag{5.288}$$

The three operators obey the Lie algebra of SU(2)

$$\left[J_i, J_j\right] = i\epsilon_{ijk}J_k\;. \tag{5.289}$$

This property is derived as follows

$$\begin{aligned} \left[J_i, J_j\right] &= \frac{1}{4}\sum_{n,n',m,m'} < n\left|\sigma_i\right|m >< n'\left|\sigma_j\right|m' > \left[b_n^\dagger b_m, b_{n'}^\dagger b_{m'}\right] \\ &= \frac{1}{4}\sum_{n,n',m,m'} < n\left|\sigma_i\right|m >< n'\left|\sigma_j\right|m' > \{b_n^\dagger\left[b_m, b_{n'}^\dagger\right] b_{m'} \\ &\quad + b_{n'}^\dagger\left[b_n^\dagger, b_{m'}\right] b_m\} \\ &= \frac{1}{4}\sum_{n,n',m,m'} < n\left|\sigma_i\right|m >< n'\left|\sigma_j\right|m' > \{b_n^\dagger b_{m'}\delta_{mn'} - b_{n'}^\dagger b_m\delta_{m'n}\} \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{4}\sum_{n,m,m'} < n\,|\sigma_i|\,m >< m\,|\sigma_j|\,m' > b_n^\dagger b_{m'} \\
&\quad -\frac{1}{4}\sum_{n',m,m'} < n'\,|\sigma_j|\,m' >< m'\,|\sigma_i|\,m > b_{n'}^\dagger b_m \\
&= \frac{1}{4}\sum_{n,m} < n\,|[\sigma_i,\sigma_j]|\,m > b_n^\dagger b_m \\
&= \frac{1}{4}\sum_{n,m,k} < n\,|2i\epsilon_{ijk}\sigma_k|\,m > b_n^\dagger b_m \\
&= i\epsilon_{ijk}\frac{1}{2}\sum_{n,m} < n\,|\sigma_k|\,m > b_n^\dagger b_m \\
&= i\epsilon_{ijk}J_k\ . \qquad (5.290)
\end{aligned}
$$

$|\Psi\,(j,m)\rangle$ as Eigenstates of J^2 and J_3

We wish to show now that the states $|\Psi\,(j,m)\rangle$ are eigenstates of J^2 and J_3. To this end we note

$$J^2 \;=\; J_1^2 \;+\; J_2^2 \;+\; J_3^2 \;=\; \frac{1}{2}\,J_+J_- \;+\; \frac{1}{2}\,J_-J_+ \;+\; J_3^2 \quad . \qquad (5.291)$$

Here we have defined

$$
\begin{aligned}
J_+ \;&=\; J_1 \;+\; iJ_2 \;&=\; b_+^\dagger b_- \\
J_- \;&=\; J_1 \;-\; iJ_2 \;&=\; b_-^\dagger b_+\ . \qquad (5.292)
\end{aligned}
$$

The commutation relationships (5.289) yield

$$[J_-,\,J_+] \;=\; [J_1 \,-\, iJ_2,\,J_1 \,+\, iJ_2] \;=\; -i\,[J_2,\,J_1] \;+\; i\,[J_1,\,J_2] \;=\; -2\,J_3 \qquad (5.293)$$

from which we can conclude the property

$$\frac{1}{2}\,J_-J_+ \;=\; \frac{1}{2}\,J_+J_- \;-\; J_3 \quad . \qquad (5.294)$$

Hence,

$$J^2 \;=\; J_+J_- \;+\; J_3^2 \;-\; J_3 \;=\; J_3\,(J_3 \,-\, 1) \;+\; J_+J_- \quad . \qquad (5.295)$$

The last result together with (5.292), (5.288) yields first the expression

$$J^2 \;=\; \frac{1}{4}\,(\,b^\dagger{}_+\,b_+ \;-\; b^\dagger{}_-b_-\,)\,(\,b^\dagger{}_+\,b_+ \;-\; b^\dagger{}_-b_- \;-\; 2\,) \;+\; b^\dagger{}_+b_-b^\dagger{}_-b_+ \quad . \qquad (5.296)$$

which results in

$$
\begin{aligned}
J^2 \;=\; \tfrac{1}{4}\,(\,&b^\dagger{}_+\,b_+\,b^\dagger{}_+\,b_+ \;+\; b^\dagger{}_+\,b_+\,b^\dagger{}_-b_- \;-\; 2\,b^\dagger{}_+\,b_+ \;-\; b^\dagger{}_-b_-\,b^\dagger{}_+\,b_+ \\
&+\; b^\dagger{}_-b_-b^\dagger{}_-b_- \;+\; 2\,b^\dagger{}_-b_- \;+\; 4\,b^\dagger{}_+b_-b^\dagger{}_-b_+\,) \qquad (5.297)
\end{aligned}
$$

The last term on the r.h.s. can be written

$$\begin{aligned} b^\dagger{}_+ b_- b^\dagger{}_- b_+ &= b^\dagger{}_+ b_+ \; + \; b^\dagger{}_+ \, b_+ b^\dagger{}_- b_- \\ &= b^\dagger{}_+ b_+ \; + \; \frac{1}{2} b^\dagger{}_+ \, b_+ b^\dagger{}_- b_- \; + \; \frac{1}{2} b^\dagger{}_- b_- b^\dagger{}_+ \, b_+ \; . \end{aligned} \tag{5.298}$$

One can then state

$$\begin{aligned} J^2 \;\; = \;\; & \frac{1}{4} \Big[(b^\dagger{}_+ b_+)^2 \; + \; b^\dagger{}_+ b_+ b^\dagger{}_- b_- \; + \; b^\dagger{}_- b_- b^\dagger{}_+ \, b_+ \\ & + (b^\dagger{}_- b_-)^2 \; + \; 2 \, b^\dagger{}_+ b_+ \; + \; 2 \, b^\dagger{}_- b_- \Big] \; . \end{aligned} \tag{5.299}$$

Defining the operator

$$\hat{k} \; = \; \frac{1}{2} \, (\, b^\dagger{}_+ b_+ \; + \; b^\dagger{}_- b_- \,) \tag{5.300}$$

one can write finally

$$J^2 \; = \; \hat{k} \, (\, \hat{k} \; + \; 1 \,) \quad . \tag{5.301}$$

Obviously, the states (5.254) are eigenstates of $b^\dagger{}_+ b_+$ and $b^\dagger{}_- b_-$ with eigenvalues $j+m$ and $j-m$, respectively, and eigenstates of $\hat{k}$ with eigenvalues j. One can then conclude

$$J^2 \, |\Psi \, (j,m) \rangle \;\; = \;\; j \, (\, j \; + \; 1 \,) \, |\Psi \, (j,m) \rangle \tag{5.302}$$

$$J_3 \, |\Psi \, (j,m) \rangle \;\; = \;\; m \, |\Psi \, (j,m) \rangle \; . \tag{5.303}$$

One can furthermore derive readily

$$J_+ |\Psi \, (j,m) \rangle \;\; = \;\; \sqrt{(j+m+1) \, (j-m)} \, |\Psi \, (j,m+1) \rangle \tag{5.304}$$

$$J_- \, |\Psi \, (j,m) \rangle \;\; = \;\; \sqrt{(j+m) \, (j-m+1)} |\Psi \, (j,m-1) \rangle \; . \tag{5.305}$$

Exercise 5.10.1: The system investigated in this section is formally identical to a 2-dimensional isotropic harmonic oscillator governed by the Hamiltonian

$$H \; = \; \hbar \omega (b_1^+ b_1 \; + \; b_2^+ b_2 \; + \; 1) \, ; \; [b_j, b_k^+] \; = \; \delta_{jk} \, , \; j \; = \; 1, 2 \; .$$

(a) Show that the eigenstates are given by

$$|n_1, n_2 \rangle \; = \; \frac{1}{\sqrt{n_1!}} \left(b_1^+ \right)^{n_1} |0\rangle_1 \, \frac{1}{\sqrt{n_2!}} \left(b_2^+ \right)^{n_2} |0\rangle_2$$

where the vacuum states are defined through $b_j |0\rangle_j \; = \; 0$. State the corresponding eigenvalues and the degree of degeneracy, i.e., the number of states to the possible energy eigenvalues.
(b) Show that the three operators

$$I_1 \; = \; \frac{1}{2} (b_1^+ b_2 \; + \; b_2^+ b_1) \, , \; I_2 \; = \; \frac{1}{2i} (b_1^+ b_2 \; - \; b_2^+ b_1) \, , \; I_3 \; = \; \frac{1}{2} (b_1^+ b_1 \; + \; b_2^+ b_2)$$

satisfy the Lie algebra of SU(2), i.e. $[I_j, I_k] \; = \; i \epsilon_{jk\ell} I_\ell$. Construct, using operators $I_\pm \; = \; I_1 \pm i I_2$ and the subspace $\{ |n_1, n_2 \rangle$, $n \; = \; n_1 + n_2$ fixed, $n_1 \; = \; 0, 1, 2, \ldots, n \}$ eigenstates of $I^2 \; = \; I_1^2 + I_2^2 + I_3^2$ and of I_3

$$I^2 || \lambda, m \rangle \; = \; \lambda || \lambda, m \rangle \, ; \; I_3 || \lambda, m \rangle \; = \; m || \lambda, m \rangle$$

where $\lambda \; = \; 0, 1 \; 2, \ldots \, , \; m \; = \; -\lambda, -\lambda + 1, \ldots, \lambda$. Show $\lambda \; = \; \frac{n_1 + n_2}{2} (\frac{n_1 + n_2}{2} \; + \; 1)$ and $m \; = \; \frac{1}{2} (n_1 \; - \; n_2)$.

5.11 Evaluation of the Elements $d^j_{m\,m'}(\beta)$ of the Wigner Rotation Matrix

The spinor algorithm allows one to derive expressions for the Wigner rotation matrix elements $d^j_{m\,m'}(\beta)$. For this purpose we note that the states $|\Psi(jm)\rangle$ in a rotated coordinate system according to (5.254) are

$$|\Psi'\left(j,m'\right)\rangle = \frac{\left(b'^{\dagger}_{+}\right)^{j+m'}}{\sqrt{(j+m')!}} \frac{\left(b'^{\dagger}_{-}\right)^{j-m'}}{\sqrt{(j-m')!}} |\Psi_0\rangle \tag{5.306}$$

where $b'^{\dagger}_{+}$ and $b'^{\dagger}_{-}$ are given by (5.251). On the other side the states $|\Psi'\left(j,m'\right)\rangle$ are related to the states $|\Psi\left(j,m\right)\rangle$ in the original coordinate system by

$$|\Psi'\left(j,m'\right)\rangle \;=\; \sum_{m=-j}^{j} d^{(j)}_{m\,m'}(\beta)\,|\Psi\left(j,m\right)\rangle\;. \tag{5.307}$$

Comparision of (5.306) and (5.307) shows that the elements of the rotation matrix can be obtained by binomial expansion of $b'^{\dagger}_{+}$ and $b'^{\dagger}_{-}$ in terms of $b^{\dagger}_{+}$ and $b^{\dagger}_{-}$. For this purpose we expand

$$\begin{aligned}
&\left(b'^{\dagger}_{+}\right)^{j+m'} \left(b'^{\dagger}_{-}\right)^{j-m'} \;= \\
&\left(\cos\tfrac{\beta}{2} b^{\dagger}_{+} \;+\; \sin\tfrac{\beta}{2} b^{\dagger}_{-}\right)^{j+m'} \left(-\sin\tfrac{\beta}{2} b^{\dagger}_{+} \;+\; \cos\tfrac{\beta}{2} b^{\dagger}_{-}\right)^{j-m'} \;= \\
&\textstyle\sum_{\sigma'=0}^{j+m'} \sum_{\sigma=0}^{j-m'} \left(\begin{array}{c} j+m' \\ \sigma' \end{array}\right) \left(\cos\tfrac{\beta}{2}\right)^{\sigma'} \left(\sin\tfrac{\beta}{2}\right)^{j+m'-\sigma'} \\
&\left(\begin{array}{c} j-m' \\ \sigma \end{array}\right) \left(-\sin\tfrac{\beta}{2}\right)^{\sigma} \left(\cos\tfrac{\beta}{2}\right)^{j-m'-\sigma} \left(b^{\dagger}_{+}\right)^{\sigma'+\sigma} \left(b^{\dagger}_{-}\right)^{2j-\sigma'-\sigma}
\end{aligned} \tag{5.308}$$

The latter sum involves terms $(b^{\dagger}_{+})^{j+m}(b^{\dagger}_{-})^{j-m}$ for $\sigma' + \sigma \;=\; j+m$ and $2j-\sigma'-\sigma \;=\; j-m$. One may expect that these two conditions restrict both σ' and σ. However, both conditions are satisfied for $\sigma' \;=\; j+m-\sigma$. The combination of σ, σ' values which yields $(b^{\dagger}_{+})^{j+m}(b^{\dagger}_{-})^{j-m}$ is then $\sigma' \;=\; j+m-\sigma$. The prefactor of $(b^{\dagger}_{+})^{j+m}(b^{\dagger}_{-})^{j-m}$ which according to (5.306,5.307) can be identified with the elements $d^{(j)}_{m\,m'}(\beta)$ of the rotation matrix can then be written

$$\begin{aligned}
d^{(j)}_{m\,m'}(\beta) \;&=\; \sqrt{\frac{(j+m)!(j-m)!}{(j+m')!(j-m')!}} \sum_{\sigma=0}^{j-m'} \left(\begin{array}{c} j+m' \\ j+m-\sigma \end{array}\right) \left(\begin{array}{c} j-m' \\ \sigma \end{array}\right) \times \\
&\times (-1)^{j-m'-\sigma} \left(\sin\frac{\beta}{2}\right)^{2j-m-m'-2\sigma} \left(\cos\frac{\beta}{2}\right)^{m+m'+2\sigma}\;.
\end{aligned} \tag{5.309}$$

In case $j = 1$ this expression yields, for example,

$$(d^{(1)}_{m'\,m}) \;=\; \left(\begin{array}{ccc} \frac{1}{2}(1+\cos\beta) & \frac{1}{\sqrt{2}}\sin\beta & \frac{1}{2}(1-\cos\beta) \\ -\frac{1}{\sqrt{2}}\sin\beta & \cos\beta & \frac{1}{\sqrt{2}}\sin\beta \\ \frac{1}{2}(1-\cos\beta) & -\frac{1}{\sqrt{2}}\sin\beta & \frac{1}{2}(1+\cos\beta) \end{array}\right) \tag{5.310}$$

and in case $j \;=\; \frac{1}{2}$ it reduces to (5.243).

5.12 Mapping of SU(2) onto SO(3)

The representation (5.310) establishes a mapping of SU(2) onto SO(3). This can be shown by applying to the matrix $A = (d^{(1)}_{m'\,m})$ in (5.310) the similarity transformation

$$\tilde{A} \;=\; U^{\dagger}\, A\, U \tag{5.311}$$

where U is the 3×3 unitary matrix which establishes the transformation

$$\begin{pmatrix} \frac{1}{\sqrt{2}}(\,-x_1 \,-\, i x_2\,) \\ x_3 \\ \frac{1}{\sqrt{2}}(\,x_1 \,-\, i x_2\,) \end{pmatrix} \;=\; U \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \;;\quad U \;=\; \begin{pmatrix} -\frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} & 0 \end{pmatrix} . \tag{5.312}$$

The choice of this transformation derives from a property shown further below, namely that the component of the vector on the l.h.s. of (5.312) transforms like an angular momentum state $|1\,m\rangle$. Hence, the matrix $\tilde{A}$ should represent a rotation around the x_2–axis in the space of vectors $(x_1, x_2, x_3) \in \mathbb{R}^{\not\#}$. Evaluation of $\tilde{A}$ yields

$$\begin{aligned} \tilde{A} \;=\; \tfrac{1}{4} \begin{pmatrix} -1 & 0 & 1 \\ i & 0 & i \\ 0 & \sqrt{2} & 0 \end{pmatrix} \begin{pmatrix} 1 \,+\, \cos\!\beta & \sqrt{2}\sin\!\beta & 1 \,-\, \cos\!\beta \\ -\sqrt{2}\sin\!\beta & 2\cos\!\beta & \sqrt{2}\sin\!\beta \\ 1 \,-\, \cos\!\beta & -\sqrt{2}\sin\!\beta & 1 \,+\, \cos\!\beta \end{pmatrix} \times \\ \times \begin{pmatrix} -1 & -i & 0 \\ 0 & 0 & \sqrt{2} \\ 1 & -i & 0 \end{pmatrix} \;=\; \tfrac{1}{4} \begin{pmatrix} -2\cos\!\beta & -2\sqrt{2}\sin\!\beta & 2\cos\!\beta \\ 2i & 0 & 2i \\ -2\sin\!\beta & 2\sqrt{2}\cos\!\beta & 2\sin\!\beta \end{pmatrix} \times \\ \times \begin{pmatrix} -1 & -i & 0 \\ 0 & 0 & \sqrt{2} \\ 1 & -i & 0 \end{pmatrix} \;=\; \begin{pmatrix} \cos\!\beta & 0 & -\sin\!\beta \\ 0 & 1 & 0 \\ \sin\!\beta & 0 & \cos\!\beta \end{pmatrix} \end{aligned} \tag{5.313}$$

which, in fact, is the expected element of SO(3), i.e., the orthogonal 3×3 matrix which describes a rotation around the x_2–axis.

It is of interest to trace the mapping from SU(2) onto SO(3) as represented by (5.310) to the SU(2) transformation assumed in deriving the general result (5.309) and the particular matrix (5.310). The SU(2) transformation entered in (5.308) and had the form (5.251). Replacing the latter transformation by its negative form, i.e.,

$$\left(b'_{+}\right)^{\dagger} \;=\; -\cos\frac{\beta}{2}\, b^{\dagger}_{+} \;-\; \sin\frac{\beta}{2}\, b^{\dagger}_{-} \quad ;\quad \left(b'_{-}\right)^{\dagger} \;=\; \sin\frac{\beta}{2}\, b^{\dagger}_{+} \;-\; \cos\frac{\beta}{2}\, b^{\dagger}_{-}\,. \tag{5.314}$$

leaves (5.308) unaltered except for a factor $(-1)^{2j}$ which multiplies then also the final result (5.309). This factor implies, however, that the the representation (5.309) does not distinguish between SU(2) transformations (5.251) and (5.314) in case of integer j–values, e.g. in case of $j = 1$. One can, therefore, conclude that the mapping from SU(2) onto SO(3) is a 2–1 mapping.

Chapter 6

Quantum Mechanical Addition of Angular Momenta and Spin

In this section we consider composite systems made up of several particles, each carrying orbital angular momentum decribed by spherical harmonics $Y_{\ell m}(\theta, \phi)$ as eigenfunctions and/or spin. Often the socalled *total angular momentum*, classically speaking the sum of all angular momenta and spins of the composite system, is the quantity of interest, since related operators, sums of orbital angular momentum and of spin operators of the particles, commute with the Hamiltonian of the composite system and, hence, give rise to good quantum numbers. We like to illustrate this for an example involving particle motion. Further below we will consider composite systems involving spin states.

Example: Three Particle Scattering

Consider the scattering of three particles A, B, C governed by a Hamiltonian **H** which depends only on the internal coordinates of the system, e.g., on the distances between the three particles, but neither on the position of the center of mass of the particles nor on the overall orientation of the three particle system with respect to a laboratory–fixed coordinate frame.
To specify the dependency of the Hamiltonian on the particle coordinates we start from the nine numbers which specify the Cartesian components of the three position vectors $\vec{r}_A$, $\vec{r}_B$, $\vec{r}_C$ of the particles. Since the Hamiltonian does not depend on the position of the center of mass $\vec{R} = (m_A\vec{r}_A + m_B\vec{r}_B + m_C\vec{r}_C)/(m_A+m_B+m_C)$, six parameters must suffice to describe the interaction of the system. The overall orientation of any three particle configuration can be specified by three parameters[1], e.g., by a rotational vector $\vec{\vartheta}$. This eleminates three further parameters from the dependency of the Hamiltonian on the three particle configuration and one is left with three parameters. How should they be chosen?
Actually there is no unique choice. We like to consider a choice which is physically most reasonable in a situation that the scattering proceeds such that particles A and B are bound, and particle C impinges on the compound AB coming from a large distance. In this case a proper choice for a description of interactions would be to consider the vectors $\vec{r}_{AB} = \vec{r}_A - \vec{r}_B$ and $\vec{\rho}_C = (m_A\vec{r}_A + m_B\vec{r}_B)/(m_A + m_B) - \vec{r}_C$, and to express the Hamiltonian in terms of $|\vec{r}_{AB}|$, $|\vec{\rho}_C|$, and $\sphericalangle(\vec{r}_{AB}, \vec{\rho}_C)$. The rotational part of the scattering motion is described then in terms of the unit vectors $\hat{r}_{AB}$ and

[1] We remind the reader that, for example, three Eulerian angles α, β, γ are needed to specify a general rotational transformation

$\hat{\rho}_C$, each of which stands for two angles. One may consider then to describe the motion in terms of products of spherical harmonics $Y_{\ell_1 m_1}(\hat{r}_{AB})\, Y_{\ell_2 m_2}(\hat{\rho}_C)$ describing rotation of the compound AB and the orbital angular momentum of C around AB.
One can describe the rotational degrees of freedom of the three-particle scattering process through the basis

$$\begin{aligned} \mathcal{B} \;\; = \;\; & \{\, Y_{\ell_1 m_1}(\hat{r}_{AB})\, Y_{\ell_2 m_2}(\hat{\rho}_C), \ell_1 \,=\, 0, 1, \ldots, \ell_{1,max}, -\ell_1 \,\leq m_1 \,\leq\, \ell_1; \\ & \ell_2 \,=\, 0, 1, \ldots, \ell_{2,max}, -\ell_2 \,\leq m_1 \,\leq\, \ell_2 \,\} \end{aligned} \tag{6.1}$$

where $\ell_{1,max}$ and $\ell_{2,max}$ denote the largest orbital and rotational angular momentum values, the values of which are determined by the size of the interaction domain Δ_V, by the total energy E, by the masses m_A, m_B, m_C, and by the moment of inertia I_{A-B} of the diatomic molecule A–B approximately as follows

$$\ell_{1,max} \;=\; \frac{\Delta}{\hbar}\sqrt{\frac{2\, m_A\, m_B\, m_C\, E}{m_A m_B \,+\, m_B m_C \,+\, m_A m_C}}\,, \quad \ell_{2,max} \;=\; \frac{1}{\hbar}\sqrt{2\, I_{A-B} E}\,. \tag{6.2}$$

The dimension $d(\mathcal{B})$ of $\mathcal{B}$ is

$$d(\mathcal{B}) \;=\; \sum_{\ell_1=0}^{\ell 1,max} (2\,\ell_1 \,+\, 1) \sum_{\ell_2=0}^{\ell 2,max} (2\,\ell_2 \,+\, 1) \;=\; (\ell_{1,max} \,+\, 1)^2\, (\ell_{2,max} \,+\, 1)^2 \tag{6.3}$$

For rather moderate values $\ell_{1,max} \,=\, \ell_{2,max} \,=\, 10$ one obtains $d(\mathcal{B}) \,=\, 14\,641$, a very large number. Such large number of dynamically coupled states would constitute a serious problem in any detailed description of the scattering process, in particular, since further important degrees of freedom, i.e., vibrations and rearrangement of the particles in reactions like $AB \,+\, C \,\rightarrow\, A \,+\, BC$, have not even be considered. The rotational symmetry of the interaction between the particles allows one, however, to separate the 14 641 dimensional space of rotational states $Y_{\ell_1 m_1}(\hat{r}_{AB})\, Y_{\ell_2 m_2}(\hat{\rho}_C)$ into subspaces $\mathcal{B}_k$, $\mathcal{B}_1 \oplus \mathcal{B}_2 \oplus \ldots \,=\, \mathcal{B}$ such, that only states within the subspaces $\mathcal{B}_k$ are coupled in the scattering process. In fact, as we will demonstrate below, the dimensions $d(\mathcal{B}_k)$ of these subspaces does not exceed 100. Such extremely useful transformation of the problem can be achieved through the choice of a new basis set

$$\mathcal{B}' \;=\; \{\, \sum_{\substack{\ell_1, m_1 \\ \ell_2, m_2}} c^{(n)}_{\ell 1, m_1; \ell_2, m_2} Y_{\ell_1 m_1}(\hat{r}_{AB})\, Y_{\ell_2 m_2}(\hat{\rho}_C),\; n \,=\, 1, 2, \ldots 14\,641\} \,. \tag{6.4}$$

The basis set which provides a maximum degree of decoupling between rotational states is of great principle interest since the new states behave in many respects like states with the attributes of a single angular momentum state: to an observer the three particle system prepared in such states my look like a two particle system governed by a single angular momentum state. Obviously, composite systems behaving like elementary objects are common, albeit puzzling, and the following mathematical description will shed light on their ubiquitous appearence in physics, in fact, will make their appearence a natural consequence of the symmetry of the building blocks of matter.
There is yet another important reason why the following section is of fundamental importance for the theory of the microscopic world governed by Quantum Mechanics, rather than by Classical Mechanics. The latter often arrives at the physical properties of composite systems by adding the

corresponding physical properties of the elementary components; examples are the total momentum or the total angular momentum of a composite object which are the sum of the (angular) momenta of the elementary components. Describing quantum mechanically a property of a composite object as a whole and relating this property to the properties of the elementary building blocks is then the quantum mechanical equivalent of the important operation of addittion. In this sense, the reader will learn in the following section how to add and subtract in the microscopic world of Quantum Physics, presumably a facility the reader would like to acquire with great eagerness.

Rotational Symmetry of the Hamiltonian

As pointed out already, the existence of a basis (6.4) which decouples rotational states is connected with the rotational symmetry of the Hamiltonian of the three particle system considered, i.e., connected with the fact that the Hamiltonian $\mathbf{H}$ does not depend on the overall orientation of the three interacting particles. Hence, rotations $\mathfrak{R}(\vec{\vartheta})$ of the wave functions $\psi(\vec{r}_{AB}, \vec{\rho}_C)$ defined through

$$\mathfrak{R}(\vec{\vartheta})\,\psi(\vec{\mathfrak{r}}_{\mathfrak{AB}}, \vec{\rho}_{\mathfrak{C}}) \;=\; \psi(\mathfrak{R}^{-1}(\vec{\vartheta})\,\vec{\mathfrak{r}}_{\mathfrak{AB}},\, \mathfrak{R}^{-1}(\vec{\vartheta})\,\vec{\rho}_{\mathfrak{C}}) \tag{6.5}$$

do not affect the Hamiltonian. To specify this property mathematically let us denote by $\mathbf{H}'$ the Hamiltonian in the rotated frame, assuming presently that $\mathbf{H}'$ might, in fact, be different from $\mathbf{H}$. It holds then $\mathbf{H}'\,\mathfrak{R}(\vec{\vartheta})\,\psi \;=\; \mathfrak{R}(\vec{\vartheta})\,\mathbf{H}\,\psi$. Since this is true for any $\psi(\vec{r}_{AB}, \vec{\rho}_C)$ it follows $\mathbf{H}'\,\mathfrak{R}(\vec{\vartheta}) = \mathfrak{R}(\vec{\vartheta})\,\mathbf{H}$, from which follows in turn the well-known result that $\mathbf{H}'$ is related to $\mathbf{H}$ through the similarity transformation $\mathbf{H}' \;=\; \mathfrak{R}(\vec{\vartheta})\,\mathbf{H}\,\mathfrak{R}^{-1}(\vec{\vartheta})$. The invariance of the Hamiltonian under overall rotations of the three particle system implies then

$$\mathbf{H} \;=\; \mathfrak{R}(\vec{\vartheta})\,\mathbf{H}\,\mathfrak{R}^{-1}(\vec{\vartheta}) \quad . \tag{6.6}$$

For the following it is essential to note that $\mathbf{H}$ is not invariant under rotations of only $\vec{r}_{AB}$ or $\vec{\rho}_C$, but solely under **simultaneous and identical** rotations of $\vec{r}_{AB}$ or $\vec{\rho}_C$.
Following our description of rotations of single particle wave functions we express (6.5) according to (5.48)

$$\mathfrak{R}(\vec{\vartheta}) \;=\; \exp\left(-\frac{\mathrm{i}}{\hbar}\vec{\vartheta}\cdot\vec{\mathcal{J}}^{(1)}\right)\,\exp\left(-\frac{\mathrm{i}}{\hbar}\vec{\vartheta}\cdot\vec{\mathcal{J}}^{(2)}\right) \tag{6.7}$$

where the generators $\vec{\mathcal{J}}^{(k)}$ are differential operators acting on $\hat{r}_{AB}$ $(k=1)$ and on $\hat{\rho}_C$ $(k=2)$. For example, according to (5.53, 5.55) holds

$$-\frac{i}{\hbar}\,\mathcal{J}_1^{(1)} \;=\; z_{AB}\frac{\partial}{\partial y_{AB}} \;-\; y_{AB}\frac{\partial}{\partial z_{AB}}\,; \quad -\frac{i}{\hbar}\,\mathcal{J}_3^{(2)} \;=\; \rho_y\frac{\partial}{\partial \rho_x} \;-\; \rho_x\frac{\partial}{\partial \rho_y}\,. \tag{6.8}$$

Obviously, the commutation relationships

$$\left[\mathcal{J}_p^{(1)},\, \mathcal{J}_q^{(2)}\right] \;=\; 0 \quad \text{for } p,q \;=\; 1,2,3 \tag{6.9}$$

hold since the components of $\vec{\mathcal{J}}^{(k)}$ are differential operators with respect to different variables. One can equivalently express therefore (6.7)

$$\mathfrak{R}(\vec{\vartheta}) \;=\; \exp\left(-\frac{\mathrm{i}}{\hbar}\vec{\vartheta}\cdot\vec{\mathbb{J}}\right) \tag{6.10}$$

where

$$\vec{\mathbb{J}} \;=\; \vec{\mathcal{J}}^{(1)} \;+\; \vec{\mathcal{J}}^{(2)} \quad . \tag{6.11}$$

By means of (6.11) we can write the condition (6.6) for rotational invariance of the Hamiltonian in the form

$$\mathbf{H} \;=\; \exp\left(-\frac{i}{\hbar}\vec{\vartheta}\cdot\vec{\mathbb{J}}\right)\,\mathbf{H}\,\exp\left(+\frac{i}{\hbar}\vec{\vartheta}\cdot\vec{\mathbb{J}}\right) \quad . \tag{6.12}$$

We consider this equation for infinitesimal rotations, i.e. for $|\vec{\vartheta}| \ll 1$. To order $O(|\vec{\vartheta}|)$ one obtains

$$\mathbf{H} \;\approx\; \left(\mathbb{1} \;-\; \frac{i}{\hbar}\vec{\vartheta}\cdot\vec{\mathbb{J}}\right)\,\mathbf{H}\,\left(\mathbb{1} \;+\; \frac{i}{\hbar}\vec{\vartheta}\cdot\vec{\mathbb{J}}\right) \;\approx\; \mathbf{H} \;+\; \frac{i}{\hbar}\,\mathbf{H}\,\vec{\vartheta}\cdot\vec{\mathbb{J}} \;-\; \frac{i}{\hbar}\,\vec{\vartheta}\cdot\vec{\mathbb{J}}\mathbf{H} \quad . \tag{6.13}$$

Since this holds for any $\vec{\vartheta}$ it follows $\mathbf{H}\vec{\mathbb{J}} - \vec{\mathbb{J}}\mathbf{H} = 0$ or, componentwise,

$$[\mathbf{H},\, \mathbb{J}_k] \;=\; 0 \quad ,\; k \;=\; 1,2,3 \quad . \tag{6.14}$$

We will refer in the following to $\mathbb{J}_k$, $k = 1,2,3$ as the three components of the *total angular momentum operator*.

The property (6.14) implies that the **total** angular momentum is conserved during the scattering process, i.e., that energy, and the eigenvalues of $\vec{\mathbb{J}}^2$ and $\mathbb{J}_3$ are good quantum numbers. To describe the scattering process of AB + C most concisely one seeks eigenstates $\mathfrak{Y}_{JM}$ of $\vec{\mathbb{J}}^2$ and $\mathbb{J}_3$ which can be expressed in terms of $Y_{\ell_1 m_1}(\hat{r}_{AB})\, Y_{\ell_2 m_2}(\hat{\rho}_C)$.

Definition of Total Angular Momentum States

The commutation property (6.14) implies that the components of the total angular momentum operator (6.12) each individually can have simultaneous eigenstates with the Hamiltonian. We suspect, of course, that the components $\mathbb{J}_k$, $k = 1,2,3$ cannot have simultaneous eigenstates among each other, a supposition which can be tested through the commutation properties of these operators. One can show readily that the commutation relationships

$$[\mathbb{J}_k, \mathbb{J}_\ell] \;=\; i\hbar\,\epsilon_{k\ell m}\mathbb{J}_m \tag{6.15}$$

are satisfied, i.e., the operators $\mathbb{J}_k$, $k = 1,2,3$ do not commute. For a proof one uses (6.9), the properties $[\mathcal{J}_k^{(n)}, \mathcal{J}_\ell^{(n)}] \;=\; i\hbar\,\epsilon_{k\ell m}\mathcal{J}_m^{(n)}$ for $n = 1,2$ together with the property $[A, B + C] = [A, B] + [A, C]$.

We recognize, however, the important fact that the $\mathbb{J}_k$ obey the Lie algebra of SO(3). According to the theorem above this property implies that one can construct eigenstates $\mathfrak{Y}_{JM}$ of $\mathbb{J}_3$ and of

$$\mathbb{J}^2 \;=\; \mathbb{J}_1^2 \;+\; \mathbb{J}_2^2 \;+\; \mathbb{J}_3^2 \tag{6.16}$$

following the procedure stated in the theorem above [c.f. Eqs. (5.71–5.81)]. In fact, we will find that the states yield the basis $\mathcal{B}'$ with the desired property of a maximal uncoupling of rotational states.

Before we apply the procedure (5.71–5.81) we want to consider the relationship between $\mathfrak{Y}_{JM}$ and $Y_{\ell_1 m_1}(\hat{r}_{AB})\, Y_{\ell_2 m_2}(\hat{\rho}_C)$. In the following we will use the notation

$$\Omega_1 \;=\; \hat{r}_{AB}\;, \quad \Omega_2 \;=\; \hat{\rho}_C\;. \tag{6.17}$$

6.1 Clebsch-Gordan Coefficients

In order to determine $\mathfrak{Y}_{JM}$ we notice that the states $Y_{\ell_1 m_1}(\Omega_1)\, Y_{\ell_2 m_2}(\Omega_2)$ are characterized by **four** quantum numbers corresponding to eigenvalues of $\left[\mathcal{J}^{(1)}\right]^2$, $\mathcal{J}_3^{(1)}$, $\left[\mathcal{J}^{(2)}\right]^2$, and $\mathcal{J}_3^{(2)}$. Since $\mathfrak{Y}_{JM}$ sofar specifies solely **two** quantum numbers, two further quantum numbers need to be specified for a complete characterization of the total angular momentum states. The two missing quantum numbers are ℓ_1 and ℓ_2 corresponding to the eigenvalues of $\left[\mathcal{J}^{(1)}\right]^2$ and $\left[\mathcal{J}^{(2)}\right]^2$. We, therefore, assume the expansion

$$\mathfrak{Y}_{JM}(\ell_1,\ell_2|\Omega_1,\Omega_2) \;=\; \sum_{m_1,m_2} (J\,M|\ell_1\,m_1\,\ell_2\,m_2)\, Y_{\ell_1 m_1}(\Omega_1)\, Y_{\ell_2 m_2}(\Omega_2) \tag{6.18}$$

where the states $\mathfrak{Y}_{JM}(\ell_1,\ell_2|\Omega_1,\hat{\rho}_C)$ are normalized. The expansion coefficients $(J\,M|\ell_1\,m_1\,\ell_2\,m_2)$ are called the *Clebsch-Gordan coefficients* which we seek to determine now. These coefficients, or the closely related *Wigner 3j-coefficients* introduced further below, play a cardinal role in the mathematical description of microscopic physical systems. Equivalent coefficients exist for other symmetry properties of multi–component systems, an important example being the symmetry groups SU(N) governing elementary particles made up of two quarks, i.e., mesons, and three quarks, i.e., baryons.

Exercise 6.1.1: Show that $\mathbb{J}^2, \mathbb{J}_3, \left(\mathcal{J}^{(1)}\right)^2, \left(\mathcal{J}^{(2)}\right)^2$, and $\vec{\mathcal{J}}^{(1)}\cdot\vec{\mathcal{J}}^{(2)}$ commute. Why can states $\mathfrak{Y}_{JM}$ then not be specified by 5 quantum numbers?

Properties of Clebsch-Gordan Coefficients

A few important properties of *Clebsch-Gordan coefficients* can be derived rather easily. We first notice that $\mathfrak{Y}_{JM}$ in (6.18) is an eigenfunction of $\mathbb{J}_3$, the eigenvalue being specified by the quantum number M, i.e.

$$\mathbb{J}_3\,\mathfrak{Y}_{JM} \;=\; \hbar M\,\mathfrak{Y}_{JM}\;. \tag{6.19}$$

Noting $\mathbb{J}_3 \;=\; \mathcal{J}_3^{(1)} \;+\; \mathcal{J}_3^{(2)}$ and applying this to the l.h.s. of (6.18) yields using the property $\mathcal{J}_3^{(k)} Y_{\ell_k m_k}(\Omega_k) \;=\; \hbar m_k\, Y_{\ell_k m_k}(\Omega_k)\,,\; k=1,2$

$$\begin{aligned} &M\,\mathfrak{Y}_{JM}(\ell_1,\ell_2|\Omega_1,\Omega_2) = \\ &\qquad \sum_{m_1,m_2} (m_1 \;+\; m_2)\,(J\,M|\ell_1\,m_1\,\ell_2\,m_2)\, Y_{\ell_1 m_1}(\Omega_1)\, Y_{\ell_2 m_2}(\Omega_2)\;. \end{aligned} \tag{6.20}$$

This equation can be satisfied only if the *Clebsch-Gordan coefficients* satisfy

$$(J\,M|\ell_1\,m_1\,\ell_2\,m_2) \;=\; 0 \quad \text{for } m_1 \;+\; m_2 \neq M \quad . \tag{6.21}$$

One can, hence, restrict the sum in (6.18) to avoid summation of vanishing terms

$$\mathfrak{Y}_{JM}(\ell_1,\ell_2|\Omega_1,\Omega_2) \;=\; \sum_{m_1} (J\,M|\ell_1\,m_1\,\ell_2\,M-m_1)\, Y_{\ell_1 m_1}(\Omega_1)\, Y_{\ell_2 m_2}(\Omega_2) \quad . \tag{6.22}$$

We will not adopt such explicit summation since it leads to cumbersum notation. However, the reader should always keep in mind that conditions equivalent to (6.21) hold.

The expansion (6.18) constitutes a change of an orthonormal basis

$$\begin{aligned}\mathfrak{B}(\ell_1,\ell_2) \;&=\; \{Y_{\ell_1 m_1}(\Omega_1)\, Y_{\ell_2 m_2}(\Omega_2), m_1 \;=\; -\ell_1, -\ell_1+1, \ldots, \ell_1\,, \\ &\quad m_2 \;=\; -\ell_2, -\ell_2+1, \ldots, \ell_2\}\,,\end{aligned} \tag{6.23}$$

corresponding to the r.h.s., to a new basis $\mathfrak{B}'(\ell_1,\ell_2)$ corresponding to the l.h.s. The orthonormality property implies

$$\int d\Omega_1 \int d\Omega_2\; Y_{\ell_1 m_1}(\Omega_1)\, Y_{\ell_2 m_2}(\Omega_2) Y_{\ell_1' m_1'}\Omega_1)\, Y_{\ell_2' m_2'}(\Omega_2) \;=\; \delta_{\ell_1 \ell_1'}\delta_{m_1 m_1'}\delta_{\ell_2 \ell_2'}\delta_{m_2 m_2'}\,. \tag{6.24}$$

The basis $\mathfrak{B}(\ell_1,\ell_2)$ has $(2\ell_1+1)(2\ell_2+1)$ elements. The basis $\mathfrak{B}'(\ell_1,\ell_2)$ is also orthonormal[2] and must have the same number of elements. For each quantum number J there should be $2J+1$ elements $\mathfrak{Y}_{JM}$ with $M = -J, -J+1, \ldots, J$. However, it is not immediately obvious what the J–values are. Since $\mathfrak{Y}_{JM}$ represents the total angular momentum state and $Y_{\ell_1 m_1}(\Omega_1)$ and $Y_{\ell_2 m_2}(\Omega_2)$ the individual angular momenta one may start from one's classical notion that these states represent angular momentum vectors $\vec{J}$, $\vec{J}^{(1)}$ and $\vec{J}^{(2)}$, respectively. In this case the range of $|\vec{J}|$–values would be the interval $[\Big||\vec{J}^{(1)}| - |\vec{J}^{(2)}|\Big|, |\vec{J}^{(1)}| + |\vec{J}^{(2)}|]$. This obviously corresponds quantum mechanically to a range of J–values $J = |\ell_1-\ell_2|, |\ell_1-\ell_2|+1, \ldots \ell_1+\ell_2$. In fact, it holds

$$\sum_{J=|\ell_1-\ell_2|}^{\ell_1+\ell_2} (\,2J\,+\,1\,) \;=\; (2\ell_1+1)\,(2\ell_2+1) \quad , \tag{6.25}$$

i.e., the basis $\mathfrak{B}'(\ell_1,\ell_2)$ should be

$$\begin{aligned}\mathfrak{B}_2 \;&=\; \{\mathfrak{Y}_{JM}(\ell_1,\ell_2|\Omega_1,\Omega_2);\; J \;=\; |\ell_1-\ell_2|,\, |\ell_1-\ell_2|+1, \ell_1+\ell_2\,, \\ &\quad M \;=\; -J, -J+1, \ldots, J\}\,.\end{aligned} \tag{6.26}$$

We will show below in an explicit construction of the *Clebsch-Gordan coefficients* that, in fact, the range of values assumed for J is correct. Our derivation below will also yield real values for the *Clebsch-Gordan coefficients*.

Exercise 6.1.2: Prove Eq. (6.25)

We want to state now two summation conditions which follow from the orthonormality of the two basis sets $\mathfrak{B}(\ell_1,\ell_2)$ and $\mathfrak{B}'(\ell_1,\ell_2)$. The property

$$\int d\Omega_1 \int d\Omega_2\; \mathfrak{Y}^*_{JM}(\ell_1,\ell_2|\Omega_1,\Omega_2)\, \mathfrak{Y}_{J'M'}(\ell_1,\ell_2|\Omega_1,\Omega_2) \;=\; \delta_{JJ'}\delta_{MM'} \tag{6.27}$$

together with (6.18) applied to $\mathfrak{Y}^*_{JM}$ and to $\mathfrak{Y}_{J'M'}$ and with (6.24) yields

$$\sum_{m_1,m_2} (J\,M|\ell_1\,m_1\,\ell_2\,m_2)^*(J'\,M'|\ell_1\,m_1\,\ell_2\,m_2) \;=\; \delta_{JJ'}\delta_{MM'} \quad . \tag{6.28}$$

[2]This property follows from the fact that the basis elements are eigenstates of hermitian operators with different eigenvalues, and that the states can be normalized.

The second summation condition starts from the fact that the basis sets $\mathfrak{B}(\ell_1, \ell_2)$ and $\mathfrak{B}'(\ell_1, \ell_2)$ span the **same** function space. Hence, it is possible to expand $Y_{\ell_1 m_1}(\Omega_1) Y_{\ell_2 m_2}(\Omega_2)$ in terms of $\mathfrak{Y}_{JM}(\ell_1, \ell_2 | \Omega_1, \Omega_2)$, i.e.,

$$Y_{\ell_1 m_1}(\Omega_1) Y_{\ell_2 m_2}(\Omega_2) \;=\; \sum_{J'=|\ell_1-\ell_2|}^{\ell_1+\ell_2} \sum_{M'=-J}^{J} c_{J'M'}\, \mathfrak{Y}_{J'M'}(\ell_1, \ell_2 | \Omega_1, \Omega_2) \;, \tag{6.29}$$

where the expansion coefficients are given by the respective scalar products in function space

$$c_{J'M'} \;=\; \int d\Omega_1 \int d\Omega_2\; \mathfrak{Y}^*_{J'M'}(\ell_1, \ell_2 | \Omega_1, \Omega_2) Y_{\ell_1 m_1}(\Omega_1) Y_{\ell_2 m_2}(\Omega_2) \quad . \tag{6.30}$$

The latter property follows from multiplying (6.18) by $\mathfrak{Y}^*_{J'M'}(\ell_1, \ell_2 | \Omega_1, \Omega_2)$ and integrating. The orthogonality property (6.27) yields

$$\delta_{JJ'} \delta_{MM'} \;=\; \sum_{m_1, m_2} (J\, M | \ell_1\, m_1\, \ell_2\, m_2)\, c_{J'M'} \quad . \tag{6.31}$$

Comparision with (6.28) allows one to conclude that the coefficients $c_{J'M'}$ are identical to $(J'\, M' | \ell_1\, m_1\, \ell_2\, m_2)^*$, i.e.,

$$\begin{aligned} & Y_{\ell_1 m_1}(\Omega_1) Y_{\ell_2 m_2}(\Omega_2) \\ &= \sum_{J'=|\ell_1-\ell_2|}^{\ell_1+\ell_2} \sum_{M'=-J}^{J} (J'\, M' | \ell_1\, m_1\, \ell_2\, m_2)^*\, \mathfrak{Y}_{J'M'}(\ell_1, \ell_2 | \Omega_1, \Omega_2) \;, \end{aligned} \tag{6.32}$$

which complements (6.18). One can show readily using the same reasoning as applied in the derivation of (6.28) from (6.18) that the *Clebsch-Gordan coefficients* obey the second summation condition

$$\sum_{JM} (J\, M | \ell_1\, m_1\, \ell_2\, m_2)^* (J\, M | \ell_1\, m'_1\, \ell_2\, m'_2) \;=\; \delta_{m_1 m'_1}\, \delta_{m_2 m'_2} \quad . \tag{6.33}$$

The latter summation has not been restricted explicitly to allowed J–values, rather the convention

$$(J\, M | \ell_1\, m_1\, \ell_2\, m_2) \;=\; 0 \quad \text{if } J \,<\, |\ell_1 - \ell_2| \,, \text{ or } J \,>\, \ell_1 + \ell_2 \tag{6.34}$$

has been assumed.

6.2 Construction of Clebsch-Gordan Coefficients

We will now construct the *Clebsch-Gordan coefficients*. The result of this construction will include all the properties previewed above. At this point we like to stress that the construction will be based on the theorems (5.71–5.81) stated above, i.e., will be based solely on the commutation properties of the operators $\vec{\mathbb{J}}$ and $\vec{\mathcal{J}}^{(k)}$. We can, therefore, also apply the results, and actually also the properties of *Clebsch-Gordan coefficients* stated above, to composite systems involving spin-$\frac{1}{2}$ states. A similar construction will also be applied to composite systems governed by other symmetry groups, e.g., the group SU(3) in case of meson multiplets involving two quarks, or baryons multiplets involving three quarks.

For the construction of $\mathfrak{Y}_{JM}$ we will need the operators

$$\mathbb{J}_\pm \;=\; \mathbb{J}_1 \,+\, i\mathbb{J}_2 \;. \tag{6.35}$$

The construction assumes a particular choice of $J \in \{|\ell_1 - \ell_2|,\, |\ell_1 - \ell_2| + 1, \ldots \ell_1 + \ell_2\}$ and for such J–value seeks an expansion (6.18) which satisfies

$$\mathbb{J}_+ \,\mathfrak{Y}_{JJ}(\ell_1, \ell_2|\Omega_1, \Omega_2) \qquad\qquad = \; 0 \tag{6.36}$$

$$\mathbb{J}_3 \,\mathfrak{Y}_{JJ}(\ell_1, \ell_2|\Omega_1, \Omega_2) \quad = \; \hbar\, J\, \mathfrak{Y}_{JJ}(\ell_1, \ell_2|\Omega_1, \Omega_2) \;. \tag{6.37}$$

The solution needs to be normalized. Having determined such $\mathfrak{Y}_{JJ}$ we then construct the whole family of functions $\mathbb{X}_J \,=\, \{\mathfrak{Y}_{JM}(\ell_1, \ell_2|\Omega_1, \Omega_2),\; M \,=\, -J, -J+1, \ldots J\}$ by applying repeatedly

$$\mathbb{J}_- \,\mathfrak{Y}_{JM+1}(\ell_1, \ell_2|\Omega_1, \Omega_2) \;=\; \hbar\,\sqrt{(J+M+1)(J-M)}\mathfrak{Y}_{JM}(\ell_1, \ell_2|\Omega_1, \Omega_2) \quad . \tag{6.38}$$

for $M \,=\, J-1, J-2, \ldots, -J$.
We embark on the suggested construction for the choice $J \,=\, \ell_1 + \ell_2$. We first seek an unnormalized solution $\mathcal{G}_{JJ}$ and later normalize. To find $\mathcal{G}_{JJ}$ we start from the observation that $\mathcal{G}_{JJ}$ represents the state with the largest possible quantum number $J \,=\, \ell_1 \,+\, \ell_2$ with the largest possible component $M \,=\, \ell_1 \,+\, \ell_2$ along the z–axis. The corresponding classical total angular momentum vector $\vec{\mathbb{J}}_{\rm class}$ would be obtained by aligning both $\vec{\mathcal{J}}^{(1)}_{\rm class}$ and $\vec{\mathcal{J}}^{(2)}_{\rm class}$ also along the z–axis and adding these two vectors. Quantum mechanically this corresponds to a state

$$\mathcal{G}_{\ell_1+\ell_2,\ell_1+\ell_2}(\ell_1, \ell_2|\Omega_1, \Omega_2) \;=\; Y_{\ell_1\ell_1}(\Omega_1)\, Y_{\ell_2\ell_2}(\Omega_2) \tag{6.39}$$

which we will try for a solution of (6.37). For this purpose we insert (6.39) into (6.37) and replace according to (6.11) $\mathbb{J}_+$ by $\mathcal{J}^{(1)}_+ \,+\, \mathcal{J}^{(2)}_+$. We obtain using (5.66,5.68)

$$\begin{aligned} &\left(\mathcal{J}^{(1)}_+ \,+\, \mathcal{J}^{(2)}_+\right)\, Y_{\ell_1\ell_1}(\Omega_1)\, Y_{\ell_2\ell_2}(\Omega_2) \\ &=\; \left(\mathcal{J}^{(1)}_+ Y_{\ell_1\ell_1}(\Omega_1)\right) Y_{\ell_2\ell_2}(\Omega_2) \,+\, Y_{\ell_1\ell_1}(\Omega_1)\left(\mathcal{J}^{(2)}_+ Y_{\ell_2\ell_2}(\Omega_2)\right) \;=\; 0 \quad . \end{aligned} \tag{6.40}$$

Similarly, we can demonstrate condition (6.25) using (6.11) and (5.64)

$$\begin{aligned} &\left(\mathcal{J}^{(1)}_3 \,+\, \mathcal{J}^{(2)}_3\right)\, Y_{\ell_1\ell_1}(\Omega_1)\, Y_{\ell_2\ell_2}(\Omega_2) \\ &=\; \left(\mathcal{J}^{(1)}_3 Y_{\ell_1\ell_1}(\Omega_1)\right) Y_{\ell_2\ell_2}(\Omega_2) \,+\, Y_{\ell_1\ell_1}(\Omega_1)\left(\mathcal{J}^{(2)}_3 Y_{\ell_2\ell_2}(\Omega_2)\right) \\ &=\; \hbar\,(\ell_1 \,+\, \ell_2)\, Y_{\ell_1\ell_1}(\Omega_1)\, Y_{\ell_2\ell_2}(\Omega_2) \;. \end{aligned} \tag{6.41}$$

In fact, we can also demonstrate using (**??**) that $\mathcal{G}_{\ell_1+\ell_2,\ell_1+\ell_2}(\ell_1, \ell_2|\Omega_1, \Omega_2)$ is normalized

$$\begin{aligned} &\int d\Omega_1 \int d\Omega_2\, \mathcal{G}_{\ell_1+\ell_2,\ell_1+\ell_2}(\ell_1, \ell_2|\Omega_1, \Omega_2) \\ &=\; \left(\int d\Omega_1 Y_{\ell_1\ell_1}(\Omega_1)\right)\left(\int d\Omega_2 Y_{\ell_2\ell_2}(\Omega_2)\right) \;=\; 1 \;. \end{aligned} \tag{6.42}$$

We, therefore, have shown

$$\mathfrak{Y}_{\ell_1+\ell_2,\ell_1+\ell_2}(\ell_1, \ell_2|\Omega_1, \Omega_2) \;=\; Y_{\ell_1\ell_1}(\Omega_1)\, Y_{\ell_2\ell_2}(\Omega_2) \;. \tag{6.43}$$

We now employ property (6.38) to construct the family of functions $\mathbb{B}_{\ell_\curlywedge+\ell_\curlyvee} = \{\mathfrak{Y}_{\ell_\curlywedge+\ell_\curlyvee\,\mathbb{M}}(\ell_\curlywedge,\ell_\curlyvee|\measuredangle_\curlywedge,\measuredangle_\curlyvee), \mathbb{M} = -(\ell_\curlywedge+\ell_\curlyvee),\ldots,(\ell_\curlywedge+\ell_\curlyvee)\}$. We demonstrate the procedure explicitly only for $M = \ell_1+\ell_2-1$. The r.h.s. of (6.38) yields with $\mathbb{J}_- = \mathcal{J}_-^{(1)} + \mathcal{J}_-^{(1)}$ the expression $\hbar\sqrt{2\ell_1}\,Y_{\ell_1\ell_1-1}(\Omega_1)Y_{\ell_2\ell_2}(\Omega_2) + \hbar\sqrt{2\ell_2}\,Y_{\ell_1\ell_1-1}(\Omega_1)Y_{\ell_2\ell_2-1}(\Omega_2)$. The l.h.s. of (6.38) yields $\hbar\sqrt{2(\ell_1+\ell_2)}\mathfrak{Y}_{\ell_1+\ell_2\ell_1+\ell_2-1}(\ell_1,\ell_2|\Omega_1,\Omega_2)$. One obtains then

$$\begin{aligned}&\mathfrak{Y}_{\ell_1+\ell_2\,\ell_1+\ell_2-1}(\ell_1,\ell_2|\Omega_1,\Omega_2) = \\ &\sqrt{\tfrac{\ell_1}{\ell_1+\ell_2}}\,Y_{\ell_1\ell_1-1}(\Omega_1)Y_{\ell_2\ell_2}(\Omega_2) + \sqrt{\tfrac{\ell_2}{\ell_1+\ell_2}}\,Y_{\ell_1\ell_1}(\Omega_1)Y_{\ell_2\ell_2-1}(\Omega_2)\,.\end{aligned} \tag{6.44}$$

This construction can be continued to obtain all $2(\ell_1+\ell_2)+1$ elements of $\mathbb{B}_{\ell_\curlywedge+\ell_\curlyvee}$ and, thereby, all the *Clebsch-Gordan coefficients* $(\ell_1+\ell_2 M|\ell_1 m_1\ell_2 m_2)$. We have provided in Table 1 the explicit form of $\mathfrak{Y}_{J\,M}(\ell_1\ell_2|\Omega_1\Omega_2)$ for $\ell_1=2$ and $\ell_2=1$ to illustrate the construction. The reader should familiarize himself with the entries of the Table, in particular, with the symmetry pattern and with the terms $Y_{\ell_1 m_1}Y_{\ell_2 m_2}$ contributing to each $\mathfrak{Y}_{J\,M}$.

We like to construct now the family of total angular momentum functions $\mathbb{B}_{\ell_\curlywedge+\ell_\curlyvee-\curlywedge} = \{\mathfrak{Y}_{\ell_\curlywedge+\ell_\curlyvee-\curlywedge\,\mathbb{M}}(\ell_\curlywedge,\ell_\curlyvee|\measuredangle_\curlywedge,\measuredangle_\curlyvee), \mathbb{M} = -(\ell_\curlywedge+\ell_\curlyvee-\curlywedge),\ldots,(\ell_\curlywedge+\ell_\curlyvee-\curlywedge)\}$. We seek for this purpose first an unnormalized solution $\mathcal{G}_{\ell_1+\ell_2-1\,\ell_1+\ell_2-1}$ of (6.36, 6.37). According to the condition (6.21) we set

$$\mathcal{G}_{\ell_1+\ell_2-1\,\ell_1+\ell_2-1}(\ell_1\ell_2|\Omega_1\Omega_2) = Y_{\ell_1\ell_1-1}(\Omega_1)Y_{\ell_2\ell_2}(\Omega_2) + c\,Y_{\ell_1\ell_1}(\Omega_1)Y_{\ell_2\ell_2-1}(\Omega_2) \tag{6.45}$$

for some unknown constant c. One can readily show that (6.37) is satisfied. To demonstrate (6.36) we proceed as above and obtain

$$\begin{aligned}&\left(\mathcal{J}_+^{(1)}Y_{\ell_1\ell_1-1}(\Omega_1)\right)Y_{\ell_2\ell_2}(\Omega_2) + c\,Y_{\ell_1\ell_1}(\Omega_1)\left(\mathcal{J}_+^{(2)}Y_{\ell_2\ell_2-1}(\Omega_2)\right)\\ &= \left(\sqrt{2\ell_1} + c\sqrt{2\ell_2}\right)Y_{\ell_1\ell_1}(\Omega_1)\,Y_{\ell_2\ell_2}(\Omega_2) = 0\,.\end{aligned} \tag{6.46}$$

To satisfy this equation one needs to choose $c = -\sqrt{\ell_1/\ell_2}$. We have thereby determined $\mathcal{G}_{\ell_1+\ell_2-1\,\ell_1+\ell_2-1}$ in (6.45). Normalization yields furthermore

$$\begin{aligned}&\mathfrak{Y}_{\ell_1+\ell_2-1\,\ell_1+\ell_2-1}(\ell_1\ell_2|\Omega_1\Omega_2)\\ &= \sqrt{\frac{\ell_2}{\ell_1+\ell_2}}\,Y_{\ell_1\ell_1-1}(\Omega_1)Y_{\ell_2\ell_2}(\Omega_2) - \sqrt{\frac{\ell_1}{\ell_1+\ell_2}}\,Y_{\ell_1\ell_1}(\Omega_1)Y_{\ell_2\ell_2-1}(\Omega_2)\,.\end{aligned} \tag{6.47}$$

This expression is orthogonal to (6.39) as required by (6.27).

Expression (6.47) can serve now to obtain recursively the elements of the family $\mathbb{B}_{\ell_\curlywedge+\ell_\curlyvee-\curlywedge}$ for $M = \ell_1+\ell_2-2, \ell_1+\ell_2-3,\ldots,-(\ell_1+\ell_2-1)$. Having constructed this family we have determined the *Clebsch-Gordan coefficients* $(\ell_1+\ell_2-1 M|\ell_1 m_1\ell_2 m_2)$. The result is illustrated for the case $\ell_1=2$, $\ell_2=1$ in Table 1.

One can obviously continue the construction outlined to determine $\mathfrak{Y}_{\ell_1+\ell_2-2\,\ell_1+\ell_2-2}$, etc. and all total angular momentum functions for a given choice of ℓ_1 and ℓ_2.

Exercise 6.2.1: Construct following the procedure above the three functions $\mathfrak{Y}_{JM}(\ell_1,\ell_2|\Omega_1,\Omega 2)$ for $M = \ell_1+\ell_2-2$ and $J = \ell_1+\ell_2, \ell_1+\ell_2-1, \ell_1+\ell_2-2$. Show that the resulting functions are orthonormal.

	$Y_{22}(\Omega_1)Y_{11}(\Omega_2)$		
$\mathcal{Y}_{33}(2,1\|\Omega_1,\Omega 2)$	1		
	$Y_{21}(\Omega_1)Y_{11}(\Omega_2)$	$Y_{22}(\Omega_1)Y_{10}(\Omega_2)$	
$\mathcal{Y}_{32}(2,1\|\Omega_1,\Omega 2)$	$\sqrt{\frac{2}{3}} \simeq 0.816497$	$\sqrt{\frac{1}{3}} \simeq 0.57735$	
$\mathcal{Y}_{22}(2,1\|\Omega_1,\Omega 2)$	$-\sqrt{\frac{1}{3}} \simeq -0.57735$	$\sqrt{\frac{2}{3}} \simeq 0.816497$	
	$Y_{20}(\Omega_1)Y_{11}(\Omega_2)$	$Y_{21}(\Omega_1)Y_{10}(\Omega_2)$	$Y_{22}(\Omega_1)Y_{1-1}(\Omega_2)$
$\mathcal{Y}_{31}(2,1\|\Omega_1,\Omega 2)$	$\sqrt{\frac{2}{5}} \simeq 0.632456$	$\sqrt{\frac{8}{15}} \simeq 0.730297$	$\sqrt{\frac{1}{15}} \simeq 0.258199$
$\mathcal{Y}_{21}(2,1\|\Omega_1,\Omega 2)$	$-\sqrt{\frac{1}{2}} \simeq -0.707107$	$\sqrt{\frac{1}{6}} \simeq 0.408248$	$\sqrt{\frac{1}{3}} \simeq 0.57735$
$\mathcal{Y}_{11}(2,1\|\Omega_1,\Omega 2)$	$\sqrt{\frac{1}{10}} \simeq 0.316228$	$-\sqrt{\frac{3}{10}} \simeq -0.547723$	$\sqrt{\frac{3}{5}} \simeq 0.774597$
	$Y_{2-1}(\Omega_1)Y_{11}(\Omega_2)$	$Y_{20}(\Omega_1)Y_{10}(\Omega_2)$	$Y_{21}(\Omega_1)Y_{1-1}(\Omega_2)$
$\mathcal{Y}_{30}(2,1\|\Omega_1,\Omega 2)$	$\sqrt{\frac{1}{5}} \simeq 0.447214$	$\sqrt{\frac{3}{5}} \simeq 0.774597$	$\sqrt{\frac{1}{5}} \simeq 0.447214$
$\mathcal{Y}_{20}(2,1\|\Omega_1,\Omega 2)$	$-\sqrt{\frac{1}{2}} \simeq -0.707107$	0	$\sqrt{\frac{1}{2}} \simeq 0.707107$
$\mathcal{Y}_{10}(2,1\|\Omega_1,\Omega 2)$	$\sqrt{\frac{3}{10}} \simeq 0.547723$	$-\sqrt{\frac{2}{5}} \simeq -0.632456$	$\sqrt{\frac{3}{10}} \simeq 0.547723$
	$Y_{2-2}(\Omega_1)Y_{11}(\Omega_2)$	$Y_{2-1}(\Omega_1)Y_{10}(\Omega_2)$	$Y_{20}(\Omega_1)Y_{1-1}(\Omega_2)$
$\mathcal{Y}_{3-1}(2,1\|\Omega_1,\Omega 2)$	$\sqrt{\frac{1}{15}} \simeq 0.258199$	$\sqrt{\frac{8}{15}} \simeq 0.730297$	$\sqrt{\frac{2}{5}} \simeq 0.632456$
$\mathcal{Y}_{2-1}(2,1\|\Omega_1,\Omega 2)$	$-\sqrt{\frac{1}{3}} \simeq -0.57735$	$-\sqrt{\frac{1}{6}} \simeq -0.408248$	$\sqrt{\frac{1}{2}} \simeq 0.707107$
$\mathcal{Y}_{1-1}(2,1\|\Omega_1,\Omega 2)$	$\sqrt{\frac{3}{5}} \simeq 0.774597$	$-\sqrt{\frac{3}{10}} \simeq -0.547723$	$\sqrt{\frac{1}{10}} \simeq 0.316228$
		$Y_{2-2}(\Omega_1)Y_{10}(\Omega_2)$	$Y_{2-1}(\Omega_1)Y_{1-1}(\Omega_2)$
$\mathcal{Y}_{3-2}(2,1\|\Omega_1,\Omega 2)$		$\sqrt{\frac{1}{3}} \simeq 0.57735$	$\sqrt{\frac{2}{3}} \simeq 0.816497$
$\mathcal{Y}_{2-2}(2,1\|\Omega_1,\Omega 2)$		$-\sqrt{\frac{2}{3}} \simeq -0.816497$	$\sqrt{\frac{1}{3}} \simeq 0.57735$
			$Y_{2-2}(\Omega_1)Y_{1-1}(\Omega_2)$
$\mathcal{Y}_{3-3}(2,1\|\Omega_1,\Omega 2)$			1

Table 6.1: Some explicit analytical and numerical values of *Clebsch-Gordan coefficients* and their relationship to the total angular momentum wave functions and single particle angular momentum wave functions.

Exercise 6.2.2: Use the construction for Clebsch-Gordan coefficients above to prove the following formulas

$$\langle J,M|\ell,m-\tfrac{1}{2},\tfrac{1}{2},+\tfrac{1}{2}\rangle \;=\; \begin{cases} \sqrt{\frac{J+M}{2J}} & \text{for } \ell = J-\frac{1}{2} \\ -\sqrt{\frac{J-M+1}{2J+2}} & \text{for } \ell = J+\frac{1}{2} \end{cases}$$

$$\langle J,M|\ell,m+\tfrac{1}{2},\tfrac{1}{2},-\tfrac{1}{2}\rangle \;=\; \begin{cases} \sqrt{\frac{J-M}{2J}} & \text{for } \ell = J-\frac{1}{2} \\ \sqrt{\frac{J+M+1}{2J+2}} & \text{for } \ell = J+\frac{1}{2} \end{cases} .$$

The construction described provides a very cumbersome route to the analytical and numerical values of the *Clebsch-Gordan coefficients*. It is actually possible to state explicit expressions for any single coefficient $(JM|\ell_1 m_1 \ell_2 m_2)$. These expressions will be derived now.

6.3 Explicit Expression for the Clebsch–Gordan Coefficients

We want to establish in this Section an explicit expression for the Clebsch–Gordan coefficients $(JM|\ell_1 m_1 \ell_2 m_2)$. For this purpsose we will employ the spinor operators introduced in Sections 5.9, 5.10.

Definition of Spinor Operators for Two Particles

In contrast to Sections 5.9, 5.10 where we studied single particle angular momentum and spin, we are dealing now with two particles carrying angular momentum or spin. Accordingly, we extent definition (5.287) to two particles

$$^{(1)}J_k \;=\; \tfrac{1}{2}\textstyle\sum_{\zeta,\zeta'} a^\dagger_\zeta < \zeta\,|\sigma_k|\,\zeta' > a_{\zeta'} \tag{6.48}$$

$$^{(2)}J_k \;=\; \tfrac{1}{2}\textstyle\sum_{\zeta,\zeta'} b^\dagger_\zeta < \zeta\,|\sigma_k|\,\zeta' > b_{\zeta'} \tag{6.49}$$

where $\zeta,\zeta' = \pm$ and the matrix elements $< \zeta\,|\sigma_k|\,\zeta' >$ are as defined in Section 5.10. The creation and annihilation operators are again of the boson type with commutation properties

$$\left[a_\zeta,\, a_{\zeta'}\right] \;=\; \left[a^\dagger_\zeta,\, a^\dagger_{\zeta'}\right] \;=\; 0 \quad,\; \left[a_\zeta,\, a^\dagger_{\zeta'}\right] \;=\; \delta_{\zeta\zeta'} \tag{6.50}$$

$$\left[b_\zeta,\, b_{\zeta'}\right] \;=\; \left[b^\dagger_\zeta,\, b^\dagger_{\zeta'}\right] \;=\; 0 \quad,\; \left[b_\zeta,\, b^\dagger_{\zeta'}\right] \;=\; \delta_{\zeta\zeta'} \quad . \tag{6.51}$$

The operators $a_\zeta, a^\dagger_\zeta$ and $b_\zeta, b^\dagger_\zeta$ refer to different particles and, hence, commute with each other

$$\left[a_\zeta,\, b_{\zeta'}\right] \;=\; \left[a^\dagger_\zeta,\, b^\dagger_{\zeta'}\right] \;=\; \left[a_\zeta,\, b^\dagger_{\zeta'}\right] \;=\; 0\,. \tag{6.52}$$

According to Section 5.10 [cf. (5.254)] the angular momentum / spin eigenstates $|\ell_1 m_1\rangle_1$ and $|\ell_2 m_2\rangle_2$ of the two particles are

$$|\ell_1 m_1\rangle_1 \;=\; \frac{\left(a^\dagger_+\right)^{\ell_1+m_1}}{\sqrt{(\ell_1+m_1)!}}\,\frac{\left(a^\dagger_-\right)^{\ell_1-m_1}}{\sqrt{(\ell_1-m_1)!}}\,|\Psi_0\rangle \tag{6.53}$$

$$|\ell_2 m_2\rangle_2 \;=\; \frac{\left(b^\dagger_+\right)^{\ell_2+m_2}}{\sqrt{(\ell_2+m_2)!}}\,\frac{\left(b^\dagger_-\right)^{\ell_2-m_2}}{\sqrt{(\ell_2-m_2)!}}\,|\Psi_0\rangle\,. \tag{6.54}$$

It holds, in analogy to Eqs. (5.302, 5.303),

$$ {}^{(1)}J^2|\ell_1 m_1\rangle_1 \;=\; \ell_1(\ell_1+1)\,|\ell_1 m_1\rangle_1 \;,\; {}^{(1)}J_3|\ell_1 m_1\rangle_1 \;=\; m_1\,|\ell_1 m_1\rangle_1 \tag{6.55} $$

$$ {}^{(2)}J^2|\ell_2 m_2\rangle_2 \;=\; \ell_2(\ell_2+1)\,|\ell_2 m_2\rangle_2 \;,\; {}^{(2)}J_3|\ell_2 m_2\rangle_2 \;=\; m_2\,|\ell_2 m_2\rangle_2 \quad . \tag{6.56} $$

The states $|\ell_1, m_1\rangle_1|\ell_2, m_2\rangle_2$, which describe a two particle system according to (6.53, 6.54), are

$$ |\ell_1, m_1\rangle_1 \quad |\ell_2, m_2\rangle_2 \;= \tag{6.57} $$

$$ \frac{\left(a_+^\dagger\right)^{\ell_1+m_1}}{\sqrt{(\ell_1+m_1)!}}\frac{\left(a_-^\dagger\right)^{\ell_1-m_1}}{\sqrt{(\ell_1-m_1)!}}\frac{\left(b_+^\dagger\right)^{\ell_2+m_2}}{\sqrt{(\ell_2+m_2)!}}\frac{\left(b_-^\dagger\right)^{\ell_2-m_2}}{\sqrt{(\ell_2-m_2)!}}|\Psi_o\rangle \;. $$

The operator of the total angular momentum/spin of the two particle system is

$$ \vec{\mathbb{J}} \;=\; {}^{(1)}\vec{J} \;+\; {}^{(2)}\vec{J} \tag{6.58} $$

with Cartesian components

$$ \mathbb{J}_k \;=\; {}^{(1)}J_k \;+\; {}^{(2)}J_k \quad ;\; k \;=\; 1,2,3\;. \tag{6.59} $$

We seek to determine states $|J, M(\ell_1,\ell_2)\rangle$ which are simultaneous eigenstates of the operators $\mathbb{J}^2$, $\mathbb{J}_3$, ${}^{(1)}J^2$, ${}^{(2)}J^2$ which, as usual are denoted by their respective quantum numbers J, M, ℓ_1, ℓ_2, i.e., for such states should hold

$$ \mathbb{J}^2|J, M(\ell_1,\ell_2)\rangle \;=\; J(J+1)\,|J, M(\ell_1,\ell_2)\rangle \tag{6.60} $$

$$ \mathbb{J}_3|J, M(\ell_1,\ell_2)\rangle \;=\; M\,|J, M(\ell_1,\ell_2)\rangle \tag{6.61} $$

$$ {}^{(1)}J^2|J, M(\ell_1,\ell_2)\rangle \;=\; \ell_1(\ell_1+1)\,|J, M(\ell_1,\ell_2)\rangle \tag{6.62} $$

$$ {}^{(2)}J^2|J, M(\ell_1,\ell_2)\rangle \;=\; \ell_2(\ell_2+1)\,|J, M(\ell_1,\ell_2)\rangle\;. \tag{6.63} $$

At this point, we like to recall for future reference that the operators ${}^{(1)}J^2, {}^{(2)}J^2$, according to (5.300), can be expressed in terms of the number operators

$$ \hat{k}_1 \;=\; \frac{1}{2}\left(a_+^\dagger a_+ \;+\; a_-^\dagger a_-\right) \quad , \quad \hat{k}_2 \;=\; \frac{1}{2}\left(b_+^\dagger b_+ \;+\; b_-^\dagger b_-\right)\;, \tag{6.64} $$

namely,

$$ {}^{(j)}J^2 \;=\; \hat{k}_j\,(\,\hat{k}_j \;+\; 1\,) \quad , j \;=\; 1,\,2\;. \tag{6.65} $$

For the operators $\hat{k}_j$ holds

$$ \hat{k}_j\,|\ell_j, m_j\rangle_j \;=\; \ell_j\,|\ell_j, m_j\rangle_j \tag{6.66} $$

and, hence,

$$ \hat{k}_j\,|J, M(\ell_1,\ell_2)\rangle \;=\; \ell_j\,|J, M(\ell_1,\ell_2)\rangle \tag{6.67} $$

We will also require below the raising and lowering operators associated with the total angular momentum operator (6.59)

$$ \mathbb{J}_\pm \;=\; \mathbb{J}_1 \;\pm\; i\,\mathbb{J}_2\;. \tag{6.68} $$

The states $|J, M(\ell_1,\ell_2)\rangle$ can be expressed in terms of Clebsch-Gordan coefficients (6.18) as follows

$$ |J, M(\ell_1,\ell_2)\rangle \;=\; \textstyle\sum_{m_1,m_2}|\ell_1, m_1\rangle_1|\ell_2, m_2\rangle_2\,(\ell_1, m_1, \ell_2, m_2|J, M(\ell_1,\ell_2)) \quad , $$

$$ |\ell_1 \;-\; \ell_2| \;\leq\; J \;\leq\; \ell_1 \;+\; \ell_2 \quad , -J \;\leq\; M \;\leq\; J\;. \tag{6.69} $$

The aim of the present Section is to determine closed expressions for the Clebsch–Gordan coefficents $(\ell_1, m_1, \ell_2, m_2|J, M(\ell_1,\ell_2))$.

The Operator $K^\dagger$

The following operator

$$K^\dagger \;=\; a_+^\dagger\, b_-^\dagger \;-\; a_-^\dagger\, b_+^\dagger \;. \tag{6.70}$$

will play a crucial role in the evaluation of the Clebsch-Gordan-Coefficients. This operator obeys the following commutation relationships with the other pertinent angular momentum / spin operators

$$\left[\hat{k}_j,\, K^\dagger\right] \;=\; \frac{1}{2}\, K^\dagger \quad ,\; j \,=\, 1,\, 2 \tag{6.71}$$

$$\left[{}^{(j)}J^2,\, K^\dagger\right] \;=\; K^\dagger\, \hat{k}_j \;+\; \frac{3}{4}\, K^\dagger \quad ,\; j \,=\, 1,\, 2 \tag{6.72}$$

$$\left[\mathbb{J}_3,\, K^\dagger\right] \;=\; 0 \tag{6.73}$$

$$\left[\mathbb{J}_\pm,\, K^\dagger\right] \;=\; 0 \;. \tag{6.74}$$

We note that, due to $\mathbb{J}^2 \,=\, \frac{1}{2}\mathbb{J}_+\mathbb{J}_- \,+\, \frac{1}{2}\mathbb{J}_-\mathbb{J}_+ \,+\, \mathbb{J}_3^2$, the relationships (6.73, 6.74) imply

$$\left[\mathbb{J}^2,\, K^\dagger\right] \;=\; 0 \;. \tag{6.75}$$

The relationships (6.71–6.73) can be readily proven. For example, using (6.64, 6.50, 6.51) one obtains

$$\begin{aligned}
\left[\hat{k}_1,\, K^\dagger\right] \;&=\; \frac{1}{2}\, \left[a_+^\dagger a_+ \,+\, a_-^\dagger a_-,\, a_+^\dagger b_-^\dagger \,-\, a_-^\dagger b_+^\dagger\right] \\
&=\; \frac{1}{2}\, a_+^\dagger\, \left[a_+,\, a_+^\dagger\right]\, b_-^\dagger \,-\, \frac{1}{2}\, a_-^\dagger\, \left[a_-,\, a_-^\dagger\right]\, b_+^\dagger \\
&=\; \frac{1}{2}\, \left(a_+^\dagger b_-^\dagger \,-\, a_-^\dagger b_+^\dagger\right) \;=\; \frac{1}{2}\, K^\dagger \;.
\end{aligned}$$

A similar calculation yields $[\hat{k}_2, K^\dagger] \,=\, \frac{1}{2} K^\dagger$. Employing (6.65) and (6.71) one can show

$$\begin{aligned}
\left[{}^{(j)}J^2,\, K^\dagger\right] \;&=\; \frac{1}{2}\, \left[\hat{k}_j(\hat{k}_j + 1),\, K^\dagger\right] \\
&=\; \frac{1}{2}\, \hat{k}_j\, \left[\hat{k}_j + 1,\, K^\dagger\right] \;+\; \frac{1}{2}\, \left[\hat{k}_j,\, K^\dagger\right]\, (\hat{k}_j + 1) \\
&=\; \frac{1}{2}\, \hat{k}_j\, K^\dagger \;+\; \frac{1}{2}\, K^\dagger\, (\hat{k}_j + 1) \\
&=\; K^\dagger\, \hat{k}_j \;+\; \frac{1}{2}\, \left[\hat{k}_j,\, K^\dagger\right] \;+\; \frac{1}{2}\, K^\dagger \\
&=\; K^\dagger\, \hat{k}_j \;+\; \frac{3}{4}\, K^\dagger \;.
\end{aligned}$$

Using $\mathbb{J}_3 \,=\, {}^{(1)}J_3 \,+\, {}^{(2)}J_3$, expressing ${}^{(k)}J_3$ through the creation and annihilation operators according to (5.288), and applying the relationships (6.71–6.73) yields

$$\begin{aligned}
\left[\mathbb{J}_3, K^\dagger\right] \;&=\; \frac{1}{2}\, \left[a_+^\dagger a_+ \,-\, a_-^\dagger a_- \,+\, b_+^\dagger b_+ \,-\, b_-^\dagger b_-,\, a_+^\dagger b_-^\dagger \,-\, a_-^\dagger b_+^\dagger\right] \\
&=\; \frac{1}{2}\, a_+^\dagger\, \left[a_+,\, a_+^\dagger\right]\, b_-^\dagger \;+\; \frac{1}{2}\, a_-^\dagger\, \left[a_-,\, a_-^\dagger\right]\, b_+^\dagger \\
&-\; \frac{1}{2}\, b_+^\dagger a_-^\dagger\, \left[b_+, b_+^\dagger\right] \;-\; \frac{1}{2}\, b_-^\dagger a_+^\dagger\, \left[b_-, b_-^\dagger\right] \;=\; 0 \;.
\end{aligned}$$

Starting from (5.292) one can derive similarly

$$\begin{aligned}\left[\mathbb{J}_+, K^\dagger\right] &= \left[a_+^\dagger a_- + b_+^\dagger b_-,\, a_+^\dagger b_- - a_-^\dagger b_+^\dagger\right] \\ &= -a_+^\dagger \left[a_-,\, a_-^\dagger\right] b_+^\dagger + b_+^\dagger a_+^\dagger \left[b_-,\, b_-^\dagger\right] = 0\,.\end{aligned}$$

The property $[\mathbb{J}_-,\, K^\dagger] = 0$ is demonstrated in an analoguous way.

Action of $K^\dagger$ on the states $|J, M(\ell_1, \ell_2)\rangle$

We want to demonstrate now that the action of $K^\dagger$ on the states $|J, M(\ell_1, \ell_2)\rangle$ produces again total angular momentum eigenstates to the same J and M quantum numbers of $\mathbb{J}^2$ and $\mathbb{J}_3$, but for different ℓ_1 and ℓ_2 quantum numbers of the operators ${}^{(1)}J^2$ and ${}^{(2)}J^2$.
The commutation properties (6.73, 6.75) ascertain that under the action of $K^\dagger$ the states $|J, M(\ell_1, \ell_2)\rangle$ remain eigenstates of $\mathbb{J}^2$ and $\mathbb{J}_3$ with the same quantum numbers. To demonstrate that the resulting states are eigenstates of ${}^{(1)}J^2$ and ${}^{(2)}J^2$ we exploit (6.72) and (6.62, 6.63, 6.67)

$$\begin{aligned}{}^{(j)}J^2 K^\dagger |J, M(\ell_1, \ell_2)\rangle &= \left(\left[{}^{(j)}J^2,\, K^\dagger\right] + K^\dagger\, {}^{(j)}J^2\right) |J, M(\ell_1, \ell_2)\rangle \\ &= \left(K^\dagger \hat{k}_j + \frac{3}{4} K^\dagger + K^\dagger \ell_j\,(\ell_j + 1)\right) |J, M(\ell_1, \ell_2)\rangle \\ &= K^\dagger \left(\ell_j + \frac{3}{4} + \ell_j\,(\ell_j + 1)\right) |J, M(\ell_1, \ell_2)\rangle \\ &= \left(\ell_j + \frac{1}{2}\right)\left(\ell_j + \frac{3}{2}\right) K^\dagger |J, M(\ell_1, \ell_2)\rangle\,.\end{aligned}$$

However, this result implies that $K^\dagger|J, M(\ell_1, \ell_2)\rangle$ is a state with quantum numbers $\ell_1 + \frac{1}{2}$ and $\ell_2 + \frac{1}{2}$, i.e., it holds

$$K^\dagger |J, M(\ell_1, \ell_2)\rangle = N |J, M(\ell_1 + \frac{1}{2}, \ell_2 + \frac{1}{2})\rangle\,. \tag{6.76}$$

Here N is an unknown normalization constant.
One can generalize property (6.76) and state

$$\left(K^\dagger\right)^n |J, M(\ell_1, \ell_2)\rangle = N' |J, M(\ell_1 + \frac{n}{2}, \ell_2 + \frac{n}{2})\rangle \tag{6.77}$$

where N' is another normalization constant. We consider now the case that $(K^\dagger)^n$ acts on the simplest total angular momentum / spin state, namely, on the state

$$|j_1 + j_2, j_1 + j_2(j_1, j_2)\rangle = |j_1, j_1\rangle_1\, |j_2, j_2\rangle_2\,, \tag{6.78}$$

a state which has been used already in the construction of Clebsch-Gordan coefficients in Section 6.2. Application of $(K^\dagger)^n$ to this state yields, according to (6.77),

$$\begin{aligned}|j_1 + j_2, j_1 + j_2 &\quad (j_1 + \frac{n}{2}, j_2 + \frac{n}{2})\rangle \\ &= N(n, j_1 1 + \frac{n}{2}, j_1 + \frac{n}{2}) \left(K^\dagger\right)^n |j_1 + j_2, j_1 + j_2(j_1, j_2)\rangle\end{aligned} \tag{6.79}$$

where we denoted the associated normalization constant by $N(n, j_1 + \frac{n}{2}, j_2 + \frac{n}{2})$.
It is now important to notice that any state of the type $|J, J(\ell_1, \ell_2)\rangle$ can be expressed through the r.h.s. of (6.79). For this purpose one needs to choose in (6.79) n, j_1, j_2 as follows

$$J \;=\; j_1 \,+\, j_2 \quad , \ell_1 \;=\; j_1 \,+\, \frac{n}{2} \quad , \ell_2 \;=\; j_2 \,+\, \frac{n}{2} \tag{6.80}$$

which is equivalent to

$$\begin{aligned} n \;&=\; \ell_1 \,+\, \ell_2 \,-\, J \\ j_1 \;&=\; \ell_1 \,-\, \frac{n}{2} \;=\; \frac{1}{2}\,(J \,+\, \ell_1 \,-\, \ell_2) \\ j_2 \;&=\; \ell_2 \,-\, \frac{n}{2} \;=\; \frac{1}{2}\,(J \,+\, \ell_1 \,-\, \ell_2)\ . \end{aligned} \tag{6.81}$$

Accodingly, holds

$$\begin{aligned} |J, J(\ell_1, \ell_2)\rangle \;&=\; N(\ell_1 + \ell_2 - J, \ell_1, \ell_2)\,\left(K^\dagger\right)^{\ell_1+\ell_2-J} \;\times \\ &\times\, |\frac{1}{2}\,(J \,+\, \ell_1 \,-\, \ell_2)\,, \frac{1}{2}\,(J \,+\, \ell_1 \,-\, \ell_2)\rangle_1 \;\times \\ &\times\, |\frac{1}{2}\,(J \,+\, \ell_2 \,-\, \ell_1)\,, \frac{1}{2}\,(J \,+\, \ell_2 \,-\, \ell_1)\rangle_2\ . \end{aligned} \tag{6.82}$$

The normalization constant appearing here is actually

$$N(\ell_1 + \ell_2 - J, \ell_1, \ell_2) \;=\; \left[\frac{(2J \,+\, 1)!}{(\ell_1 \,+\, \ell_2 \,-\, J)!\,(\ell_1 \,+\, \ell_2 \,+\, J \,+\, 1)!}\right]^{\frac{1}{2}}\ . \tag{6.83}$$

The derivation of this expression will be provided further below (see page 158 ff).

Strategy for Generating the States $|J, M(\ell_1, \ell_2)\rangle$

Our construction of the states $|J, M(\ell_1, \ell_2)\rangle$ exploits the expression (6.82) for $|J, J(\ell_1, \ell_2)\rangle$. The latter states, in analogy to the construction (5.104, 5.105) of the spherical harmonics, allow one to obtain the states $|J, M(\ell_1, \ell_2)\rangle$ for $-J \,\leq\, M \,\leq\, J$ as follows

$$|J, M(\ell_1, \ell_2)\rangle \;=\; \Delta(J, M)\ (\mathbb{J}_-)^{J-M}\ |J, J(\ell_1, \ell_2)\rangle \tag{6.84}$$

$$\Delta(J, M) \;=\; \left[\frac{(J \,+\, M)!}{(2J)!\,(J \,-\, M)!}\right]^{\frac{1}{2}}\ . \tag{6.85}$$

Combining (6.84) with (6.82, 6.57) and exploiting the fact that $\mathbb{J}_-$ and $K^\dagger$ commute [c.f. (6.74)] yields

$$\begin{aligned} |J, M(\ell_1, \ell_2)\rangle \;&=\; \frac{N(\ell_1 + \ell_2 - J, \ell_1, \ell_2)\,\Delta(J, M)}{\sqrt{(J + \ell_1 - \ell_2)!\,(J + \ell_2 - \ell_1)!}} \;\times \\ &\times \left(K^\dagger\right)^{\ell_1+\ell_2-J} (\mathbb{J}_-)^{J-M} \left(a_+^\dagger\right)^{J+\ell_1-\ell_2} \left(b_+^\dagger\right)^{J+\ell_2-\ell_1}\ |\Psi_o\rangle \end{aligned} \tag{6.86}$$

Our strategy for the evaluation of the Clebsch-Gordan-coefficients is to expand (6.86) in terms of monomials

$$\left(a_+^\dagger\right)^{\ell_1+m_1}\left(a_-^\dagger\right)^{\ell_1-m_1}\left(b_+^\dagger\right)^{\ell_2+m_2}\left(b_-^\dagger\right)^{\ell_2-m_2}|\Psi_o\rangle\;, \tag{6.87}$$

i.e., in terms of $|\ell_1, m_1\rangle_1|\ell_2, m_2\rangle_2$ [cf. (6.57)]. Comparision with (6.69) yields then the Clebsch-Gordan-coefficients.

Expansion of an Intermediate State

We first consider the expansion of the following factor appearing in (6.86)

$$|G_{st}^r\rangle \;=\; \mathbb{J}_-^r\left(a_+^\dagger\right)^s\left(b_+^\dagger\right)^t|\Psi_o\rangle \tag{6.88}$$

in terms of monomials (6.87). For this purpose we introduce the generating function

$$I(\lambda,x,y) \;=\; \exp\left(\lambda\mathbb{J}_-\right)\exp\left(x\,a_+^\dagger\right)\exp\left(x\,b_+^\dagger\right)|\Psi_o\rangle\;. \tag{6.89}$$

Taylor expansion of the two exponential operators yields immediately

$$I(\lambda,x,y) \;=\; \sum_{r,s,t}\frac{\lambda^r x^s y^t}{r!s!t!}\,|G_{st}^r\rangle\;, \tag{6.90}$$

i.e., $I(\lambda,x,y)$ is a generating function for the states $|G_{st}^r\rangle$ defined in (6.88).
The desired expansion of $|G_{st}^r\rangle$ can be obtained from an alternate evaluation of $I(\lambda,x,y)$ which is based on the properties

$$a_\zeta f(a_\zeta^\dagger)\,|\Psi_o\rangle \;=\; \frac{\partial}{\partial a_\zeta^\dagger}f(a_\zeta^\dagger)\,|\Psi_o\rangle \qquad b_\zeta f(b_\zeta^\dagger)\,|\Psi_o\rangle \;=\; \frac{\partial}{\partial b_\zeta^\dagger}f(b_\zeta^\dagger)\,|\Psi_o\rangle \tag{6.91}$$

which, in analogy to (5.264), follows from the commutation properties (6.50–6.52). One obtains then using

$$\mathbb{J}_- \;=\; a_-^\dagger a_+ \;+\; b_-^\dagger b_+ \tag{6.92}$$

and noting $[a_+,\,a_-^\dagger] \,=\, [b_+,\,b_-^\dagger] \,=\, 0$ [cf. (6.50)]

$$\begin{aligned}
\exp\left(\,\lambda\mathbb{J}_-\,\right)\,f(a_+^\dagger)\,g(b_+^\dagger)\,|\Psi_o\rangle \;&=\; \exp\left(\,a_-^\dagger a_+\,\right)\,f(a_+^\dagger)\exp\left(\,b^\dagger b_+\,\right)\,g(b_+^\dagger)\,|\Psi_o\rangle \\
&=\; \sum_u\frac{\lambda^u}{u!}\left(a_-^\dagger a_+\right)^u f(a_+^\dagger)\;\times \\
&\qquad\times\sum_v\frac{\lambda^v}{v!}\left(b_-^\dagger b_+\right)^v g(b_+^\dagger)\,|\Psi_o\rangle \\
&=\; \sum_u\frac{\left(\lambda\,a_-^\dagger\right)^u}{u!}\left(\frac{\partial}{\partial a_+^\dagger}\right)^u f(a_+^\dagger)\;\times \\
&\qquad\times\sum_v\frac{\left(\lambda\,b_-^\dagger\right)^v}{v!}\left(\frac{\partial}{\partial b_+^\dagger}\right)^v g(b_+^\dagger)\,|\Psi_o\rangle \\
&=\; f(a_+^\dagger\,+\,\lambda\,a_-^\dagger)\,g(b_+^\dagger\,+\,\lambda\,b_-^\dagger)\,|\Psi_o\rangle\;.
\end{aligned} \tag{6.93}$$

We conclude

$$I(\lambda, x, y) \;=\; \exp\left(a_+^\dagger \,+\, \lambda\, a_-^\dagger \right) \exp\left(b_+^\dagger \,+\, \lambda\, b_-^\dagger \right) |\Psi_o\rangle \,. \tag{6.94}$$

One can infer from this result the desired expressions for $|G_{st}^r\rangle$. Expanding the exponentials in (6.94) yields

$$\begin{aligned} I(\lambda, x, y) \;&=\; \sum_{s,t} \frac{x^s}{s!}\,\frac{y^t}{t!} \left(a_+^\dagger \,+\, \lambda\, a_-^\dagger \right)^s \left(b_+^\dagger \,+\, \lambda\, b_-^\dagger \right)^t |\Psi_o\rangle \\ &=\; \sum_{s,t} \frac{x^s}{s!}\,\frac{y^t}{t!} \sum_t \sum_v \left(\begin{array}{c} s \\ u \end{array} \right) \left(\begin{array}{c} t \\ v \end{array} \right) \times \\ &\qquad \times \left(a_+^\dagger \right)^{s-u} \lambda^u \left(a_-^\dagger \right)^u \left(b_+^\dagger \right)^{t-v} \lambda^v \left(b_-^\dagger \right)^v |\Psi_o\rangle \\ &=\; \sum_{r,s,t} \frac{\lambda^r}{r!}\,\frac{x^s}{s!}\,\frac{y^t}{t!} \sum_q r! \left(\begin{array}{c} s \\ q \end{array} \right) \left(\begin{array}{c} t \\ r-q \end{array} \right) \times \\ &\qquad \times \left(a_+^\dagger \right)^{s-q} \left(a_-^\dagger \right)^q \left(b_+^\dagger \right)^{t-r+q} \left(b_-^\dagger \right)^{r-q} |\Psi_o\rangle \end{aligned} \tag{6.95}$$

Comparision with (6.90) allows one to infer

$$\begin{aligned} |G_{s,t}^r\rangle \;&=\; \sum_q r! \left(\begin{array}{c} s \\ q \end{array} \right) \left(\begin{array}{c} t \\ r-q \end{array} \right) \times \\ &\qquad \times \left(a_+^\dagger \right)^{s-q} \left(a_-^\dagger \right)^q \left(b_+^\dagger \right)^{t-r+q} \left(b_-^\dagger \right)^{r-q} |\Psi_o\rangle \end{aligned} \tag{6.96}$$

and, using the definition (6.88), one can write the right factor in (6.86)

$$\begin{aligned} &(\mathbb{J}_-)^{J-M} \left(a_+^\dagger \right)^{J+\ell_1-\ell_2} \left(b_+^\dagger \right)^{J+\ell_2-\ell_1} |\Psi_o\rangle \qquad (6.97) \\ &= \sum_q \frac{(J-M)!(J+\ell_1-\ell_2)!(J+\ell_2-\ell_1)!}{q!(J+\ell_1-\ell_2-q)!(J-M-q)!(M+\ell_2-\ell_1+q)!} \;\times \\ &\times \left(a_+^\dagger \right)^{J+\ell_1-\ell_2-q} \left(a_-^\dagger \right)^q \left(b_+^\dagger \right)^{M+\ell_2-\ell_1+q} \left(b_-^\dagger \right)^{J-M-q} |\Psi_o\rangle \end{aligned}$$

Final Result

Our last step is to apply the operator $(K^\dagger)^{\ell_1+\ell_2-J}$ to expression (6.97), to obtain the desired expansion of $|J, M(\ell_1, \ell_2)\rangle$ in terms of states $|\ell_1, m_1\rangle_1 |\ell_2, m_2\rangle_2$. With $K^\dagger$ given by (6.70) holds

$$\begin{aligned} \left(K^\dagger \right)^{\ell_1+\ell_2-J} \;&=\; \sum_s \left(\begin{array}{c} \ell_1+\ell_2-J \\ s \end{array} \right) (-1)^s \\ &\qquad \left(a_+^\dagger \right)^{\ell_1+\ell_2-J-s} \left(b_-^\dagger \right)^{\ell_1+\ell_2-J-s} \left(a_-^\dagger \right)^s \left(b_+^\dagger \right)^s \,. \end{aligned} \tag{6.98}$$

Operation of this operator on (6.97) yields, using the commutation property (6.50),

$$|J, M(\ell_1, \ell_2)\rangle \;=\; \sum_{s,q} (-1)^s \tag{6.99}$$

$$\frac{(\ell_1+\ell_2-J)!(J-M)!(J+\ell_1-\ell_2)!(J+\ell_2-\ell_1)!}{s!(\ell_1+\ell_2-J-s)!q!(J+\ell_1-\ell_2-q)!(J-M-q)!(M+\ell_2-\ell_1+q)!}$$
$$\left(a_+^\dagger\right)^{2\ell_1-q-s}\left(a_-^\dagger\right)^{q+s}\left(b_+^\dagger\right)^{M+\ell_2-\ell_1+q+s}\left(b_-^\dagger\right)^{\ell_1+\ell_2-M-q-s}|\Psi_o\rangle$$

The relationships (6.53,6.54) between creation operator monomials and angular momentum states allow one to write this

$$\begin{aligned}|J,M(\ell_1,\ell_2)\rangle \;=\; & \frac{N(\ell_1+\ell_2-J,\ell_1,\ell_2)\,\Delta(J,M)}{\sqrt{(J+\ell_1-\ell_2)!\,(J+\ell_2-\ell_1)!}}\sum_{s,q}(-1)^s\;\times \\ \times & \frac{(\ell_1+\ell_2-J)!(J-M)!(J+\ell_1-\ell_2)!(J+\ell_2-\ell_1)!}{s!(\ell_1+\ell_2-J-s)!q!(J+\ell_1-\ell_2-q)!(J-M-q)!(M+\ell_2-\ell_1+q)!} \\ \times & \sqrt{(2\ell_1-q-s)!(q+s)!} \\ \times & \sqrt{(M+\ell_2-\ell_1+q+s)!(\ell_1+\ell_2-M-q-s)!} \\ \times & |\ell_1,\ell_1-q-s\rangle_1\,|\ell_2,M-\ell_1+q+s\rangle_2\end{aligned} \tag{6.100}$$

One can conclude that this expression reproduces (6.69) if one identifies

$$m_1 \;=\; \ell_1\;-q\;-\;s\quad,\; m_2 \;=\; M\;-\;\ell_1\;+\;q\;+\;s\,. \tag{6.101}$$

Note that $m_1 + m_2 = M$ holds. The summation over q corresponds then to the summation over m_1, m_2 in (6.69) since, according to (6.101), $q = \ell_1 - m_1 - s$ and $m_2 = M - m_1$. The Clebsch-Gordan coefficents are then finally

$$\begin{aligned}&(\ell_1,m_1,\ell_2,m_2|J,M) \;= \\ &\sqrt{2J+1}\,\left[\frac{\ell_1+\ell_2-J)!(\ell_1-\ell_2+J)!(-\ell_1+\ell_2+J)!}{(\ell_1+\ell_2+J+1)!}\right]^{\frac{1}{2}} \\ &\times\;[(\ell_1+m_1)!(\ell_1-m_1)!(\ell_2+m_2)!(\ell_2-m_2)!(J+M)!(J-M)!]^{\frac{1}{2}} \\ &\times\sum_s\;\frac{(-1)^s}{s!(\ell_1-m_1-s)!(\ell_2+m_2-s)!} \\ &\qquad\times\;\frac{1}{(\ell_1+\ell_2-J-s)!(J-\ell_1-m_2+s)!(J-\ell_2+m_1+s)!}\end{aligned} \tag{6.102}$$

The Normalization

We want to determine now the expression (6.83) of the normalization constant $N(\ell_1+\ell_2-J,\ell_1,\ell_2)$ defined through (6.82). For this purpose we introduce

$$j_1 \;=\; \frac{1}{2}(J\;+\;\ell_1\;-\;\ell_2)\;,\; j_2 \;=\; \frac{1}{2}(J\;+\;\ell_2\;-\;\ell_1)\;,\; n \;=\; \ell_1\;+\;\ell_2\;-\;J\;. \tag{6.103}$$

To determine $N = N(\ell_1+\ell_2-J,\ell_1,\ell_2)$ we consider the scalar product $\langle J,J(\ell_1,\ell_2)|J,J(\ell_1,\ell_2)\rangle = 1$. Using (6.82) and (6.103) this can be written

$$1 \;=\; N^2\,\langle\psi(j_1,j_2,n)|\psi(j_1,j_2,n)\rangle \tag{6.104}$$

where

$$|\psi(j_1,j_2,n)\rangle \;=\; \left(K^\dagger\right)^n |j_1,j_1\rangle_1 |j_2,j_2\rangle_2 \;. \tag{6.105}$$

The first step of our calculation is the expansion of $\psi(j_1,j_2,n)$ in terms of states $|j_1',m_1\rangle_1|j_2',m_2\rangle_2$. We employ the expression (6.57) for these states and the expression (6.70) for the operator $K^\dagger$. Accordingly, we obtain

$$\begin{aligned}
|\psi(j_1,j_2,n)\rangle \;&=\; \frac{1}{\sqrt{(2j_1)!(2j_2)!}} \sum_s \left(\begin{array}{c} n \\ s \end{array}\right) \left(a_+^\dagger b_-^\dagger\right)^{n-s} (-1)^s \left(a_-^\dagger b_+^\dagger\right)^s \\
&\qquad \left(a_+^\dagger\right)^{2j_1} \left(b_+^\dagger\right)^{2j_2} |\Psi_o\rangle \;= \\
&\frac{n!}{\sqrt{(2j_1)!(2j_2)!}} \sum_s \frac{(-1)^s \sqrt{(2j_1+n-s)!s!(2j_2+s)!(n-s)!}}{s!(n-s)!} \\
&\frac{\left(a_+^\dagger\right)^{2j_1-n-s} \left(a_-^\dagger\right)^s \left(b_+^\dagger\right)^{2j_2+s} \left(b_-^\dagger\right)^{n-s}}{\sqrt{(2j_1+n-s)!s!(2j_2+s)!(n-s)!}} |\Psi_o\rangle \;.
\end{aligned} \tag{6.106}$$

The orthonormality of the states occurring in the last expression allows one to write (6.104)

$$\begin{aligned}
1 \;&=\; N^2 \frac{(n!)^2}{(2j_1)!(2j_2)!} \sum_s \frac{(2j_1+n-s)!(2j_2+s)!}{s!(n-s)!} \\
&=\; (n!)^2 \sum_s \left(\begin{array}{c} 2j_1+n-s \\ 2j_1 \end{array}\right) \left(\begin{array}{c} 2j_2+s \\ 2j_2 \end{array}\right)
\end{aligned} \tag{6.107}$$

The latter sum can be evaluated using

$$\left(\frac{1}{1-\lambda}\right)^{n_1+1} \;=\; \sum_{m_1} \left(\begin{array}{c} n_1+m_1 \\ n_1 \end{array}\right) \lambda^{m_1} \tag{6.108}$$

a property which follows from

$$\frac{\partial^\nu}{\partial\lambda^\nu} \left(\frac{1}{1-\lambda}\right)^{n_1+1}\Bigg|_{\lambda=0} \;=\; \frac{(n_1+\nu)!}{n_1!} \tag{6.109}$$

and Taylor expansion of the left hand side of (6.108). One obtains then, applying (6.108) twice,

$$\left(\frac{1}{1-\lambda}\right)^{n_1+1} \left(\frac{1}{1-\lambda}\right)^{n_2+1} \;=\; \sum_{m_1,m_2} \left(\begin{array}{c} n_1+m_1 \\ n_1 \end{array}\right) \left(\begin{array}{c} n_2+m_2 \\ n_2 \end{array}\right) \lambda^{m_1+m_2} \tag{6.110}$$

which can be written

$$\left(\frac{1}{1-\lambda}\right)^{n_1+n_2+2} \;=\; \sum_r \left[\sum_s \left(\begin{array}{c} n_1+r-s \\ n_1 \end{array}\right) \left(\begin{array}{c} n_2+s \\ n_2 \end{array}\right)\right] \lambda^r \tag{6.111}$$

Comparision with (6.108) yields the identity

$$\sum_s \left(\begin{array}{c} n_1+r-s \\ n_1 \end{array}\right) \left(\begin{array}{c} n_2+s \\ n_2 \end{array}\right) \;=\; \left(\begin{array}{c} n_1+n_2+r+1 \\ n_1+n_2+1 \end{array}\right) \;. \tag{6.112}$$

Applying this to (6.107) yields

$$\begin{aligned} 1 &= N^2\,(n!)^2 \begin{pmatrix} 2j_1+2j_2+n+1 \\ 2j_1+2j_2+1 \end{pmatrix} \\ &= N^2\,\frac{n!(2j_1+2j_2+n+1)!}{(2j_1+2j_2+1)!}\ . \end{aligned} \tag{6.113}$$

Using the identities (6.103) one obtains the desired result (6.83).

6.4 Symmetries of the Clebsch-Gordan Coefficients

The Clebsch-Gordan coefficients obey symmetry properties which reflect geometrical aspects of the operator relationship (6.11)

$$\vec{\mathbb{J}} \;=\; \vec{\mathcal{J}}^{(1)} \;+\; \vec{\mathcal{J}}^{(2)} \quad . \tag{6.114}$$

For example, interchanging the operators $\vec{\mathcal{J}}^{(1)}$ and $\vec{\mathcal{J}}^{(2)}$ results in

$$\vec{\mathbb{J}} \;=\; \vec{\mathcal{J}}^{(2)} \;+\; \vec{\mathcal{J}}^{(1)} \quad . \tag{6.115}$$

This relationship is a trivial consequence of (6.114) as long as $\vec{\mathbb{J}}$, $\vec{\mathcal{J}}^{(1)}$, and $\vec{\mathcal{J}}^{(2)}$ are vectors in $\mathbb{R}^{\not\#}$. For the quantum mechanical addition of angular momenta the Clebsch Gordan coefficients $(\ell_1, m_1, \ell_2, m_2|J, M)$ corresponding to (6.114) show a simple relationship to the Clebsch Gordan coefficients $(\ell_2, m_2, \ell_1, m_1|J, M)$ corresponding to (6.115), namely,

$$(\ell_1, m_1, \ell_2, m_2|J, M) \;=\; (-1)^{\ell_1+\ell_2-J}\,(\ell_2, m_2, \ell_1, m_1|J, M)\ . \tag{6.116}$$

If one takes the negatives of the operators in (6.114) one obtains

$$-\,\vec{\mathbb{J}} \;=\; -\,\vec{\mathcal{J}}^{(1)} \;-\; \vec{\mathcal{J}}^{(2)} \quad . \tag{6.117}$$

The respective Clebsch-Gordan coefficients $(\ell_1, -m_1, \ell_2, -m_1 2|J, -M)$ are again related in a simple manner to the coefficients $(\ell_1, m_1, \ell_2, m_2|J, M)$

$$(\ell_1, m_1, \ell_2, m_2|J, M) \;=\; (-1)^{\ell_1+\ell_2-J}\,(\ell_1, -m_1, \ell_2, -m_2|J, -M)\ . \tag{6.118}$$

Finally, one can interchange also the operator $\vec{\mathbb{J}}$ on the l.h.s. of (6.114) by, e.g., $\vec{\mathcal{J}}^{(1)}$ on the r.h.s. of this equation

$$\vec{\mathcal{J}}^{(1)} \;=\; \vec{\mathcal{J}}^{(2)} \;-\; \vec{\mathbb{J}} \quad . \tag{6.119}$$

The corresponding symmetry property of the Clebsch-Gordan coefficients is

$$(\ell_1, m_1, \ell_2, m_2|J, M) \;=\; (-1)^{\ell_2+m_2}\,\sqrt{\frac{2J+1}{2\ell_1+1}}\,(\ell_2, -m_2, J, M|\ell_1, m_1)\ . \tag{6.120}$$

The symmetry properties (6.116), (6.118), and (6.120) can be readily derived from the expression (6.102) of the Clebsch-Gordan coefficients. We will demonstrate this now.
To derive relationship (6.116) one expresses the Clebsch-Gordan coefficient on the r.h.s. of (6.116) through formula (6.102) by replacing (ℓ_1, m_1) by (ℓ_2, m_2) and, vice versa, (ℓ_2, m_2) by (ℓ_1, m_1), and

seeks then to relate the resulting expression to the original expression (6.102) to prove identity with the l.h.s. Inspecting (6.102) one recognizes that only the sum

$$S(\ell_1, m_1, \ell_2, m_2|J, M) \;=\; \sum_s \frac{(-1)^s}{s!(\ell_1 - m_1 - s)!(\ell_2 + m_2 - s)!} \times \frac{1}{(\ell_1 + \ell_2 - J - s)!(J - \ell_1 - m_2 + s)!(J - \ell_2 + m_1 + s)!} \tag{6.121}$$

is affected by the change of quantum numbers, the factor in front of S being symmetric in (ℓ_1, m_1) and (ℓ_2, m_2). Correspondingly, (6.116) implies

$$S(\ell_1, m_1, \ell_2, m_2|J, M) \;=\; (-1)^{\ell_1+\ell_2-J}\, S(\ell_2, m_2, \ell_1, m_1|J, M)\;. \tag{6.122}$$

To prove this we note that S on the r.h.s. reads, according to (6.121),

$$S(\ell_2, m_2, \ell_1, m_1|J, M) \;=\; \sum_s \frac{(-1)^s}{s!(\ell_2 - m_2 - s)!(\ell_1 + m_1 - s)!} \times \frac{1}{(\ell_1 + \ell_2 - J - s)!(J - \ell_2 - m_1 + s)!(J - \ell_1 + m_2 + s)!}\;. \tag{6.123}$$

Introducing the new summation index

$$s' \;=\; \ell_1 \,+\, \ell_2 \,-\, J \,-\, s \tag{6.124}$$

and using the equivalent relationships

$$s \;=\; \ell_1 \,+\, \ell_2 \,-\, J \,-\, s'\,, \qquad -s \;=\; J \,-\, -\ell_1 \,-\, \ell_2 \,+\, s' \tag{6.125}$$

to express s in terms of s' in (6.123) one obtains

$$\begin{aligned} &S(\ell_2, m_2, \ell_1, m_1|J, M) \;= \\ &\quad (-1)^{\ell_1+\ell_2-J} \sum_{s'} \frac{(-1)^{-s'}}{(\ell_1 + \ell_2 - J - s')!(J - \ell_1 - m_2 + s')!(J - \ell_2 + m_1 + s')!} \\ &\qquad \times \frac{1}{s'!(\ell_1 - m_1 - s')!(\ell_2 + m_2 - s')!}\;. \end{aligned} \tag{6.126}$$

Now it holds that $\ell_1 + \ell_2 - J$ in (6.124) is an integer, irrespective of the individual quantum numbers ℓ_1, ℓ_2, J being integer or half-integer. This fact can best be verified by showing that the construction of the eigenstates of $(\vec{\mathcal{J}}^{(1)} \,+\, \vec{\mathcal{J}}^{(2)})^2$ and $(\vec{\mathcal{J}}^{(1)} \,+\, \vec{\mathcal{J}}^{(2)})_3$ in Sect. 6.2 does, in fact, imply this property. Since also s in (6.102) and, hence, in (6.122) is an integer, one can state that s', as defined in (6.124), is an integer and, accordingly, that

$$(-1)^{-\,s'} \;=\; (-1)^{s'} \tag{6.127}$$

holds in (6.126). Reordering the factorials in (6.126) to agree with the ordering in (6.121) leads one to conclude the property (6.122) and, hence, one has proven (6.116).

To prove (6.118) we note that in the expression (6.102) for the Clebsch-Gordan coefficients the prefactor of S, the latter defined in (6.121), is unaltered by the change $m_1,\ m_2,\ M \to -m_1,\ -m_2,\ -M$. Hence, (6.118) implies

$$S(\ell_1, m_1, \ell_2, m_2|J, M) \ = \ (-1)^{\ell_1+\ell_2-J}\, S(\ell_1, -m_1, \ell_2, -m_1 2|J, -M)\ . \tag{6.128}$$

We note that according to (6.121) holds

$$\begin{aligned} S(\ell_1, -m_1, \ell_2, -m_1 2|J, -M) \ &= \ \sum_s \ \frac{(-1)^s}{s!(\ell_1+m_1-s)!(\ell_2-m_2-s)!} \\ &\times \ \frac{1}{(\ell_1+\ell_2-J-s)!(J-\ell_1+m_2+s)!(J-\ell_2-m_1+s)!}\ . \end{aligned} \tag{6.129}$$

Introducing the new summation index s' as defined in (6.124) and using the relationships (6.125) to replace, in (6.129), s by s' one obtains

$$\begin{aligned} &S(\ell_1, -m_1, \ell_2, -m_2|J, -M) \ = \\ &\quad (-1)^{\ell_1+\ell_2-J} \sum_s \ \frac{(-1)^{-s'}}{(\ell_1+\ell_2-J-s')!(J-\ell_2+m_1+s')!(J-\ell_1-m_2+s')!} \\ &\qquad \times \ \frac{1}{s'!(\ell_2+m_2-s')!(\ell_1-m_1-s')!}\ . \end{aligned} \tag{6.130}$$

For reasons stated already above, (6.127) holds and after reordering of the factorials in (6.130) to agree with those in (6.121) one can conclude (6.128) and, hence, (6.118).
We want to prove finally the symmetry property (6.120). Following the strategy adopted in the proof of relationships (6.116) and (6.118) we note that in the expression (6.102) for the Clebsch-Gordan coefficients the prefactor of S, the latter defined in (6.121), is symmetric in the pairs of quantum numbers (ℓ_1, m_1), (ℓ_2, m_2) and (J, M), except for the factor $\sqrt{2J+1}$ which singles out J. However, in the relationship (6.120) this latter factor is already properly 'repaired' such that (6.120) implies

$$S(\ell_1, m_1, \ell_2, m_2|J, M) \ = \ (-1)^{\ell_2+m_2}\, S(\ell_2, -m_2, J, M|\ell_1, m_1)\ . \tag{6.131}$$

According to (6.121) holds

$$\begin{aligned} S(\ell_2, -m_2, J, M|\ell_1, m_1) \ &= \ \sum_s \ \frac{(-1)^s}{s!(\ell_2+m_2-s)!(J+M-s)!} \\ &\times \ \frac{1}{(\ell_2+J-\ell_1-s)!(\ell_1-\ell_2-M+s)!(\ell_1-J-m_2+s)!}\ . \end{aligned} \tag{6.132}$$

Introducing the new summation index

$$s' \ = \ \ell_2 \ + \ m_2 \ - \ s \tag{6.133}$$

and, using the equivalent relationships

$$s \ = \ \ell_2 \ + \ m_2 \ - \ s'\ , \qquad -s \ = \ -\ell_2 \ - \ m_2 \ + \ s' \tag{6.134}$$

to replace s by s' in (6.132), one obtains

$$\begin{aligned} S(\ell_1,-m_1,\ell_2,-m_1 2|J,-M) \;=\; & \\ (-1)^{\ell_2+m_2}\sum_s \frac{(-1)^{-s'}}{(\ell_2+m_2-s')!s'!(J-\ell_2+m_1+s')!} & \\ \times \frac{1}{(J-\ell_1-m_2+s')!(\ell_1-m_1-s')!(\ell_1+\ell_2-J-s')!}\,. & \end{aligned} \tag{6.135}$$

Again for the reasons stated above, (6.127) holds and after reordering of the factorials in (6.135) to agree with those in (6.121) one can conclude (6.131) and, hence, (6.120).

6.5 Example: Spin–Orbital Angular Momentum States

Relativistic quantum mechanics states that an electron moving in the Coulomb field of a nucleus experiences a coupling $\sim \vec{\mathcal{J}}\cdot\vec{S}$ between its angular momentum, described by the operator $\vec{\mathcal{J}}$ and wave functions $Y_{\ell m}(\hat{r})$, and its spin-$\frac{1}{2}$, described by the operator $\vec{S}$ and wave function $\chi_{\frac{1}{2}\pm\frac{1}{2}}$. As a result, the eigenstates of the electron are given by the eigenstates of the total angular momentum-spin states

$$\mathcal{Y}_{jm}(\ell,\tfrac{1}{2}|\hat{r}) \;=\; \sum_{m',\sigma}(\ell,m',\tfrac{1}{2},\sigma|j,m)\,Y_{\ell m'}(\hat{r})\,\chi_{\frac{1}{2}\sigma} \tag{6.136}$$

which have been defined in (6.18). The states are simultaneous eigenstates of $(\mathcal{J}^{(tot)})^2$, $\mathcal{J}_3^{(tot)}$, $\mathcal{J}^2$, and S^2 and, as we show below, also of the spin-orbit coupling term $\sim\vec{\mathcal{J}}\cdot\vec{S}$. Here $\mathcal{J}^{(tot)}$ is defined as

$$\vec{\mathcal{J}}^{(tot)} \;=\; \vec{\mathcal{J}} \;+\; \vec{S}\,. \tag{6.137}$$

Here we assume for $\vec{S}$ the same units as for $\vec{\mathcal{J}}$, namely, $\hbar$, i.e., we define

$$\vec{S} \;=\; \frac{\hbar}{2}\,\vec{\sigma} \tag{6.138}$$

rather than (5.223).

Two-Dimensional Vector Representation

One can consider the functions $\chi_{\frac{1}{2}\pm\frac{1}{2}}$ to be represented alternatively by the basis vectors of the space $\mathbb{C}^{\not\in}$

$$\chi_{\frac{1}{2}\frac{1}{2}} \;=\; \begin{pmatrix}1\\0\end{pmatrix}\,, \qquad \chi_{\frac{1}{2}-\frac{1}{2}} \;=\; \begin{pmatrix}0\\1\end{pmatrix}\,. \tag{6.139}$$

The states $\mathcal{Y}_{jm}(\ell,\frac{1}{2}|\hat{r})$, accordingly, can then also be expressed as two-dimensional vectors. Using

$$m' \;=\; m \;-\; \sigma\,; \qquad \sigma \;=\; \pm\tfrac{1}{2} \tag{6.140}$$

one obtains

$$\begin{aligned} \mathcal{Y}_{jm}(\ell,\tfrac{1}{2}|\hat{r}) \;\;=\;\; & (\ell,m-\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2}|j,m)\,Y_{\ell m-\frac{1}{2}}(\hat{r})\begin{pmatrix}1\\0\end{pmatrix} \\ & +\,(\ell,m+\tfrac{1}{2},-\tfrac{1}{2},\tfrac{1}{2}|j,m)\,Y_{\ell m+\frac{1}{2}}(\hat{r})\begin{pmatrix}0\\1\end{pmatrix} \end{aligned} \tag{6.141}$$

or

$$\mathcal{Y}_{jm}(\ell,\tfrac{1}{2}|\hat{r}) \;=\; \begin{pmatrix} (\ell, m-\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}|j,m)\, Y_{\ell m-\frac{1}{2}}(\hat{r}) \\ (\ell, m+\tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2}|j,m)\, Y_{\ell m+\frac{1}{2}}(\hat{r}) \end{pmatrix} . \tag{6.142}$$

In this expression the quantum numbers (ℓ, m') of the angular momentum state are integers. According to (6.140), m is then half-integer and so must be j. The triangle inequalities (6.220) state in the present case $|\ell - \frac{1}{2}| \leq j \leq \ell + \frac{1}{2}$ and, therefore, we conclude $j = \ell \pm \frac{1}{2}$ or, equivalently, $\ell = j \pm \frac{1}{2}$. The different Clebsch-Gordon coefficients in (6.141) have the values

$$(j-\tfrac{1}{2}, m-\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}|j,m) \;=\; \sqrt{\frac{j\,+\,m}{2j}} \tag{6.143}$$

$$(j-\tfrac{1}{2}, m+\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}|j,m) \;=\; \sqrt{\frac{j\,-\,m}{2j}} \tag{6.144}$$

$$(j+\tfrac{1}{2}, m-\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}|j,m) \;=\; -\sqrt{\frac{j\,-\,m\,+1}{2j\,+\,2}} \tag{6.145}$$

$$(j+\tfrac{1}{2}, m+\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}|j,m) \;=\; \sqrt{\frac{j\,+\,m\,+1}{2j\,+\,2}} \tag{6.146}$$

which will be derived below (see pp. 170). Accordingly, the spin-orbital angular momentum states (6.141, 6.142) are

$$\mathcal{Y}_{jm}(j-\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \;=\; \begin{pmatrix} \sqrt{\frac{j+m}{2j}}\, Y_{j-\frac{1}{2}m-\frac{1}{2}}(\hat{r}) \\ \sqrt{\frac{j-m}{2j}}\, Y_{j-\frac{1}{2}m+\frac{1}{2}}(\hat{r}) \end{pmatrix} \tag{6.147}$$

$$\mathcal{Y}_{jm}(j+\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \;=\; \begin{pmatrix} -\sqrt{\frac{j-m+1}{2j+2}}\, Y_{j+\frac{1}{2}m-\frac{1}{2}}(\hat{r}) \\ \sqrt{\frac{j+m+1}{2j+2}}\, Y_{j+\frac{1}{2}m+\frac{1}{2}}(\hat{r}) \end{pmatrix} . \tag{6.148}$$

Eigenvalues

For the states (6.147, 6.148) holds

$$(\vec{\mathcal{J}} \,+\, \vec{S})^2\, \mathcal{Y}_{jm}(j\mp\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \;=\; \hbar^2\, j(j\,+\,1)\, \mathcal{Y}_{jm}(j\mp\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \tag{6.149}$$

$$(\vec{\mathcal{J}} \,+\, \vec{S})_3\, \mathcal{Y}_{jm}(j\mp\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \;=\; \hbar\, m\, \mathcal{Y}_{jm}(j\mp\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \tag{6.150}$$

$$\mathcal{J}^2\, \mathcal{Y}_{jm}(j\mp\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \;=\; \hbar^2\, (j\,\mp\tfrac{1}{2})\, (\,j\mp\tfrac{1}{2}\,+\,1)\, \mathcal{Y}_{jm}(j\mp\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \tag{6.151}$$

$$S^2\, \mathcal{Y}_{jm}(j\mp\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \;=\; \frac{3}{4}\,\hbar^2\, \mathcal{Y}_{jm}(j\mp\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \tag{6.152}$$

Furthermore, using

$$(\mathcal{J}^{(tot)})^2 \;=\; (\,\vec{\mathcal{J}} \,+\, \vec{S}\,)^2 \;=\; \mathcal{J}^2 \,+\, S^2 \,+\, 2\vec{\mathcal{J}}\cdot\vec{S} \tag{6.153}$$

or, equivalently,

$$2\vec{\mathcal{J}}\cdot\vec{S} \;=\; (\mathcal{J}^{(tot)})^2 \,-\, \mathcal{J}^2 \,-\, S^2 \tag{6.154}$$

one can readily show that the states $\mathcal{Y}_{jm}(\ell, \frac{1}{2}|\hat{r})$ are also eigenstates of $\vec{\mathcal{J}} \cdot \vec{S}$. Employing (6.149, 6.151, 6.152) one derives

$$\begin{aligned} & 2\vec{\mathcal{J}} \cdot \vec{S}\, \mathcal{Y}_{jm}(j - \tfrac{1}{2}, \tfrac{1}{2}|\hat{r}) \\ &= \hbar^2 \left[j(j+1) - (j - \tfrac{1}{2})(j + \tfrac{1}{2}) - \tfrac{3}{4} \right] \mathcal{Y}_{jm}(j - \tfrac{1}{2}, \tfrac{1}{2}|\hat{r}) \\ &= \hbar^2 \left(j - \tfrac{1}{2} \right) \mathcal{Y}_{jm}(j - \tfrac{1}{2}, \tfrac{1}{2}|\hat{r}) \end{aligned} \tag{6.155}$$

and

$$\begin{aligned} & 2\vec{\mathcal{J}} \cdot \vec{S}\, \mathcal{Y}_{jm}(j + \tfrac{1}{2}, \tfrac{1}{2}|\hat{r}) \\ &= \hbar^2 [j(j+1) - (j + \tfrac{1}{2})(j + \tfrac{3}{2}) - \tfrac{3}{4}] \mathcal{Y}_{jm}(j + \tfrac{1}{2}, \tfrac{1}{2}|\hat{r}) \\ &= \hbar^2 \left(-j - \tfrac{3}{2} \right) \mathcal{Y}_{jm}(j + \tfrac{1}{2}, \tfrac{1}{2}|\hat{r}) \, . \end{aligned} \tag{6.156}$$

Orthonormality Properties

The construction (6.141) in terms of Clebsch-Gordon coefficients produces normalized states. Since eigenstates of hermitean operators, i.e., of $(\vec{\mathcal{J}} + \vec{S})^2$, $(\vec{\mathcal{J}} + \vec{S})_3$, $\mathcal{J}^2$ with different eigenvalues are orthogonal, one can conclude the orthonormality property

$$\int_{-\pi}^{+\pi} \sin\theta d\theta \int_0^{2\pi} d\phi \, [\, \mathcal{Y}^*_{j'm'}(\ell', \tfrac{1}{2}|\theta, \phi) \,]^T \, \mathcal{Y}^*_{jm}(\ell, \tfrac{1}{2}|\theta, \phi) \; = \; \delta_{jj'} \delta_{mm'} \delta_{\ell\ell'} \tag{6.157}$$

where we have introduced the angular variables θ, ϕ to represent $\hat{r}$ and used the notation $[\cdots]^T$ to denote the transpose of the two-dimensional vectors $\mathcal{Y}^*_{j'm'}(\ell', \frac{1}{2}|\theta, \phi)$ which defines the scalar product

$$\begin{pmatrix} a^* \\ b^* \end{pmatrix}^T \begin{pmatrix} c \\ d \end{pmatrix} \; = \; \begin{pmatrix} a^* & b^* \end{pmatrix} \begin{pmatrix} c \\ d \end{pmatrix} \; = \; a * c \; + \; b^* \, d \, . \tag{6.158}$$

The Operator $\sigma \cdot \hat{r}$

Another important property of the spin-orbital angular momentum states (6.147, 6.148) concerns the effect of the operator $\vec{\sigma} \cdot \hat{r}$ on these states. In a representation defined by the states (6.139), this operator can be represented by a 2×2 matrix.
We want to show that the operator $\vec{\sigma} \cdot \hat{r}$ in the basis

$$\begin{aligned} & \{ (\mathcal{Y}_{jm}(j - \tfrac{1}{2}, \tfrac{1}{2}|\hat{r}), \mathcal{Y}_{jm}(j + \tfrac{1}{2}, \tfrac{1}{2}|\hat{r})) \, , \\ & \qquad j \; = \; \tfrac{1}{2}, \, \tfrac{3}{2} \, \ldots \, ; \, m \; = \; -j, \, -j+1, \, \ldots + j \, \} \end{aligned} \tag{6.159}$$

assumes the block-diagonal form

$$\vec{\sigma} \cdot \hat{r} \; = \; \begin{pmatrix} 0 & -1 & & & \\ -1 & 0 & & & \\ & & 0 & -1 & \\ & & -1 & 0 & \\ & & & & \ddots \end{pmatrix} \tag{6.160}$$

where the blocks operate on two-dimensional subspaces spanned by $\{\mathcal{Y}_{jm}(j-\frac{1}{2},\frac{1}{2}|\hat{r}),\,\mathcal{Y}_{jm}(j+\frac{1}{2},\frac{1}{2}|\hat{r})\}$. We first demonstrate that $\vec{\sigma}\cdot\hat{r}$ is block-diagonal. This property follows from the commutation relationships

$$[\mathcal{J}_k^{(tot)},\,\vec{\sigma}\cdot\vec{r}\,]\;=\;0\,,\quad k\,=\,1,\,2,\,3 \tag{6.161}$$

where $\mathcal{J}_k^{tot}$ is defined in (6.137) To prove this we consider the case $k=1$. For the l.h.s. of (6.161) holds, using (6.137),

$$\begin{aligned}&[\,\mathcal{J}_1\,+\,S_1,\,\sigma_1x_1\,+\,\sigma_2x_2\,+\,\sigma_3x_3\,]\\ &\quad=\quad\sigma_2\,[\,\mathcal{J}_1,\,x_2\,]\,+\,[\,S_1,\,\sigma_2\,]\,x_2\,+\,\sigma_3\,[\,\mathcal{J}_1,\,x_3\,]\,+\,[\,S_1,\,\sigma_3\,]\,x_3\,.\end{aligned} \tag{6.162}$$

The commutation properties [cf. (5.53) for $\mathcal{J}_1$ and (5.228), (6.138) for $\vec{\sigma}$ and S_1]

$$[\,\mathcal{J}_1,\,x_2\,]\;=\;-i\hbar\,[\,x_2\partial_3\,-\,x_3\partial_2,\,x_2\,]\;=\;i\hbar\,x_3 \tag{6.163}$$

$$[\,\mathcal{J}_1,\,x_3\,]\;=\;-i\hbar\,[\,x_2\partial_3\,-\,x_3\partial_2,\,x_3\,]\;=\;-\,i\hbar\,x_2 \tag{6.164}$$

$$[S_1,\,\sigma_2\,]\;=\;\frac{\hbar}{2}\,[\,\sigma_1,\,\sigma_2\,]\;=\;i\hbar\sigma_3 \tag{6.165}$$

$$[S_1,\,\sigma_3\,]\;=\;\frac{\hbar}{2}\,[\,\sigma_1,\,\sigma_3\,]\;=\;-\,i\hbar\sigma_2 \tag{6.166}$$

allow one then to evaluate the commutator (6.161) for $k=1$

$$[\mathcal{J}_1^{(tot)},\,\vec{\sigma}\cdot\vec{r}\,]\;=\;i\hbar\,(\,\sigma_2x_3\,+\,\sigma_3x_2\,-\,\sigma_3x_2\,-\,\sigma_2x_3\,)\;=\;0\,. \tag{6.167}$$

One can carry out this algebra in a similar way for the $k=2,\,3$ and, hence, prove (6.161).
Since the differential operators in $\mathcal{J}_k^{(tot)}$ do not contain derivatives with respect to r, the property (6.161) applies also to $\vec{\sigma}\cdot\hat{r}$, i.e., it holds

$$[\mathcal{J}_k^{(tot)},\,\vec{\sigma}\cdot\hat{r}]\;=\;0\,,\quad k\,=\,1,\,2\,3\,. \tag{6.168}$$

From this follows

$$\begin{aligned}&\left(\mathcal{J}^{(tot)}\right)^2\,\vec{\sigma}\cdot\hat{r}\;\mathcal{Y}_{jm}(j\pm\tfrac{1}{2},\tfrac{1}{2}|\hat{r})\\ &\quad=\quad\vec{\sigma}\cdot\hat{r}\;\left(\mathcal{J}^{(tot)}\right)^2\,\mathcal{Y}_{jm}(j\pm\tfrac{1}{2},\tfrac{1}{2}|\hat{r})\\ &\quad=\quad\hbar^2\,j(j+1)\,\vec{\sigma}\cdot\hat{r}\,\mathcal{Y}_{jm}(j\pm\tfrac{1}{2},\tfrac{1}{2}|\hat{r})\,,\end{aligned} \tag{6.169}$$

i.e., $\vec{\sigma}\cdot\hat{r}\,\mathcal{Y}_{jm}(j\pm\frac{1}{2},\frac{1}{2}|\hat{r})$ is an eigenstate of $(\mathcal{J}^{(tot)})^2$ with eigenvalue $\hbar^2j(j+1)$. One can prove similarly that this state is also an eigenstate of $\mathcal{J}_3^{(tot)}$ with eigenvalue $\hbar m$. Since in the space spanned by the basis (6.159) only two states exist with such eigenvalues, namely, $\mathcal{Y}_{jm}(j\pm\frac{1}{2},\frac{1}{2}|\hat{r})$, one can conclude

$$\vec{\sigma}\cdot\hat{r}\,\mathcal{Y}_{jm}(j+\tfrac{1}{2},\tfrac{1}{2}|\hat{r})\;=\;\alpha_{++}(jm)\,\mathcal{Y}_{jm}(j+\tfrac{1}{2},\tfrac{1}{2}|\hat{r})\;+\;\alpha_{+-}(jm)\,\mathcal{Y}_{jm}(j-\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \tag{6.170}$$

and, similarly,

$$\vec{\sigma}\cdot\hat{r}\,\mathcal{Y}_{jm}(j-\tfrac{1}{2},\tfrac{1}{2}|\hat{r})\;=\;\alpha_{-+}(jm)\,\mathcal{Y}_{jm}(j+\tfrac{1}{2},\tfrac{1}{2}|\hat{r})\;+\;\alpha_{--}(jm)\,\mathcal{Y}_{jm}(j-\tfrac{1}{2},\tfrac{1}{2}|\hat{r})\,. \tag{6.171}$$

We have denoted here that the expansion coefficients $\alpha_{\pm\pm}$, in principle, depend on j and m. We want to demonstrate now that the coefficients $\alpha_{\pm\pm}$, actually, do not depend on m. This property follows from

$$[\,\mathcal{J}_{\pm}^{(tot)},\,\vec{\sigma}\cdot\hat{r}\,]\;=\;0 \tag{6.172}$$

which is a consequence of (6.168) and the definition of $\mathcal{J}_{\pm}^{(tot)}$ [c.f. (6.35)]. We will, hence, use the notation $\alpha_{\pm\pm}(j)$

Exercise 6.5.1: Show that (6.172) implies that the coefficients $\alpha_{\pm\pm}$ in (6.170, 6.171) are independent of m.

We want to show now that the coeffients $\alpha_{++}(j)$ and $\alpha_{--}(j)$ in (6.170, 6.171) vanish. For this purpose we consider the parity of the operator $\vec{\sigma}\cdot\hat{r}$ and the parity of the states $\mathcal{Y}_{jm}(j\pm\frac{1}{2},\frac{1}{2}|\hat{r})$, i.e., their property to change only by a factor ±1 under spatial inversion. For $\vec{\sigma}\cdot\hat{r}$ holds

$$\vec{\sigma}\cdot\hat{r}\;\rightarrow\;\vec{\sigma}\cdot(-\hat{r})\;=\;-\,\vec{\sigma}\cdot\hat{r}\,, \tag{6.173}$$

i.e., $\vec{\sigma}\cdot\hat{r}$ has odd parity. Replacing the $\hat{r}$-dependence by the corresponding (θ,ϕ)-dependence and noting the inversion symmetry of spherical harmonics [c.f. (5.166)]

$$Y_{j+\frac{1}{2}m\pm\frac{1}{2}}(\pi-\theta,\pi+\phi)\;=\;(-1)^{j+\frac{1}{2}}Y_{j+\frac{1}{2}m\pm\frac{1}{2}}(\theta,\phi) \tag{6.174}$$

one can conclude for $\mathcal{Y}_{jm}(j+\frac{1}{2},\frac{1}{2}|\hat{r})$ as given by (6.142)

$$\mathcal{Y}_{jm}(j+\tfrac{1}{2},\tfrac{1}{2}|\theta,\phi)\;\rightarrow\;\mathcal{Y}_{jm}(j+\tfrac{1}{2},\tfrac{1}{2}|\pi-\theta,\pi+\phi)\;=\;(-1)^{j+\frac{1}{2}}\mathcal{Y}_{jm}(j+\tfrac{1}{2},\tfrac{1}{2}|\theta,\phi)\,. \tag{6.175}$$

Similarly follows for $\mathcal{Y}_{jm}(j-\frac{1}{2},\frac{1}{2}|\hat{r})$

$$\mathcal{Y}_{jm}(j-\tfrac{1}{2},\tfrac{1}{2}|\theta,\phi)\;\rightarrow\;(-1)^{j-\frac{1}{2}}\mathcal{Y}_{jm}(j+\tfrac{1}{2},\tfrac{1}{2}|\theta,\phi)\,. \tag{6.176}$$

We note that $\mathcal{Y}_{jm}(j+\frac{1}{2},\frac{1}{2}|\hat{r})$ and $\mathcal{Y}_{jm}(j-\frac{1}{2},\frac{1}{2}|\hat{r})$ have opposite parity. Since $\vec{\sigma}\cdot\hat{r}$ has odd parity, i.e., when applied to the states $\mathcal{Y}_{jm}(j\pm\frac{1}{2},\frac{1}{2}|\hat{r})$ changes their parity, we can conclude $\alpha_{++}(j)=\alpha_{--}(j)=0$. The operator $\vec{\sigma}\cdot\hat{r}$ in the two-dimensional subspace spanned by $\mathcal{Y}_{jm}(j\pm\frac{1}{2},\frac{1}{2}|\hat{r})$ assumes then the form

$$\vec{\sigma}\cdot\hat{r}\;=\;\begin{pmatrix}0 & \alpha_{+-}(j)\\ \alpha_{-+}(j) & 0\end{pmatrix}\,. \tag{6.177}$$

Since $\vec{\sigma}\cdot\hat{r}$ must be a hermitean operator it must hold $\alpha_{-+}(j)=\alpha_{+-}^{*}(j)$.
According to (5.230) one obtains

$$(\,\vec{\sigma}\cdot\hat{r}\,)^2\;=\;\mathbb{1}\,. \tag{6.178}$$

This implies $|\alpha_{+-}(j)|=1$ and, therefore, one can write

$$\vec{\sigma}\cdot\hat{r}\;=\;\begin{pmatrix}0 & e^{i\beta(j)}\\ e^{-i\beta(j)} & 0\end{pmatrix}\,,\qquad \beta(j)\,\in\,\mathbb{R}\,. \tag{6.179}$$

One can demonstrate that $\vec{\sigma}\cdot\hat{r}$ is, in fact, a real operator. For this purpose one considers the operation of $\vec{\sigma}\cdot\hat{r}$ for the special case $\phi=0$. According to the expressions (6.147, 6.148) for

$\mathcal{Y}_{jm}(j \pm \frac{1}{2}, \frac{1}{2}|\theta, \phi)$ and (5.174–5.177) one notes that for $\phi = 0$ the spin-angular momentum states are entirely real such that $\vec{\sigma} \cdot \hat{r}$ must be real as well. One can conclude then

$$\vec{\sigma} \cdot \hat{r} \;=\; \pm \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{6.180}$$

where the sign could depend on j.
We want to demonstrate finally that the "$-$"-sign holds in (6.180). For this purpose we consider the application of $\vec{\sigma} \cdot \hat{r}$ in the case of $\theta = 0$. According to (5.180) and (6.147, 6.148) the particular states $\mathcal{Y}_{j\frac{1}{2}}(j - \frac{1}{2}, \frac{1}{2}|\hat{r})$ and $\mathcal{Y}_{j\frac{1}{2}}(j + \frac{1}{2}, \frac{1}{2}|\hat{r})$ at $\theta = 0$ are

$$\mathcal{Y}_{j\frac{1}{2}}(j - \tfrac{1}{2}, \tfrac{1}{2}|\theta = 0, \phi) \;=\; \begin{pmatrix} \sqrt{\frac{j+\frac{1}{2}}{4\pi}} \\ 0 \end{pmatrix} \tag{6.181}$$

$$\mathcal{Y}_{j\frac{1}{2}}(j + \tfrac{1}{2}, \tfrac{1}{2}|\theta = 0, \phi) \;=\; \begin{pmatrix} -\sqrt{\frac{j+\frac{1}{2}}{4\pi}} \\ 0 \end{pmatrix} . \tag{6.182}$$

Since $\vec{\sigma} \cdot \hat{r}$, given by

$$\vec{\sigma} \cdot \hat{r} \;=\; \sigma_1 \sin\theta \cos\phi \;+\; \sigma_2 \sin\theta \sin\phi \;+\; \sigma_3 \cos\theta \,, \tag{6.183}$$

in case $\theta = 0$ becomes in the standard representation with respect to the spin-$\frac{1}{2}$ states $\chi_{\frac{1}{2}\pm\frac{1}{2}}$ [c.f. (5.224)]

$$\vec{\sigma} \cdot \hat{r} \;=\; \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}_{\text{space } \chi_{\frac{1}{2}\pm\frac{1}{2}}} , \qquad \text{for } \theta = 0 \tag{6.184}$$

one can conclude from (6.181, 6.182)

$$\vec{\sigma} \cdot \hat{r}\, \mathcal{Y}_{j\frac{1}{2}}(j - \tfrac{1}{2}, \tfrac{1}{2}|\theta = 0, \phi) \;=\; -\mathcal{Y}_{j\frac{1}{2}}(j + \tfrac{1}{2}, \tfrac{1}{2}|\theta = 0, \phi) \,, \qquad \text{for } \theta = 0. \tag{6.185}$$

We have, hence, identified the sign of (6.180) and, therefore, have proven (6.160). The result can also be stated in the compact form

$$\vec{\sigma} \cdot \hat{r}\, \mathcal{Y}_{jm}(j \pm \tfrac{1}{2}, \tfrac{1}{2}|\hat{r}) \;=\; -\mathcal{Y}_{jm}(j \mp \tfrac{1}{2}, \tfrac{1}{2}|\hat{r}) \tag{6.186}$$

The Operator $\vec{\sigma} \cdot \hat{\vec{p}}$

The operator $\vec{\sigma} \cdot \hat{\vec{p}}$ plays an important role in relativistic quantum mechanics. We want to determine its action on the wave functions $f(r)\, \mathcal{Y}_{jm}(j \pm \frac{1}{2}, \frac{1}{2}|\hat{r})$. Noting that $\hat{\vec{p}} = -i\hbar\nabla$ is a first order differential operator it holds

$$\vec{\sigma} \cdot \hat{\vec{p}}\, f(r)\, \mathcal{Y}_{jm}(j \pm \tfrac{1}{2}, \tfrac{1}{2}|\hat{r}) \;=\; ((\vec{\sigma} \cdot \hat{\vec{p}} f(r)))\mathcal{Y}_{jm}(j \pm \tfrac{1}{2}, \tfrac{1}{2}|\hat{r}) \;+\; f(r)\, \vec{\sigma} \cdot \hat{\vec{p}} \mathcal{Y}_{jm}(j \pm \tfrac{1}{2}, \tfrac{1}{2}|\hat{r}) \tag{6.187}$$

Here $((\cdots))$ denotes again confinement of the diffusion operator ∂_r to within the double bracket. Since $f(r)$ is independent of θ and ϕ follows

$$((\vec{\sigma} \cdot \hat{\vec{p}} f(r))) \;=\; -i\hbar((\,\partial_r f(r)\,))\, \vec{\sigma} \cdot \hat{r} \,. \tag{6.188}$$

Using (6.186) for both terms in (6.187) one obtains

$$\vec{\sigma}\cdot\hat{\vec{p}}\,f(r)\,\mathcal{Y}_{jm}(j\pm\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \;=\; \Big[\,i\hbar\partial_r f(r) \;-\; f(r)\,\vec{\sigma}\cdot\hat{\vec{p}}\,\vec{\sigma}\cdot\hat{r}\Big]\,\mathcal{Y}_{jm}(j\mp\tfrac{1}{2},\tfrac{1}{2}|\hat{r})\,. \tag{6.189}$$

The celebrated property of the Pauli matrices (5.230) allows one to express

$$\vec{\sigma}\cdot\hat{\vec{p}}\,\vec{\sigma}\cdot\hat{r} \;=\; \hat{\vec{p}}\cdot\hat{r} \;+\; i\,\vec{\sigma}\cdot(\hat{\vec{p}}\times\hat{r})\,. \tag{6.190}$$

For the test function $h(\vec{r})$ holds

$$\begin{aligned}\hat{\vec{p}}\cdot\hat{r}\,h \;&=\; -i\hbar\nabla\cdot\left(\frac{\vec{r}}{r}\,h\right) \;=\; -i\hbar\frac{h}{r}\,\nabla\cdot\vec{r} \;-\; i\hbar\,\vec{r}\cdot\nabla\frac{h}{r}\\ &=\; -i\hbar\,\frac{3}{r}h \;-\; i\hbar\,h\,\vec{r}\cdot\nabla\frac{1}{r} \;-\; i\hbar\,\hat{r}\cdot\nabla h\,.\end{aligned} \tag{6.191}$$

Using $\nabla(1/r) \;=\; -\vec{r}/r^3$ and $\hat{r}\cdot\nabla h \;=\; \partial_r h$ one can conclude

$$\hat{\vec{p}}\cdot\hat{r}\,h \;=\; -i\hbar\left(\frac{2}{r} \;+\; \partial_r\right)\,h \tag{6.192}$$

The operator $\hat{\vec{p}}\times\hat{r}$ in (6.190) can be related to the angular momentum operator. To demonstrate this we consider one of its cartision components, e.g.,

$$\begin{aligned}(\,\hat{\vec{p}}\times\hat{r}\,)_1\,h \;&=\; -i\hbar\,(\partial_2\frac{x_3}{r} \;-\; \partial_3\frac{x_2}{r})\,h\\ &=\; -i\hbar\,\frac{1}{r}\,(\partial_2 x_3 \;-\; \partial_3 x_2)\,h - i\hbar\,h\,(x_3\partial_2\frac{1}{r} \;-\; x_2\partial_3\frac{1}{r})\,.\end{aligned} \tag{6.193}$$

Using $\partial_2(1/r) \;=\; -x_2/r^3$, $\partial_3(1/r) \;=\; -x_3/r^3$ and $\partial_2 x_3 \;=\; x_3\partial_2$, $\partial_3 x_2 \;=\; x_2\partial_3$ we obtain

$$(\,\hat{\vec{p}}\times\hat{r}\,)_1\,h \;=\; -i\hbar\,\frac{1}{r}\,(x_3\partial_2 \;-\; x_2\partial_3)\,h \;=\; -\frac{1}{r}\,\mathcal{J}_1\,h \tag{6.194}$$

where $\mathcal{J}_1$ is defined in (5.53). Corresponding results are obtained for the other components of $\hat{\vec{p}}\times\hat{r}$ and, hence, we conclude the intuitively expected identity

$$\hat{\vec{p}}\times\hat{r} \;=\; -\,\frac{1}{r}\vec{\mathcal{J}}\,. \tag{6.195}$$

Altogether we obtain, using $\partial_r\mathcal{Y}_{jm}(j\pm\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \;=\; 0$ and $\vec{\sigma} \;=\; 2\vec{S}/\hbar$,

$$\vec{\sigma}\cdot\hat{\vec{p}}\,f(r)\,\mathcal{Y}_{jm}(j\pm\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \;=\; i\,[\,\hbar\partial_r \;+\; \frac{2\hbar}{r} \;+\; \frac{1}{r}\,\frac{\vec{\mathcal{J}}\cdot\vec{S}}{\hbar}\,]\,f(r)\,\mathcal{Y}_{jm}(j\mp\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \tag{6.196}$$

Using (6.155, 6.156) this yields finally

$$\vec{\sigma}\cdot\hat{\vec{p}}\,f(r)\,\mathcal{Y}_{jm}(j+\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \;=\; i\hbar\left[\,\partial_r \;+\; \frac{j+\frac{3}{2}}{r}\right]\,f(r)\,\mathcal{Y}_{jm}(j-\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \tag{6.197}$$

$$\vec{\sigma}\cdot\hat{\vec{p}}\,g(r)\,\mathcal{Y}_{jm}(j-\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \;=\; i\hbar\left[\,\partial_r \;+\; \frac{\frac{1}{2}-j}{r}\right]\,g(r)\,\mathcal{Y}_{jm}(j+\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \tag{6.198}$$

To demonstrate the validity of this key result we note that according to (5.230, 5.99) holds

$$(\vec{\sigma}\cdot\hat{\vec{p}})^2 \;=\; -\hbar^2\nabla^2 \;=\; \frac{\hbar^2}{r}\,\frac{\partial^2}{\partial r^2}\,r \;+\; \frac{\mathcal{J}^2}{r^2}\;. \tag{6.199}$$

We want to show that eqs. (6.197, 6.198), in fact, are consistent with this identity. We note

$$\begin{aligned}
&(\vec{\sigma}\cdot\hat{\vec{p}})^2 f(r)\,\mathcal{Y}_{jm}(j+\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \\
&= \hbar\,\vec{\sigma}\cdot\hat{\vec{p}}\,[\,i\partial_r \;+\; \frac{i}{r}(j+\frac{3}{2})\,]\,f(r)\,\mathcal{Y}_{jm}(j-\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \\
&= \hbar^2\,[\,i\partial_r \;+\; \frac{i}{r}(\frac{1}{2}-j)\,]\,[\,i\partial_r \;+\; \frac{i}{r}(j+\frac{3}{2})\,]\,f(r)\,\mathcal{Y}_{jm}(j+\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \\
&= \hbar^2\,[\,-\partial_r^2 \;-\; \frac{2}{r}\partial_r \;+\; \frac{j+\frac{3}{2}}{r^2} \;-\; \frac{(\frac{1}{2}-j)(j+\frac{3}{2})}{r^2}\,]\;f(r)\,\mathcal{Y}_{jm}(j+\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \\
&= \hbar^2\,[\,-\partial_r^2 \;-\; \frac{2}{r}\partial_r \;+\; \frac{j^2+2j+\frac{3}{4}}{r^2}\,]\;f(r)\,\mathcal{Y}_{jm}(j+\tfrac{1}{2},\tfrac{1}{2}|\hat{r})
\end{aligned} \tag{6.200}$$

and, using (5.101), $j^2+2j+\frac{3}{4} \;=\; (j+\frac{1}{2})(j+\frac{3}{2})$, as well as (6.151), i.e.,

$$\hbar^2\,(j+\frac{1}{2})(j+\frac{3}{2})\,\mathcal{Y}_{jm}(j+\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \;=\; \mathcal{J}^2\,\mathcal{Y}_{jm}(j+\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \tag{6.201}$$

yields

$$(\vec{\sigma}\cdot\hat{\vec{p}})^2\,f(r)\,\mathcal{Y}_{jm}(j+\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \;=\; \left(-\frac{\hbar^2}{r}\,\partial_r^2 r \;+\; \frac{\mathcal{J}^2}{r^2}\right)\,f(r)\,\mathcal{Y}_{jm}(j+\tfrac{1}{2},\tfrac{1}{2}|\hat{r})$$

which agrees with (6.199).

Evaluation of Relevant Clebsch-Gordan Coefficients

We want to determine now the Clebsch-Gordan coefficients (6.143–6.146). For this purpose we use the construction method introduced in Sec. 6.2. We begin with the coefficients (6.143, 6.144) and, adopting the method in Sec. 6.2, consider first the case of the largest m-value $m = j$. In this case holds, according to (6.43),

$$\mathcal{Y}_{jj}(j-\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \;=\; Y_{j-\frac{1}{2}j-\frac{1}{2}}(\hat{r})\,\chi_{\frac{1}{2}\frac{1}{2}}\;. \tag{6.202}$$

The Clebsch-Gordan coefficients are then

$$(j-\tfrac{1}{2},j-\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2}|j,j) \;=\; 1 \tag{6.203}$$

$$(j-\tfrac{1}{2},j+\tfrac{1}{2},\tfrac{1}{2},-\tfrac{1}{2}|j,j) \;=\; 0 \tag{6.204}$$

$$\tag{6.205}$$

which agrees with the expressions (6.143, 6.144) for $m = j$.
For $m = j-1$ one can state, according to (6.45),

$$\mathcal{Y}_{jj-1}(j-\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \;=\; \sqrt{\frac{2j-1}{2j}}\;Y_{j-\frac{1}{2}j-\frac{3}{2}}\,\chi_{\frac{1}{2}\frac{1}{2}} \;+\; \sqrt{\frac{1}{2j}}\;Y_{j-\frac{1}{2}j-\frac{1}{2}}\,\chi_{\frac{1}{2}-\frac{1}{2}} \tag{6.206}$$

The corresponding Clebsch-Gordan coefficients are then

$$(j-\tfrac{1}{2}, j-\tfrac{3}{2}, \tfrac{1}{2}, \tfrac{1}{2}|j, j-1) \;=\; \sqrt{\frac{2j-1}{2j}} \tag{6.207}$$

$$(j-\tfrac{1}{2}, j-\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}|j, j-1) \;=\; \sqrt{\frac{1}{2j}} \tag{6.208}$$

$$\tag{6.209}$$

which again agrees with the expressions (6.143, 6.144) for $m = j-1$.
Expression (6.206), as described in Sec. 6.2, is obtained by applying the operator [c.f. (6.137)]

$$\mathcal{J}_-^{(tot)} \;=\; \mathcal{J}_- \;+\; S_- \tag{6.210}$$

to (6.202). The further Clebsch-Gordan coefficients $(\cdots|jj-2)$, $(\cdots|jj-3)$, etc., are obtained by iterating the application of (6.210). Let us verify then the expression (6.143, 6.144) for the Clebsch-Gordan coefficients by induction. (6.143, 6.144) implies for $j = m$

$$\mathcal{Y}_{jm}(j-\tfrac{1}{2}, \tfrac{1}{2}|\hat{r}) \;=\; \sqrt{\frac{j+m}{2j}}\, Y_{j-\frac{1}{2}m-\frac{1}{2}}\, \chi_{\frac{1}{2}\frac{1}{2}} \;+\; \sqrt{\frac{j-m}{2j}}\, Y_{j-\frac{1}{2}m+\frac{1}{2}}\, \chi_{\frac{1}{2}-\frac{1}{2}}\;. \tag{6.211}$$

Applying $\mathcal{J}_-^{(tot)}$ to the l.h.s. and $\mathcal{J}_- + S_-$ to the r.h.s. [c.f. (6.210)] yields

$$\begin{aligned}
&\sqrt{(j+m)(j-m+1)}\, \mathcal{Y}_{jm-1}(j-\tfrac{1}{2}, \tfrac{1}{2}|\hat{r}) \;= \\
&\quad \sqrt{\frac{j+m}{2j}}\, \sqrt{(j+m-1)(j-m+1)}\, Y_{j-\frac{1}{2}m-\frac{3}{2}}\, \chi_{\frac{1}{2}\frac{1}{2}} \\
&\quad + \sqrt{\frac{j+m}{2j}}\, Y_{j-\frac{1}{2}m-\frac{1}{2}}\, \chi_{\frac{1}{2}-\frac{1}{2}} \\
&\quad + \sqrt{\frac{j-m}{2j}}\, \sqrt{(j+m)(j-m)}\, Y_{j-\frac{1}{2}m-\frac{1}{2}}\, \chi_{\frac{1}{2}-\frac{1}{2}}
\end{aligned}$$

or

$$\begin{aligned}
&\mathcal{Y}_{jm-1}(j-\tfrac{1}{2}, \tfrac{1}{2}|\hat{r}) \;= \\
&\quad \sqrt{\frac{j+m-1}{2j}}\, Y_{j-\frac{1}{2}m-\frac{3}{2}}\, \chi_{\frac{1}{2}\frac{1}{2}} \;+ \\
&\quad \left((j\,-\,m) \sqrt{\frac{1}{2j(j-m+1)}} \;+\; \sqrt{\frac{1}{2j(j-m+1)}} \right)\, Y_{j-\frac{1}{2}m-\frac{1}{2}}\, \chi_{\frac{1}{2}-\frac{1}{2}} \\
&\quad = \sqrt{\frac{j+m-1}{2j}}\, Y_{j-\frac{1}{2}m-\frac{3}{2}}\, \chi_{\frac{1}{2}\frac{1}{2}} \;+\; \sqrt{\frac{j-m+1}{2j}}\, Y_{j-\frac{1}{2}m-\frac{1}{2}}\, \chi_{\frac{1}{2}-\frac{1}{2}}\;.
\end{aligned}$$

This implies

$$(j-\tfrac{1}{2}, m-\tfrac{3}{2}, \tfrac{1}{2}, \tfrac{1}{2}|j, m-1) \;=\; \sqrt{\frac{j\,+\,m\,-\,1}{2j}} \tag{6.212}$$

$$(j-\tfrac{1}{2}, m-\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}|j, m-1) \quad = \quad \sqrt{\frac{j-m+1}{2j}} \tag{6.213}$$

which is in agreement with (6.143, 6.144) for $j = m - 1$. We have, hence, proven (6.143, 6.144) by induction.
The Clebsch-Gordan coefficients (6.146) can be obtained from (6.143) by applying the symmetry relationships (6.116, 6.120). The latter relationships applied together read

$$\begin{aligned}(\ell_1, m_1, \ell_2, m_2|\ell_3, m_3) \quad &= \quad (-1)^{\ell_2+\ell_3-\ell_1+\ell_2+m_2} \ \times \\ &\times \sqrt{\frac{2\ell_3+1}{2\ell_1+1}}\,(\ell_3, m_3, \ell_2, -m_2|\ell_1, m_1)\end{aligned} \tag{6.214}$$

For

$$(j, m, \tfrac{1}{2}, \tfrac{1}{2}|j+\tfrac{1}{2}, m+\tfrac{1}{2}) \ = \ \sqrt{\frac{j+m+1}{2j+1}}\,, \tag{6.215}$$

which follows from (6.143), the relationship (6.214) yields

$$(j, m, \tfrac{1}{2}, \tfrac{1}{2}|j+\tfrac{1}{2}, m+\tfrac{1}{2}) \ = \ \sqrt{\frac{2j+2}{2j+1}}(j+\tfrac{1}{2}, m+\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}|j, m) \tag{6.216}$$

and, using (6.215), one obtains

$$(j+\tfrac{1}{2}, m+\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}|j, m) \ = \ \sqrt{\frac{j+m+1}{2j+2}}\,. \tag{6.217}$$

Similarly, one can obtain expression (6.145) from (6.144).

6.6 The $3j$–Coefficients

The *Clebsch-Gordan coefficients* describe the quantum mechanical equivalent of the addition of two classical angular momentum vectors $\vec{\mathcal{J}}^{(1)}_{\text{class}}$ and $\vec{\mathcal{J}}^{(2)}_{\text{class}}$ to obtain the total angular momentum vector $\vec{\mathcal{J}}^{(\text{tot})}_{\text{class}} = \vec{\mathcal{J}}^{(1)}_{\text{class}} + \vec{\mathcal{J}}^{(2)}_{\text{class}}$. In this context $\vec{\mathcal{J}}^{(1)}_{\text{class}}$ and $\vec{\mathcal{J}}^{(2)}_{\text{class}}$ play the same role, leading quantum mechanically to a symmetry of the *Clebsch-Gordan coefficients* $(JM|\ell_1 m_1 \ell_2 m_2)$ with respect to exchange of $\ell_1 m_1$ and $\ell_2 m_2$. However, a higher degree of symmetry is obtained if one rather considers classically to obtain a vector $\vec{\mathcal{J}}^{(-\text{tot})}_{\text{class}}$ with the property $\vec{\mathcal{J}}^{(1)}_{\text{class}} + \vec{\mathcal{J}}^{(2)}_{\text{class}} + \vec{\mathcal{J}}^{(-\text{tot})}_{\text{class}} = 0$. Obviously, all three vectors $\vec{\mathcal{J}}^{(1)}_{\text{class}}$, $\vec{\mathcal{J}}^{(2)}_{\text{class}}$ and $\vec{\mathcal{J}}^{(-\text{tot})}_{\text{class}}$ play equivalent roles.
The coefficients which are the quantum mechanical equivalent to $\vec{\mathcal{J}}^{(1)}_{\text{class}} + \vec{\mathcal{J}}^{(2)}_{\text{class}} + \vec{\mathcal{J}}^{(-\text{tot})}_{\text{class}} = 0$ are the $3j$–coefficients introduced by Wigner. They are related in a simple manner to the *Clebsch-Gordan coefficients*

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \ = \ (-1)^{j_1-j_2+m_3}(2j_3+1)^{-\frac{1}{2}}(j_3 - m_3|j_1\, m_1, j_2\, m_2) \tag{6.218}$$

where we have replaced the quantum numbers $J, M, \ell_1, m_1, \ell_2, m_2$ by the set $j_1, m_1, j_2, m_2, j_3, m_3$ to reflect in the notation the symmetry of these quantities.

We first like to point out that conditions (6.21, 6.34) imply

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = 0 \text{ if not } m_1 + m_2 + m_3 = 0 \tag{6.219}$$

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = 0 \text{ if not } |j_1 - j_2| \leq j_3 \leq j_1 + j_2 \,. \tag{6.220}$$

The latter condition $|j_1 - j_2| \leq j_3 \leq j_1 + j_2$, the so-called triangle condition, states that j_1, j_2, j_3 form the sides of a triangle and the condition is symmetric in the three quantum numbers.
According to the definition of the $3j$–coefficients one would expect symmetry properties with respect to exchange of j_1, m_1, j_2, m_2 and j_3, m_3 and with respect to sign reversals of all three values m_1, m_2, m_3, i.e. with respect to alltogether 12 symmetry operations. These symmetries follow the equations

$$\begin{aligned} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} &= \begin{pmatrix} j_2 & j_3 & j_1 \\ m_2 & m_3 & m_1 \end{pmatrix} = (-1)^{j_1+j_2+j_3} \begin{pmatrix} j_2 & j_1 & j_3 \\ m_2 & m_1 & m_3 \end{pmatrix} \\ &= (-1)^{j_1+j_2+j_3} \begin{pmatrix} j_1 & j_2 & j_3 \\ -m_1 & -m_2 & -m_3 \end{pmatrix} \end{aligned} \tag{6.221}$$

where the the results of a cyclic, anti-cyclic exchange and of a sign reversal are stated. In this way the values of 12 $3j$-coefficients are related.
However, there exist even further symmetry properties, discovered by Regge, for which no known classical analogue exists. To represent the full symmetry one expresses the $3j$–coefficients through a 3×3–matrix, the Regge-symbol, as follows

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = \begin{bmatrix} -j_1 + j_2 + j_3 & j_1 - j_2 + j_3 & j_1 + j_2 - j_3 \\ j_1 - m_1 & j_2 - m_2 & j_3 - m_3 \\ j_1 + m_1 & j_2 + m_2 & j_3 + m_3 \end{bmatrix} . \tag{6.222}$$

The Regge symbol vanishes, except when all elements are non-negative integers and each row and column has the same integer value $\Sigma = j_1 + j_2 + j_3$. The Regge symbol also vanishes in case that two rows or columns are identical and Σ is an odd integer. The Regge symbol reflects a remarkable degree of symmetry of the related $3j$–coefficients: One can exchange rows, one can exchange columns and one can reflect at the diagonal (transposition). In case of a non-cyclic exchange of rows and columns the $3j$–coefficent assumes a prefactor $(-1)^{\Sigma}$. These symmetry operations relate altogether 72 $3j$–coefficients.
The reader may note that the entries of the Regge-symbol, e.g., $-j_1 + j_2 + j_3$, are identical to the integer arguments which enter the analytical expression (6.102) of the Clebsch-Gordan coefficients, safe for the prefactor $\sqrt{2j_3 + 1}$ which is cancelled according to the definition (6.218) relating $3j$-coefficients and Clebsch-Gordan coefficients. The two integer entries $J - \ell_1 - m_2$ and $J - \ell_2 + m_1$ in (6.102) are obtained each through the difference of two entries of the Regge-symbol.
Because of its high degree of symmetry the Regge symbol is very suited for numerical evaluations of the $3j$–coefficents. For this purpose one can use the symmetry transformations to place the smallest element into the upper left corner of the Regge symbol. Assuming this placement the Regge symbol can be determined as follows (n_{11} is the smallest element!)

$$\begin{bmatrix} n_{11} & n_{12} & n_{13} \\ n_{21} & n_{22} & n_{23} \\ n_{31} & n_{32} & n_{33} \end{bmatrix} = (-1)^{n_{23}+n_{32}} \sqrt{\frac{n_{12}!n_{13}!n_{21}!n_{31}!}{(\Sigma+1)!n_{11}!n_{22}!n_{33}!n_{23}!n_{32}!))}} \sum_{n=0}^{n_{11}} s_n \tag{6.223}$$

where

$$\Sigma = n_{11} + n_{12} + n_{13} = j_1 + j_2 + j_3 \tag{6.224}$$

$$s_0 = \frac{n_{23}!}{(n_{23}-n_{11})!}\,\frac{n_{32}!}{(n_{32}-n_{11})!} \tag{6.225}$$

$$s_n = -\frac{(n_{11}+1-n)(n_{22}+1-n)(n_{33}+1-n)}{n(n_{23}-n_{11}+n)(n_{32}-n_{11}+n)}\,s_{n-1}\ . \tag{6.226}$$

We like to state finally a few explicit analytical expressions for *Clebsch-Gordan coefficients* which were actually obtained using (6.223-6.226) by means of a symbolic manipulation package (Mathematica)

$$\begin{aligned}(1m|2m_1 1m_2) &= \frac{(-1)^{1+m+m_1}\ \delta(m,m_1+m_2)\sqrt{(2-m_1)!}\sqrt{(2+m_1)!}}{\sqrt{10}\sqrt{(1-m)!}\sqrt{(1+m)!}\sqrt{(1-m_2)!}\sqrt{(1+m_2)!}} \\ &\quad 1\geq|m|\wedge 1\geq|m_2|\wedge 2\geq|m_1|\end{aligned} \tag{6.227}$$

$$\begin{aligned}(2m|2m_1 1m_2) &= \frac{(-1)^{m+m_1}\,(m+2m_2)\,\delta(m,m_1+m_2)\sqrt{(2-m_1)!\ \ (2+m_1)!}}{\sqrt{6}\sqrt{(2-m)!}\sqrt{(2+m)!}\sqrt{(1-m_2)!}\sqrt{(1+m_2)!}} \\ &\quad 1\geq|m_2|\wedge 2\geq|m|\wedge 2\geq|m_1|\end{aligned} \tag{6.228}$$

$$\begin{aligned}(3m|2m_1 1m_2) &= \frac{(-1)^{2m_1-2m_2}\sqrt{7}\ \delta(m,m_1+m_2)\sqrt{(3-m)!}\sqrt{(3+m)!}}{\sqrt{105}\sqrt{(2-m_1)!}\sqrt{(2+m_1)!}\sqrt{(1-m_2)!}\sqrt{(1+m_2)!}} \\ &\quad 1\geq|m_2|)\wedge 2\geq|m_1|\wedge 3\geq|m|\end{aligned} \tag{6.229}$$

$$\begin{aligned}(0m|\frac{1}{2}m_1\frac{1}{2}m_2) &= \frac{i(-1)^{1-m_2}\ \delta(0,m)\ \delta(-m_1,m_2)}{\sqrt{2}} \\ &\quad \frac{1}{2}\geq|m_1|)\end{aligned} \tag{6.230}$$

$$\begin{aligned}(1m|\frac{1}{2}m_1\frac{1}{2}m_2) &= \frac{(-1)^{2m_1-2m_2}\sqrt{3}\ \delta(m,m_1+m_2)\sqrt{(1-m)!}\sqrt{(1+m)!}}{\sqrt{6}\sqrt{(\frac{1}{2}-m_1)!}\sqrt{(\frac{1}{2}+m_1)!}\sqrt{(\frac{1}{2}-m_2)!}\sqrt{(\frac{1}{2}+m_2)!}} \\ &\quad \frac{1}{2}\geq|m_1|\wedge\frac{1}{2}\geq|m_2|\wedge 1\geq|m|\end{aligned} \tag{6.231}$$

Here is an explicit value of a Clebsch-Gordan coefficient:

$$\begin{aligned}&(70\tfrac{1}{2}\,-15\tfrac{1}{2}|120\,-10\,60\tfrac{1}{2}\,-5\tfrac{1}{2}) = \\ &\frac{479318529350314729494020934\sqrt{127}\sqrt{142}}{\sqrt{3583426199008157363513502706871897199698473122224133404619835 5}} \\ &\simeq\ 0.10752786393409395427444450130056540562826159542886\end{aligned} \tag{6.232}$$

We also illustrate the numerical values of a sequence of 3j-coefficients in Figure 6.1.

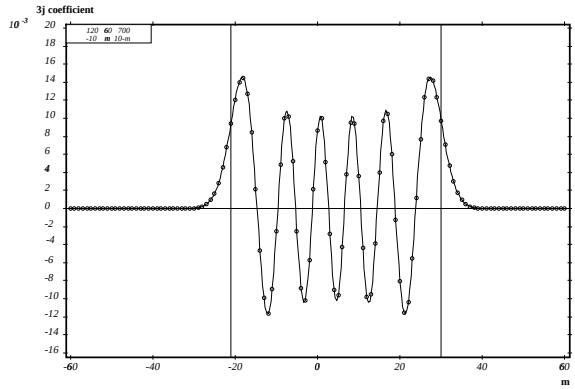


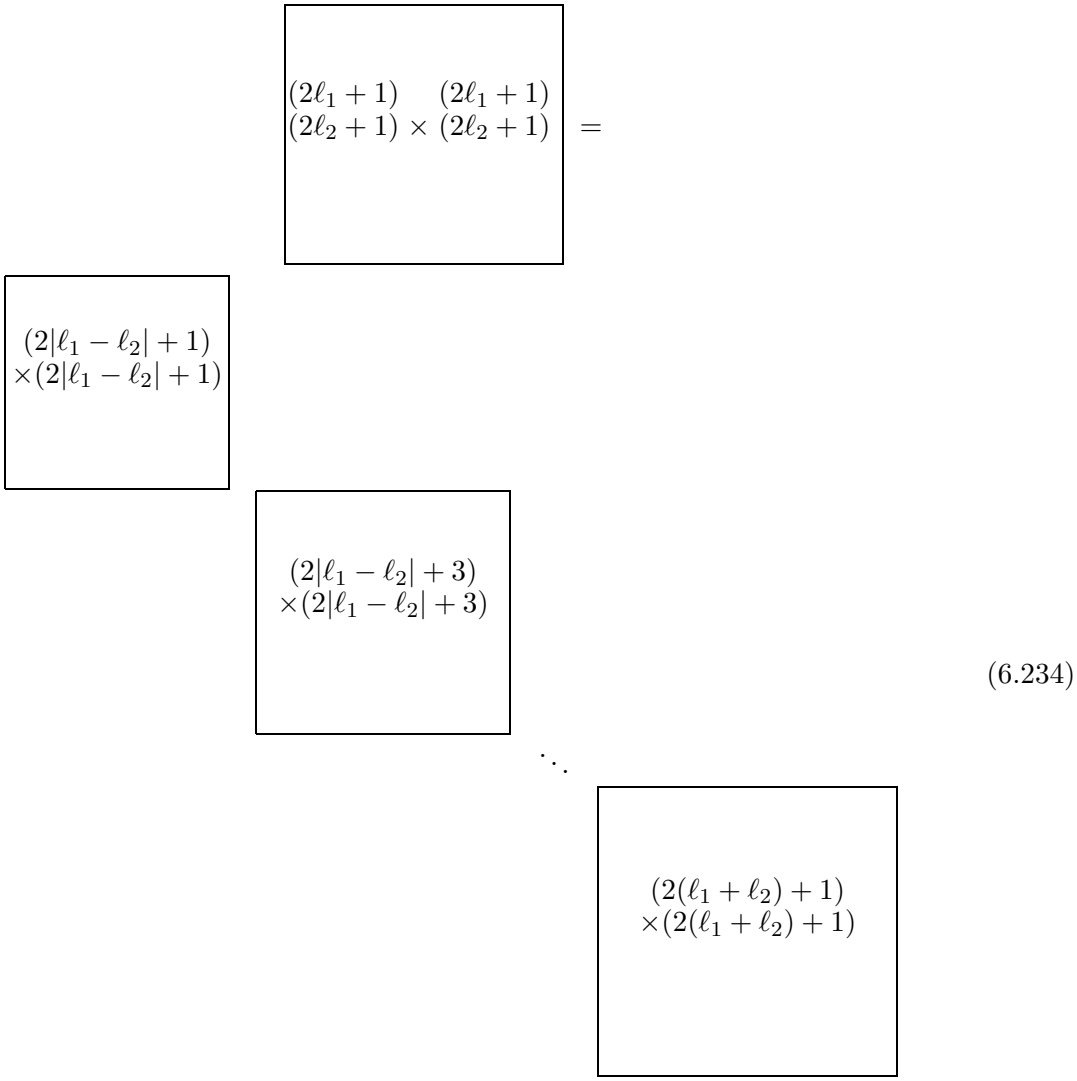
Figure 6.1: Oscillatory behavior of 3j-coefficients.

Irreducible Representation

We had stated above that the spherical harmonics $Y_{\ell m}(\Omega)$), eigenfunctions of the single particle angular momentum operators $\mathcal{J}^2$ and $\mathcal{J}_3$, provide the irreducible representation for $\mathcal{D}(\vartheta)$, i.e. the rotations in single particle function space. Similarly, the 2–particle total angular momentum wave functions $\mathfrak{Y}_{JM}(\ell_1, \ell_2|\Omega_1, \Omega_2)$ provide the irreducible representation for the rotation $\mathfrak{R}(\vartheta)$ defined in (6.5), i.e. rotations in 2–particle function space. If we define the matrix representation of $\mathfrak{R}(\vartheta)$ by $\mathfrak{D}(\vartheta)$, then for a basis $\{Y_{\ell_1 m_1}(\Omega_1)Y_{\ell_2 m_2}(\Omega_2), \ell_1, \ell_2 = 1, 2, \ldots, m_1 = -\ell_1, \ldots + \ell_1, m_2 = -\ell_2, \ldots + \ell_2\}$ the matrix has the blockdiagonal form

$$\mathcal{D}(\vartheta) = \begin{pmatrix} \boxed{1\cdot 1 \times 1\cdot 1} & & & \\ & \boxed{1\cdot 3 \times 1\cdot 3} & & \\ & & \ddots & \\ & & & \boxed{\begin{matrix}(2\ell_1+1) & (2\ell_1+1)\\ (2\ell_2+1) & \times\ (2\ell_2+1)\end{matrix}} \\ & & & & \ddots \end{pmatrix}. \tag{6.233}$$

For the basis $\{\mathfrak{Y}_{JM}(\ell_1,\ell_2|\Omega_1,\Omega_2),\ \ell_1,\ell_2 = 1,2,\ldots,\ J = |\ell_1-\ell_2|,\ldots,\ell_1+\ell_2,\ M = -J,\ldots J\}$ each of the blocks in (6.233) is further block-diagonalized as follows



(6.234)

Partitioning in smaller blocks is not possible.

Exercise 6.6.1: Prove Eqs. (6.233,6.234)

Exercise 6.6.2: How many overall singlet states can be constructed from four spin–$\frac{1}{2}$ states $|\frac{1}{2}m_1\rangle_{(1)}|\frac{1}{2}m_2\rangle_{(2)}|\frac{1}{2}m_3\rangle_{(3)}|\frac{1}{2}m_4\rangle_{(4)}$? Construct these singlet states in terms of the product wave functions above.

Exercise 6.6.3: Two triplet states $|1m_1\rangle_{(1)}|1m_2\rangle_{(2)}$ are coupled to an overall singlet state $\mathfrak{Y}_{00}(1,1)$. Show that the probability of detecting a triplet substate $|1m_1\rangle_{(1)}$ for arbitrary polarization (m_2–value) of the other triplet is $\frac{1}{3}$.

6.7 Tensor Operators and Wigner-Eckart Theorem

In this Section we want to discuss operators which have the property that they impart angular momentum and spin properties onto a quantum state. Such operators T can be characterized through their behaviour under rotational transformations.
Let $T|\psi\rangle$ denote the state obtained after the operator T has been applied and let $\mathcal{R}(\vec{\vartheta})$ denote a rotation in the representation of SO(3) or SU(2) which describes rotational transformations of the quantum states under consideration, e.g. the operator (5.42) in case (i) of the position representation of single particle wave functions or the operator (5.222) in case (ii) of single particle spin operators. Note that in the examples mentioned the operator T would be defined within the same representation as $\mathcal{R}(\vec{\vartheta})$. This implies, for example, that in

case (i) T is an operator $\mathbb{C}_\infty(\nVdash) \to \mathbb{C}_\infty(\nVdash)$ acting on single particle wave functions, e.g. a multiplicative operator $T\,\psi(\vec{r}) = f(\vec{r})\,\psi(\vec{r})$ or a differential operator $T\,\psi(\vec{r}) = (\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2})\,\psi(\vec{r})$;

case (ii) the operator T could be a spin operator S_k defined in (5.223), e.g. $S^2 = S_1^2 + S_2^2 + S_3^2$ or any other polynomial of S_k.

The operator T may also act on multi-particle states like $Y_{\ell_1 m_1}(\hat{r}_1)Y_{\ell_2 m_2}(\hat{r}_2)$. In fact, some of the examples considered below involve tensor operators T of this type.
Rotations transform $|\psi\rangle$ as $|\psi'\rangle = \mathcal{R}(\vec{\vartheta})|\psi\rangle$ and $T|\psi\rangle$ as $\mathcal{R}(\vec{\vartheta})T|\psi\rangle$. The latter can be written $T'|\psi'\rangle$ where T' denotes T in the rotated frame given by

$$T' = \mathcal{R}(\vec{\vartheta})\,T\,\mathcal{R}^{-1}(\vec{\vartheta}) \quad . \tag{6.235}$$

The property that T imparts onto states $|\psi\rangle$ angular momentum or spin corresponds to T behaving as an angular momentum or spin state multiplying $|\psi\rangle$. The latter implies that T transforms like an angular momentum or spin state $|\ell m\rangle$, i.e. that T belongs to a family of operators $\{T_{kq},\, q = -k, -k+1, \ldots k\}$ such that

$$T'_{kq} = \sum_{q'=-k}^{k} \mathcal{D}^{(k)}_{q'q}(\vec{\vartheta})\,T_{kq'} \; . \tag{6.236}$$

In this equation $\mathcal{D}^{(k)}_{q'q}(\vec{\vartheta})$ denotes the rotation matrix

$$\mathcal{D}^{(k)}_{q'q}(\vec{\vartheta}) = \langle kq'|\mathcal{R}(\vec{\vartheta})|kq\rangle \; . \tag{6.237}$$

The operators $T \in \{T_{kq},\, q = -k, -k+1, \ldots k\}$ with the transformation property (6.236, 6.237) are called *tensor operators of rank k*.

Examples of Tensor Operators

The multiplicative operators $\mathbb{C}_\infty(\nVdash) \to \mathbb{C}_\infty(\nVdash)$

$$\mathbb{Y}_{\urcorner\!\shortmid\!\shortmid}(\vec{\diagdown}) \stackrel{\text{def}}{=} \diagdown^{\urcorner}\,\widetilde{\mathbb{Y}_{\urcorner\!\shortmid\!\shortmid}(\diagdown)} \tag{6.238}$$

are tensor operators of rank k. Examples are

$$\begin{aligned}
\mathbb{Y}_{00} &= \frac{1}{\sqrt{4\pi}} \\
\mathbb{Y}_{1\pm1} &= \mp\sqrt{\frac{3}{8\pi}} r \sin\theta\, e^{\pm i\phi} = \mp\sqrt{\frac{3}{8\pi}}\,(x_1 \pm i x_2) \\
\mathbb{Y}_{10} &= \sqrt{\frac{3}{4\pi}} r\cos\theta = \sqrt{\frac{3}{4\pi}}\, x_3 \\
\mathbb{Y}_{2\pm2} &= \frac{1}{4}\sqrt{\frac{15}{2\pi}} r^2 \sin^2\theta\, e^{\pm 2i\phi} = \frac{1}{4}\sqrt{\frac{15}{2\pi}}\,(x_1 \pm i x_2)^2 \\
\mathbb{Y}_{2\pm1} &= \mp\sqrt{\frac{15}{8\pi}} r^2 \sin\theta\, e^{\pm i\phi} = \mp\sqrt{\frac{15}{8\pi}}\,(x_1 \pm i x_2)\, x_3 \\
\mathbb{Y}_{20} &= \sqrt{\frac{5}{4\pi}} r^2 \left(\frac{3}{2}\cos^2\theta - \frac{1}{2}\right) = \sqrt{\frac{5}{4\pi}}\left(\frac{3}{2}x_3^2 - \frac{r^2}{2}\right)
\end{aligned} \tag{6.239}$$

These operators can be expressed in terms of the coordinates x_1, x_2, x_3. The fact that these operators form tensor operators of rank 1, 2, 3 follows from the transformation properties of the spherical harmonics derived in Section 1.3.

Exercise 6.7.1: Show that the following set of spin operators

$$\begin{aligned}
T_{00} &= 1 \\
T_{1\pm1} &= \mp\frac{1}{\sqrt{2}}\, S_\pm \\
T_{10} &= S_3 \\
T_{2\pm2} &= \left(S^\pm\right)^2 \\
T_{2\pm1} &= \mp(S_3 S^\pm + S^\pm S_3) \\
T_{20} &= \sqrt{\frac{2}{3}}\,(3S_3^2 - S^2)
\end{aligned}$$

are tensor operators of rank 0, 1, 2.

Exercise 6.7.2: Express the transformed versions of the following operators (a) $T = x_1^2 - x_2^2$ and (b) $S_2 S_3$ in terms of Wigner rotation matrices and untransformed operators.

For the following it is important to note that the rotation matrix elements (6.237) do not require that the rotation $\mathcal{R}^{-1}(\vec{\vartheta})$ is expressed in terms of Euler angles according to (5.203), but rather any rotation and any parametrization can be assumed. In fact, we will assume presently that the rotation is chosen as follows

$$\mathcal{R}(\vec{\vartheta}) = \exp\left(\vartheta_+ L_+ + \vartheta_- L_- + \vartheta_3 L_3\right) \tag{6.240}$$

where we have defined $\vartheta_\pm = \frac{1}{2}(\vartheta_1 \mp i\vartheta_2)$ and $L_\pm = L_1 \pm i L_2$. This choice of parametrization allows us to derive conditions which are equivalent to the property (6.235 - 6.237), but are far easier to ascertain for any particular operator.

The conditions can be derived if we consider the property (6.235 - 6.237) for transformations characterized by infinitesimal values of $\vartheta_+, \vartheta_-, \vartheta_3$. We first consider a rotation with $\vec{\vartheta} = (\vartheta_+, 0, 0)^T$ in case of small ϑ_+. The property (6.235 - 6.237) yields first

$$\mathcal{R}(\vec{\vartheta})\, T_{kq}\, \mathcal{R}^{-1}(\vec{\vartheta}) \;=\; \sum_{q'=-k}^{k} \langle kq'|\mathcal{R}(\vec{\vartheta})|kq\rangle\, T_{kq'}\;. \tag{6.241}$$

Using $\mathcal{R}(\vec{\vartheta}) = \mathbb{1} + \vartheta_+ L_+ + O(\vartheta_+^2)$ this equation can be rewritten neglecting terms of order $O(\vartheta_+^2)$

$$(\mathbb{1} + \vartheta_+ L_+)\, T_{kq}\, (\mathbb{1} - \vartheta_+ L_+) \;=\; \sum_{q'=-k}^{k} \langle kq'|\mathbb{1} + \vartheta_+ L_+|kq\rangle\, T_{kq'}\;. \tag{6.242}$$

from which follows by means of $\langle kq'|\mathbb{1}|kq\rangle = \delta_{qq'}$ and by subtracting T_{kq} on both sides of the equation

$$\vartheta_+\, [L_+, T_{kq}] \;=\; \sum_{q'=-k}^{k} \vartheta_+ \langle kq'|L_+|kq\rangle\, T_{kq'}\;. \tag{6.243}$$

From (5.80) follows $\langle kq'|L_+|kq\rangle = -i\sqrt{(k+q+1)(k-q)}\delta_{q'\,q+1}$ and, hence,

$$[L_+, T_{kq}] \;=\; -iT_{k\,q+1}\,\sqrt{(k+q+1)(k-q)}\;. \tag{6.244}$$

Similar equations can be derived for infinitesimal rotations of the form $\vec{\vartheta} = (0, \vartheta_-, 0)^T$, $(0, 0, \vartheta_3)^T$. Expressing the results in terms of the angular momentum operators J_+, J_-, J_3 yields

$$[J_+, T_{kq}] \;=\; \hbar\, T_{kq+1}\,\sqrt{(k+q+1)(k-q)} \tag{6.245}$$

$$[J_-, T_{kq}] \;=\; \hbar\, T_{kq-1}\,\sqrt{(k+q)(k-q+1)} \tag{6.246}$$

$$[J_3, T_{kq}] \;=\; \hbar\, q\, T_{kq}\;. \tag{6.247}$$

These properties often can be readily demonstrated for operators and the transformation properties (6.235 - 6.237) be assumed then.

Exercise 6.7.3: Derive Eqs. (6.246, 6.247).
Exercise 6.7.4: Is the 1-particle Hamiltonian

$$H \;=\; -\frac{\hbar^2}{2m}\nabla^2 \;+ V(|\vec{r}|) \tag{6.248}$$

a tensor operator?
Exercise 6.7.5: Consider a system of two spin-$\frac{1}{2}$ particles for which the first spin is described by the operator $\vec{S}^{(1)}$ and the second spin by the operator $\vec{S}^{(2)}$. Show that $\vec{S}^{(1)} \cdot \vec{S}^{(2)}$ is a tensor operator of rank 0 in the space of the products of the corresponding spin states $|\frac{1}{2}m_1\rangle_{(1)}|\frac{1}{2}m_2\rangle_{(2)}$. For this purpose state first the proper rotation operator $\mathcal{R}(\vec{\vartheta})$ and note then that $\vec{S}^{(1)} \cdot \vec{S}^{(2)}$ commutes with the generators of the rotation of $|\frac{1}{2}m_1\rangle_{(1)}|\frac{1}{2}m_2\rangle_{(2)}$.
Exercise 6.7.6: Form tensor operators of rank 1 in terms of the three components of ∇ acting on the space of 1-particle wave functions.

6.8 Wigner-Eckart Theorem

A second important property of tensor operators T_{kq} beside (6.235 - 6.237) is that their matrix elements $\langle \ell_1 m_1 \gamma_1 | T_{kq} | \ell_2 m_2 \gamma_2 \rangle$ obey simple relationships expressed in terms of Clebsch-Gordan coefficients. $|\ell_1 m_1 \gamma_1\rangle$ denotes an angular momentum (spin) state, possibly the total angular momentum–spin state of a compositie system, which is characterized also by a set of other quantum numbers γ_1 which are not affected by the rotational transformation $\mathcal{R}(\vec{\vartheta})$. The relationships among the matrix elements $\langle \ell_1 m_1 \gamma_1 | T_{kq} | \ell_2 m_2 \gamma_2 \rangle$ are stated by the Wigner–Eckart theorem which we will derive now. Starting point of the derivation is the fact that the states $T_{kq}|\ell_2 m_2 \gamma_2\rangle$ behave like angular momentum states of a composite system of two particles each carrying angular momentum or spin, i.e. behave like $|kq\rangle|\ell_1 m_1\rangle$. To prove this we consider the transformation of $T_{kq}|\ell_2 m_2 \gamma_2\rangle$

$$\begin{aligned} \mathcal{R}(\vec{\vartheta})\, T_{kq}\, |\ell_2 m_2 \gamma_2\rangle &= \mathcal{R}(\vec{\vartheta})\, T_{kq}\, \mathcal{R}^{-1}(\vec{\vartheta})\, \mathcal{R}(\vec{\vartheta})\, |\ell_2 m_2 \gamma_2\rangle \\ &= \sum_{q'\, m_2'} \mathcal{D}^{(k)}_{q'q} T_{kq'}\, \mathcal{D}^{(\ell_2)}_{m_2' m_2} |\ell_2 m_2' \gamma_2\rangle \end{aligned} \tag{6.249}$$

which demonstrates, in fact, the stated property. One can, hence, construct states $\Phi_{\ell_1 m_1}(k, \ell_2|\gamma_2)$ which correspond to total angular momentum states. These states according to (6.18) are defined through

$$\Phi_{\ell_1 m_1}(k, \ell_2|\gamma_2) \;=\; \sum_{q, m_2} (\ell_1 m_1 | kq\, \ell_2 m_2)\, T_{kq}\, |\ell_2 m_2 \gamma_2\rangle \;. \tag{6.250}$$

We want to show now that these states are eigenstates of $J^2 = J_1^2 + J_2^2 + J_3^2$ and of J_3 where J_1, J_2, J_3 are the generators of the rotation $\mathcal{R}(\vec{\vartheta})$. Before we proceed we like to point out that the states $|\ell_2 m_2 \gamma_2\rangle$ are also eigenstates of J^2, J_3, i.e.

$$J^2\, |\ell_2 m_2 \gamma_2\rangle \;=\; \hbar^2 \ell_2(\ell_2+1)\, |\ell_2 m_2 \gamma_2\rangle \quad ; \quad J_3\, |\ell_2 m_2 \gamma_2\rangle \;=\; \hbar\, m_2\, |\ell_2 m_2 \gamma_2\rangle \;. \tag{6.251}$$

The corresponding property for $\Phi_{\ell_1 m_1}(k, \ell_2|\gamma_2)$ can be shown readily as follows using (6.247), $J_3|\ell_2 m_2 \gamma_2\rangle = \hbar\, m_2 \gamma_2 |\ell_2 m_2\rangle$ and the property (6.21) of Clebsch-Gordan coefficients

$$\begin{aligned} J_3\, \Phi_{\ell_1 m_1}(k, \ell_2|\gamma_2) &= \sum_{q, m_2} (\ell_1 m_1 | kq\, \ell_2 m_2)\, (\underbrace{J_3 T_{kq} - T_{kq} J_3}_{=\hbar q T_{kq}} + T_{kq} J_3)\, |\ell_2 m_2 \gamma_2\rangle \\ &= \hbar \sum_{q, m_2} \underbrace{(\ell_1 m_1 | kq\, \ell_2 m_2)}_{\sim \delta_{m_1\, q+m_2}} (q + m_2)\, T_{kq}\, |\ell_2 m_2 \gamma_2\rangle \\ &= \hbar\, m_1 \sum_{q, m_2} (\ell_1 m_1 | kq\, \ell_2 m_2)\, T_{kq}\, |\ell_2 m_2 \gamma_2\rangle \end{aligned} \tag{6.252}$$

Similarly, one can show that $\Phi_{\ell_1 m_1}(k, \ell_2|\gamma_2)$ is an eigenstate of J^2 with eigenvalue $\hbar^2 \ell_1(\ell_1+1)$. The hermitian property of J_3 and J^2 implies that states $\Phi_{\ell_1 m_1}(k, \ell_2|\gamma_2)$ are orthogonal to $|\ell_1' m_1'\rangle$ in case of different quantum numbers ℓ_1, m_1, i.e.

$$\langle \ell_1' m_1' | \Phi_{\ell_1 m_1}(k, \ell_2|\gamma_2) \;=\; C\, \delta_{\ell_1 \ell_1'} \delta_{m_1 m_1'} \tag{6.253}$$

Exercise 6.8.1: Show that $\Phi_{\ell_1 m_1}(k, \ell_2|\gamma_2)$ is an eigenstate of J^2 with eigenvalue $\hbar^2 \ell_1(\ell_1+1)$.

In order to evaluate the matrix elements $\langle \ell_1 m_1 \gamma_1 | T_{kq} | \ell_2 m_2 \gamma_2 \rangle$ we express using the equivalent of (6.32)

$$T_{kq} \, |\ell_2 m_2 \gamma_2\rangle \;=\; \sum_{\ell_1 m_1} (\ell_1 m_1 | k q \ell_2 m_2) \, \Phi_{\ell_1 m_1}(k \ell_2 | \gamma_2) \tag{6.254}$$

and orthogonality property (6.253)

$$\langle \ell_1 m_1 \gamma_1 | T_{kq} | \ell_2 m_2 \gamma_2 \rangle \;=\; \langle \ell_1 m_1 \gamma_1 | \Phi_{\ell_1 m_1}(k \ell_2 | \gamma_2) \rangle \, (\ell_1 m_1 | k q \ell_2 m_2) \; . \tag{6.255}$$

At this point the important property can be proven that $\langle \ell_1 m_1 \gamma_1 | \Phi_{\ell_1 m_1}(k \ell_2 | \gamma_2) \rangle$ is independent of m_1, i.e. the matrix elements $\langle \ell_1 m_1 \gamma_1 | T_{kq} | \ell_2 m_2 \gamma_2 \rangle$ can be reduced to an m_1–independent factor, its m_1–dependence being expressed solely through a Clebsch-Gordan coefficient. To prove this property we consider $\langle \ell_1 m_1 \gamma_1 | \Phi_{\ell_1 m_1}(k \ell_2 | \gamma_2) \rangle$ for a different m_1 value, say $m_1 + 1$. Using

$$|\ell_1 m_1 + 1 \gamma_1\rangle \;=\; \frac{1}{\sqrt{(\ell_1 + m_1 + 1)(\ell_1 - m_1)}} \, J_+ |\ell_1 m_1 \gamma_1\rangle \tag{6.256}$$

and noting that the operator adjoint to J_+ is J_-, one obtains

$$\begin{gathered} \langle \ell_1 m_1 + 1 \gamma_1 | \Phi_{\ell_1 m_1 + 1}(k \ell_2 | \gamma_2) \;= \\ \tfrac{1}{\sqrt{(\ell_1 + m_1 + 1)(\ell_1 - m_1)}} \, \langle \ell_1 m_1 \gamma_1 | J_- \Phi_{\ell_1 m_1 + 1}(k \ell_2 | \gamma_2) \rangle = \\ \langle \ell_1 m_1 \gamma_1 | \Phi_{\ell_1 m_1}(k \ell_2 | \gamma_2) \rangle \end{gathered} \tag{6.257}$$

which establishes the m_1–independence of $\langle \ell_1 m_1 \gamma_1 | \Phi_{\ell_1 m_1}(k \ell_2 | \gamma_2) \rangle$. In order to express the m_1–independence explicitly we adopt the following notation

$$\langle \ell_1 m_1 \gamma_1 | \Phi_{\ell_1 m_1}(k \ell_2 | \gamma_2) \rangle \;=\; (-1)^{k - \ell_2 + \ell_1} \frac{1}{\sqrt{2\ell_1 + 1}} \, \langle \ell_1, \gamma_1 || T_k || \ell_2, \gamma_2 \rangle \; . \tag{6.258}$$

We can then finally express the matrix elements of the tensor operators T_{kq} as follows

$$\begin{gathered} \langle \ell_1 m_1, \gamma_1 | T_{kq} | \ell_2 m_2, \gamma_2 \rangle \;= \\ (\ell_1 m_1 | k q \ell_2 m_2) \, (-1)^{k - \ell_2 + \ell_1} \tfrac{1}{\sqrt{2\ell_1 + 1}} \, \langle \ell_1, \gamma_1 || T_k || \ell_2, \gamma_2 \rangle \end{gathered} \tag{6.259}$$

The socalled reduced matrix element $\langle \ell_1, \gamma_1 || T_k || \ell_2, \gamma_2 \rangle$ is determined by applying (6.259) to a combination of magnetic quantum numbers m_1', q', m_2', e.g. $m_1' = q' = m_2' = 0$, for which the l.h.s. can be evaluated as easily as possible. One can then evaluate also the corresponding Clebsch-Gordan coefficient $(\ell_1 m_1' | k q' \ell_2 m_2')$ and determine

$$\langle \ell_1, \gamma_1 || T_k || \ell_2, \gamma_2 \rangle \;=\; \sqrt{2\ell_1 + 1} \frac{\langle \ell_1 m_1', \gamma_1 | T_{kq'} | \ell_2 m_2', \gamma_2 \rangle}{(-1)^{k - \ell_2 + \ell_1} (\ell_1 m_1' | k q' \ell_2 m_2')} \tag{6.260}$$

Exercise 6.8.2: Determine the matrix elements of the gradient operator ∇ of the type

$$\int d^3 r F(\vec{r}) \nabla G(\vec{r}) \tag{6.261}$$

when the functions $F(\vec{r})$ and $G(\vec{r})$ are of the type $f(r)Y_{\ell m}(\hat{r})$. For this purpose relate ∇ to a tensor operator T_{1q}, evaluate the matrix for $m_1 = q = m_2 = 0$ using

$$\cos\theta\, Y_{\ell m}(\theta,\phi) = \sqrt{\tfrac{(\ell+1-m)(\ell+1+m)}{(2\ell+1)(2\ell+3)}}\, Y_{\ell+1m}(\theta,\phi) + \sqrt{\tfrac{(\ell-m)(\ell+m)}{(2\ell-1)(2\ell+1)}}\, Y_{\ell-1m}(\theta,\phi)$$

$$\sin\theta\, Y_{\ell m}(\theta,\phi) = \tfrac{\ell(\ell+1)}{\sqrt{(2\ell+1)(2\ell+3)}}\, Y_{\ell+1m}(\theta,\phi) - \tfrac{\ell(\ell-1)}{\sqrt{2\ell-1)(2\ell+1)}}\, Y_{\ell-1m}(\theta,\phi)$$

and express the remaining matrix elements using the Wigner–Eckart theorem. (The necessary evaluations are cumbersome, but a very useful exercise!)

Chapter 7

Motion in Spherically Symmetric Potentials

We describe in this section the stationary bound states of quantum mechanical particles in spherically symmetric potentials $V(r)$, i.e., in potentials which are solely a function of r and are independent of the angles θ, ϕ. Four examples will be studied. The first potential

$$V(r) \;=\; \begin{cases} 0 & 0 \le r \le R \\ \infty & r > R \end{cases} \tag{7.1}$$

confines a freely moving particle to a spherical box of radius R. The second potential is of the square well type

$$V(r) \;=\; \begin{cases} -V_o & 0 \le r \le R \\ 0 & r > R \end{cases} \quad . \tag{7.2}$$

The third potential

$$V(r) \;=\; \frac{1}{2} m\omega^2 r^2 \tag{7.3}$$

describes an isotropic harmonic oscillator . The fourth potential

$$V(r) \;=\; -\,\frac{Z\,e^2}{r} \tag{7.4}$$

governs the motion of electrons in hydrogen-type atoms.
Potential (7.4) is by far the most relevant of the four choices. It leads to the stationary electronic states of the hydrogen atom ($Z = 1$). The corresponding wave functions serve basis functions for multi-electron systems in atoms, molecules, and crystals. The potential (7.3) describes the motion of a charge in a uniformely charged sphere and can be employed to describe the motion of protons and neutrons in atomic nuclei[1]. The potentials (7.1, 7.2) serve as schematic descriptions of quantum particles, for example, in case of the so-called bag model of hadronic matter.

[1]See, for example, *Simple Models of Complex Nuclei / The Shell Model and the Interacting Boson Model* by I. Talmi (Harwood Academic Publishers, Poststrasse 22, 7000 Chur, Switzerland, 1993)

7.1 Radial Schrödinger Equation

A classical particle moving in a potential $V(r)$ is governed by the Newtonian equation of motion

$$m\dot{\vec{v}} \;=\; -\,\hat{e}_r\,\partial_r V(r)\;. \tag{7.5}$$

In the case of an angular independent potential angular momentum $\vec{J} = m\,\vec{r}\times\vec{v}$ is a constant of motion. In fact, the time variation of $\vec{J}$ can be written, using (7.5) and $\dot{\vec{r}} = \vec{v}$,

$$\frac{d}{dt}\vec{J} \;=\; \vec{v}\times m\vec{v} \;+\; \vec{r}\times m\dot{\vec{v}} \;=\; 0 \;+\; \vec{r}\times\hat{e}_r\,(-\partial_r V(r)) \;=\; 0\;. \tag{7.6}$$

Since $\dot{J}_k$ is also equal to the poisson bracket $\{H, J_k\}$ where H is the Hamiltonian, one can conclude

$$\dot{J}_k \;=\; \{H, J_k\} \;=\; 0\;, \quad k \;=\; 1,2,3. \tag{7.7}$$

The correspondence principle dictates then for the quantum mechanical Hamiltonian operator $\hat{H}$ and angular momentum operators $\mathcal{J}_k$

$$[\hat{H},\,\mathcal{J}_k] \;=\; 0\;, \quad k \;=\; 1,2,3\;. \tag{7.8}$$

This property can also be proven readily employing the expression of the kinetic energy operator [c.f. (5.99)]

$$-\frac{\hbar^2}{2m}\nabla^2 \;=\; -\frac{\hbar^2}{2m}\,\frac{1}{r}\,\partial_r^2\,r \;+\; \frac{\mathcal{J}^2}{2mr^2}\;, \tag{7.9}$$

the expressions (5.85– 5.87) for $\mathcal{J}_k$ as well as the commutation property (5.61) of $\mathcal{J}^2$ and $\mathcal{J}_k$. Accordingly, stationary states $\psi_{E,\ell,m}(\vec{r})$ can be chosen as simultaneous eigenstates of the Hamiltonian operator $\hat{H}$ as well as of $\mathcal{J}^2$ and $\mathcal{J}_3$, i.e.,

$$\hat{H}\,\psi_{E,\ell,m}(\vec{r}) \;=\; E\,\psi_{E,\ell,m}(\vec{r}) \tag{7.10}$$

$$\mathcal{J}^2\psi_{E,\ell,m}(\vec{r}) \;=\; \hbar^2\ell(\ell+1)\,\psi_{E,\ell,m}(\vec{r}) \tag{7.11}$$

$$\mathcal{J}_3\psi_{E,\ell,m}(\vec{r}) \;=\; \hbar\,m\,\psi_{E,\ell,m}(\vec{r})\;. \tag{7.12}$$

In classical mechanics one can exploit the conservation of angular momentum to reduce the equation of motion to an equation governing solely the radial coordinate of the particle. For this purpose one concludes first that the conservation of angular momentum $\vec{J}$ implies a motion of the particle confined to a plane. Employing in this plane the coordinates r, θ for the distance from the origin and for the angular position, one can state

$$J \;=\; m\,r^2\,\dot{\theta}\;. \tag{7.13}$$

The expression of the kinetic energy

$$\frac{\vec{p}^{\,2}}{2m} \;=\; \frac{1}{2}m\,\dot{r}^2 \;+\; \frac{1}{2}m\,r^2\dot{\theta}^2 \;=\; \frac{1}{2}m\,\dot{r}^2 \;+\; \frac{J^2}{2mr^2} \tag{7.14}$$

and conservation of energy yield

$$\frac{1}{2}m\,\dot{r}^2 \;+\; \frac{J^2}{2mr^2} \;+\; V(r) \;=\; E\;. \tag{7.15}$$

This is a differential equation which governs solely the radial coordinate. It can be solved by integration of

$$\dot{r} \;=\; \pm \left\{ \frac{2}{m} \left[E \,-\, V(r) \,-\, \frac{J^2}{2mr^2} \right] \right\}^{\frac{1}{2}} \;. \tag{7.16}$$

Once, $r(t)$ is determined the angular motion follows from (7.13), i.e., by integration of

$$\dot{\theta} \;=\; \frac{J}{mr^2(t)} \,. \tag{7.17}$$

In analogy to the classical description one can derivem, in the present case, for the wave function of a quantum mechanical particle a differential equation which governs solely the r-dependence. Employing the kinetic energy operator in the form (7.9) one can write the stationary Schrödinger equation (7.10), using (7.11),

$$\left[-\frac{\hbar^2}{2m} \frac{1}{r} \partial_r^2 \, r \;+\; \frac{\mathcal{J}^2}{2mr^2} \;+\; V(r) \;-\; E_{\ell,m} \right] \psi_{E,\ell,m}(\vec{r}) \;=\; 0 \,. \tag{7.18}$$

Adopting for $\psi_{E,\ell,m}(\vec{r})$ the functional form

$$\psi_{E,\ell,m}(\vec{r}) \;=\; v_{E,\ell,m}(r)\, Y_{\ell m}(\theta,\phi) \tag{7.19}$$

where $Y_{\ell m}(\theta,\phi)$ are the angular momentum eigenstates defined in Section 5.4, equations (7.11, 7.12) are obeyed and one obtains for (7.10)

$$\left[-\frac{\hbar^2}{2m} \frac{1}{r} \partial_r^2 \, r \;+\; \frac{\hbar^2 \ell(\ell+1)}{2mr^2} \;+\; V(r) \;-\; E_{\ell,m} \right] v_{E,\ell,m}(r) \;=\; 0 \,. \tag{7.20}$$

Since this equation is independent of the quantum number m we drop the index m on the radial wave function $v_{E,\ell,m}(r)$ and $E_{\ell,m}$.
One can write (7.20) in the form of the one-dimensional Schrödinger equation

$$\left[-\frac{\hbar^2}{2m} \partial_r^2 \;+\; V_{\text{eff}}(r) \;-\; E \right] \phi_E(r) \;=\; 0 \,. \tag{7.21}$$

where

$$V_{\text{eff}}(r) \;=\; V(r) \;+\; \frac{\hbar^2 \ell(\ell+1)}{2mr^2} \tag{7.22}$$

$$\phi_E(r) \;=\; r\, v_{E,\ell,m}(r) \,. \tag{7.23}$$

This demonstrates that the function $r\, v_{E,\ell,m}(r)$ describes the radial motion as a one-dimensional motion in the interval $[0,\, \infty[$ governed by the effective potential (7.22) which is the original potential $V(r)$ with an added rotational barrier potential $\frac{\hbar^2 \ell(\ell+1)}{2mr^2}$. This barrier, together with the original potential, can exclude particles from the space with small r values, but can also trap particles in the latter space giving rise to strong scattering resonances (see Section **??**).
Multiplying (7.20) by $-2mr/\hbar^2$ yields the so-called *radial Schrödinger equation*

$$\left[\partial_r^2 \;-\; \frac{\ell(\ell+1)}{r^2} \;-\; U(r) \;-\; \kappa_\ell^2 \right] r\, v_{\kappa,\ell}(r) \;=\; 0 \,. \tag{7.24}$$

where we defined

$$U(r) \;=\; -\frac{2m}{\hbar^2}\,V(r) \tag{7.25}$$

$$\kappa_\ell^2 \;=\; -\frac{2m}{\hbar^2}\,E_\ell \tag{7.26}$$

In case $E < 0$, κ assumes real values. We replaced in (7.20) the index E by the equivalent index κ.

Boundary Conditions

In order to solve (7.20) one needs to specify proper boundary conditions[2]. For $r \to 0$ one may assume that the term $\ell(\ell+1)/r^2$ becomes larger than the potential $U(r)$. In this case the solution is governed my

$$\left[\partial_r^2 \;-\; \frac{\ell(\ell+1)}{r^2}\right] r\,v_{\kappa,\ell}(r) \;=\; 0\,, \quad r \to 0 \tag{7.27}$$

and, accordingly, assumes the general form

$$r\,v_{\kappa,\ell}(r) \;\sim\; A\,r^{\ell+1} \;+\; B\,r^{-\ell} \tag{7.28}$$

or

$$v_{\kappa,\ell}(r) \;\sim\; A\,r^{\ell} \;+\; B\,r^{-\ell-1} \tag{7.29}$$

Only the first term is admissable. This follows for $\ell > 0$ from consideration of the integral which measures the total particle density. The radial part of this integral is

$$\int_0^\infty dr\,r^2\,v_{\kappa,\ell}^2(r) \tag{7.30}$$

and, hence, the term $Br^{-(\ell+1)}$ would contribute

$$|B|^2 \int_0^\epsilon dr\,r^{-2\ell} \tag{7.31}$$

which, for $\ell > 0$ is not integrable. For $\ell = 0$ the contribution of $Br^{-(\ell+1)}$ to the complete wave function is, using the expression (5.182) for Y_{00},

$$\psi_{E,\ell,m}(\vec{r}) \;\sim\; \frac{B}{\sqrt{4\pi}|\vec{r}|}\,. \tag{7.32}$$

The total kinetic energy resulting from this contribution is, according to a well-known result in *Classical Electromagnetism*[3],

$$\frac{\hbar^2}{2m}\,\nabla^2\psi_{E,\ell,m}(\vec{r}) \;\sim\; \sqrt{4\pi}\,B\,\delta(x_1)\delta(x_2)\delta(x_3)\,. \tag{7.33}$$

[2] A detailed discussion of the proper boundary conditions, in particular, at $r = 0$ is found in the excellent monographs *Quantum Mechanics I, II* by A. Galindo and P. Pascual (Springer, Berlin, 1990)

[3] We refer here to the fact that the function $\Phi(\vec{r}) = 1/r$ is the solution of the Poisson equation $\nabla^2 = -4\pi\delta(x)\delta(y)\delta(z)$; see, for example, "Classical Electrodynamics, 2nd Ed." by J.D. Jackson (John Wiley, New York, 1975).

Since there is no term in the stationary Schrödinger equation which could compensate this δ-function contribution we need to postulate that the second term in (7.29) is not permissible. One can conclude that the solution of the radial Schrödinger equation must obey

$$r\, v_{\kappa,\ell}(r) \;\to\; 0 \quad \text{for } r \to 0\,. \tag{7.34}$$

The boundary conditions for $r \to \infty$ are governed by two terms in the radial Schrödinger equation, namely,

$$\left[\,\partial_r^2 \;-\; \kappa_\ell^2\,\right] r\, v_{\kappa,\ell}(r) \;=\; 0 \quad \text{for } r \to \infty\;. \tag{7.35}$$

We have assumed here $\lim_{r\to\infty} V(r) \;=\; 0$ which is the convention for potentials. The solution of this equation is

$$r\, v_{\kappa,\ell}(r) \;\sim\; A\, e^{-\kappa r} \;+\; B\, e^{+\kappa r} \quad \text{for } r \to \infty\,. \tag{7.36}$$

For bound states κ is real and, hence, the second contribution is not permissible. We conclude, therefore, that the asymptotic boundary condition for the solution of the radial Schrödinger equation (7.20) is

$$r\, v_{\kappa,\ell}(r) \;\sim\; e^{-\kappa r} \quad \text{for } r \to \infty\,. \tag{7.37}$$

Degeneracy of Energy Eigenvalues

We have noted above that the differential operator appearing on the l.h.s. of the in radial Schrödinger equation (7.24) is independent of the angular momentum quantum number m. This implies that the energy eigenvalues associated with stationary bound states of radially symmetric potentials with identical ℓ, but different m quantum number, assume the same values. This behaviour is associated with the fact that any rotational transformation of a stationary state leaves the energy of a stationary state unaltered. This property holds since (7.8) implies

$$[\hat{H},\, \exp(-\frac{i}{\hbar}\vartheta\cdot\vec{\mathcal{J}})\,] \;=\; 0\,. \tag{7.38}$$

Applying the rotational transformation $\exp(-\frac{i}{\hbar}\vartheta\cdot\vec{\mathcal{J}})$ to (7.10) yields then

$$\hat{H}\, \exp(-\frac{i}{\hbar}\vartheta\cdot\vec{\mathcal{J}})\, \psi_{E,\ell,m}(\vec{r}) \;=\; E\, \exp(-\frac{i}{\hbar}\vartheta\cdot\vec{\mathcal{J}})\, \psi_{E,\ell,m}(\vec{r})\,, \tag{7.39}$$

i.e., any rotational transformation produces energetically degenerate stationary states. One might also apply the operators $\mathcal{J}_\pm \;=\; \mathcal{J}_1 \pm i\mathcal{J}_2$ to (7.10) and obtain for $-\ell < m < \ell$

$$\hat{H}\, \mathcal{J}_\pm\, \psi_{E,\ell,m}(\vec{r}) \;=\; E\, \mathcal{J}_\pm\, \psi_{E,\ell,m}(\vec{r})\,, \tag{7.40}$$

which, together with the identities (5.172, 5.173), yields

$$\hat{H}\, \psi_{E,\ell,m\pm1}(\vec{r}) \;=\; E\, \psi_{E,\ell,m\pm1}(\vec{r}) \tag{7.41}$$

where E is the same eigenvalue as in (7.40). One expect, therefore, that the stationary states for spherically symmetric potentials form groups of $2\ell+1$ energetically degenerate states, so-called multiplets where $\ell \;=\; 0, 1, 2, \ldots$. Following a convention from atomic spectroscopy, one refers to the multiplets with $\ell \;=\; 0, 1, 2, 3$ as the s, p, d, f-multipltes, respectively.
In the remainder of this section we will solve the radial Schrödinger equation (7.20) for the potentials stated in (7.1–7.4). We seek to describe bound states for the particles, i.e., states with $E < 0$. States with $E > 0$, which play a key role in scattering processes, will be described in Section **??**.

7.2 Free Particle Described in Spherical Coordinates

We consider first the case of a particle moving in a force-free space described by the potential

$$V(r) \equiv 0 . \tag{7.42}$$

The stationary Schrödinger equation for this potential reads

$$\left[-\frac{\hbar^2}{2m} \nabla^2 - E \right] \psi_E(\vec{r}) = 0 . \tag{7.43}$$

Stationary States Expressed in Cartesian Coordinates The general solution of (7.43), as expressed in (3.74–3.77), is

$$\psi(\vec{k}|\vec{r}) = N e^{i\vec{k}\cdot\vec{r}} \tag{7.44}$$

where

$$E = \frac{\hbar^2 k^2}{2m} \geq 0 . \tag{7.45}$$

The possible energies can assume continuous values. N in (7.44) is some suitably chosen normalization constant; the reader should be aware that (7.44) does not represent a localized particle and that the function is not square integrable. One chooses N such that the orthonormality property

$$\int_{\Omega_\infty} d^3r \, \psi^*(\vec{k}'|\vec{r})\psi(\vec{k}|\vec{r}) = \delta(\vec{k}' - \vec{k}) \tag{7.46}$$

holds. The proper normalization constant is $N = (2\pi)^{-3/2}$.
In case of a force-free motion momentum is conserved. In fact, the Hamiltonian in the present case

$$H_o = -\frac{\hbar^2}{2m} \nabla^2 \tag{7.47}$$

commutes with the momentum operator $\hat{\vec{p}} = (\hbar/i)\nabla$ and, accordingly, the eigenfunctions of (7.45) can be chosen as simultaneous eigenfunction of the momentum operator. In fact, it holds

$$\hat{\vec{p}} N e^{i\vec{k}\cdot\vec{r}} = \hbar\vec{k} \, N e^{i\vec{k}\cdot\vec{r}} . \tag{7.48}$$

as one can derive using in (7.48) Cartesian coordinates, i.e.,

$$\nabla = \begin{pmatrix} \partial_1 \\ \partial_2 \\ \partial_3 \end{pmatrix} , \quad \vec{k}\cdot\vec{r} = k_1x_1 + k_2x_2 + k_3x_3 . \tag{7.49}$$

Stationary States Expressed in Spherical Coordinates Rather than specifying energy through $k = |\vec{k}|$ and the direction of the momentum through $\hat{k} = \vec{k}/|\vec{k}|$ one can exploit the fact that the angular momentum operators $\mathcal{J}^2$ and $\mathcal{J}_3$ given in (5.97) and in (5.92), respectively, commute with H_o as defined in (7.47). This latter property follows from (5.100) and (5.61). Accordingly, one can choose stationary states of the free particle which are eigenfunctions of (7.47) as well as eigenfunctions of $\mathcal{J}^2$ and $\mathcal{J}_3$ described in Sect. 5.4.

The corresponding stationary states, i.e., solutions of (7.43) are given by wave functions of the form

$$\psi(k,\ell,m|\vec{r}) \;=\; v_{k,\ell}(r)\,Y_{\ell m}(\theta,\phi) \tag{7.50}$$

where the radial wave functions obeys [c.f. (7.20)]

$$\left[-\frac{\hbar^2}{2m}\,\frac{1}{r}\,\partial_r^2\,r \;+\; \frac{\hbar^2\ell(\ell+1)}{2mr^2} \;-\; E_{\ell,m}\right] v_{k,\ell}(r) \;=\; 0\,. \tag{7.51}$$

Using (7.45) and multiplying (7.20) by $-2mr/\hbar^2$ yields the radial Schrödinger equation

$$\left[\partial_r^2 \;-\; \frac{\ell(\ell+1)}{r^2} \;+\; k^2\right] r\,v_{k,\ell}(r) \;=\; 0\,. \tag{7.52}$$

We want to determine now the solutions $v_{k,\ell}(r)$ of this equation.
We first notice that the solution of (7.52) is actually only a function of kr, i.e., one can write $v_{k,\ell}(r) \;=\; j_\ell(kr)$. In fact, one can readily show, introducing the new variable $z \;=\; kr$, that (7.52) is equivalent to

$$\left[\frac{d^2}{dz^2} \;-\; \frac{\ell(\ell+1)}{z^2} \;+\; 1\right] z\,j_\ell(z) \;=\; 0\,. \tag{7.53}$$

According to the discussion in Sect. 7.1 the regular solution of this equation, at small r, behaves like

$$j_\ell(z) \;\sim\; z^\ell \quad \text{for} \quad r \;\to\; 0\,. \tag{7.54}$$

There exists also a so-called irregular solution of (7.53), denoted by $n_\ell(z)$ which behaves like

$$n_\ell(z) \;\sim\; z^{-\ell-1} \quad \text{for} \quad r \;\to\; 0\,. \tag{7.55}$$

We will discuss further below also this solution, which near $r \;=\; 0$ is inadmissable in a quantum mechanical wave function, but admissable for $r \;\neq\; 0$.
For large z values the solution of (7.53) is governed by

$$\left[\frac{d^2}{dz^2} \;+\; 1\right] z\,j_\ell(z) \;=\; 0 \quad \text{for} \quad r \;\to\; \infty \tag{7.56}$$

the general solution of which is

$$j_\ell(z) \;\sim\; \frac{1}{z}\,\sin(z \;+\; \alpha) \quad \text{for} \quad r \;\to\; \infty \tag{7.57}$$

for some phase α.
We note in passing that the functions $g_\ell(z) \;=\; j_\ell(z), n_\ell(z)$ obey the differential equation equivalent to (7.53)

$$\left[\frac{d^2}{dz^2} \;+\; \frac{2}{z}\,\frac{d}{dz} \;-\; \frac{\ell(\ell+1)}{z^2} \;+\; 1\right] g_\ell(z) \;=\; 0\,. \tag{7.58}$$

Noting that $\sin(z \;+\; \alpha)$ can be written as an infinite power series in z we attempt to express the solution of (7.53) for arbitary z values in the form

$$j_\ell(z) \;=\; z^\ell\,f(z^2)\,, \quad f(z^2) \;=\; \sum_{n=0}^{\infty} a_n\,z^{2n}\,. \tag{7.59}$$

The unknown expansion coefficients can be obtained by inserting this series into (7.53). We have introduced here the assumption that the factor f in (7.59) depends on z^2. This follows from

$$\frac{d^2}{dz^2}\, z^{\ell+1} f(z) \;=\; z^{\ell+1}\, \frac{d^2}{dz^2} f(z) \;+\; 2(\ell+1) z^{\ell} \frac{d}{dz} f \;+\; \ell(\ell+1) z^{\ell-1} f \tag{7.60}$$

from which we can conclude

$$\left(\frac{d^2}{dz^2} \;+\; (\ell+1)\, \frac{2}{z}\, \frac{d}{dz} \;+\; 1 \right)\, f(z) \;=\; 0\,. \tag{7.61}$$

Introducing the new variable $v = z_2$ yields, using

$$\frac{1}{z}\frac{d}{dz} \;=\; 2\frac{d}{dv}\,, \quad \frac{d^2}{dz^2} \;=\; 4v\frac{d^2}{dv^2} \;+\; 2\frac{d}{dv}\,, \tag{7.62}$$

the differential equation

$$\left(\frac{d^2}{dv^2} \;+\; \frac{2\ell+3}{2v}\, \frac{d}{dv} \;+\; \frac{1}{4v} \right)\, f(v) \;=\; 0 \tag{7.63}$$

which is consistent with the functional form in (7.59). The coefficients in the series expansion of $f(z^2)$ can be obtained from inserting $\sum_{n=0}^{\infty} a_n\, z^{2n}$ into (7.63) ($v = z^2$)

$$\sum_n \left(a_n\, n\,(n-1)\, v^{n-2} \;+\; \frac{1}{2}\,(2\ell+3)\, a_n\, v^{n-2} \;+\; \frac{1}{4}\, a_n\, v^{n-1} \right) \;=\; 0 \tag{7.64}$$

Changing the summation indices for the first two terms in the sum yields

$$\sum_n \left(a_{n+1}\, n\,(n-1) \;+\; \frac{1}{2}\,(2\ell+3)\, a_n \;+\; \frac{1}{4}\, a_n \right)\, v^{n-1} \;=\; 0\,. \tag{7.65}$$

In this expression each term $\sim v^{n-1}$ must vanish individually and, hence,

$$a_{n+1} \;=\; -\frac{1}{2}\, \frac{1}{(n+1)\,(2n \,+\, 2\ell \,+\, 3)}\, a_n \tag{7.66}$$

One can readily derive

$$a_1 \;=\; -\frac{1}{2}\frac{1}{1!\,(2\ell+3)}\, a_0\,, \quad a_2 \;=\; \frac{1}{4}\, \frac{1}{2!\,(2\ell+3)(2\ell+5)}\, a_0\,. \tag{7.67}$$

The common factor a_0 is arbitrary. Choosing

$$a_0 \;=\; \frac{1}{1\cdot 3\cdot 5\cdot (2\ell+1)}\,. \tag{7.68}$$

the ensuing functions ($\ell = 0, 1, 2, \ldots$)

$$j_\ell(z) \;=\; \frac{z^\ell}{1\cdot 3\cdot 5\cdots(2\ell+1)} \left[1 \;-\; \frac{\frac{1}{2}z^2}{1!(2\ell+3)} \;+\; \frac{(\frac{1}{2}z^2)^2}{2!(2\ell+3)(2\ell+5)} \;-\; +\cdots \right] \tag{7.69}$$

are called *regular spherical Bessel functions*.

One can derive similarly for the solution (7.55) the series expansion ($\ell = 0, 1, 2, \ldots$)

$$n_\ell(z) \;=\; -\,\frac{1\cdot 3\cdot 5\cdots(2\ell-1)}{z^{\ell+1}}\;\left[1\;-\;\frac{\frac{1}{2}z^2}{1!(1-2\ell)}\;+\;\frac{(\frac{1}{2}z^2)^2}{2!(1-2\ell)(3-2\ell)}\;-+\cdots\right]. \tag{7.70}$$

These functions are called *irregular spherical Bessel functions*.

Exercise 7.2.1: Demonstrate that (7.70) is a solution of (7.52) obeying (7.55).

The Bessel functions (7.69, 7.70) can be expressed through an infinite sum which we want to specify now. For this purpose we write (7.69)

$$\begin{aligned} j_\ell(z) \quad &= \quad \left(\frac{z}{2}\right)^\ell \frac{1}{\frac{1}{2}\cdot\frac{3}{2}\cdot\frac{5}{2}\cdots(\ell+\frac{1}{2})} \\ &\qquad \left[1\;+\frac{\left(\frac{iz}{2}\right)^2}{1!(\ell+\frac{3}{2})}\;+\;\frac{\left(\left(\frac{iz}{2}\right)^4}{2!(\ell+\frac{3}{2})(2\ell+\frac{5}{2})}\;+\cdots\right] \end{aligned} \tag{7.71}$$

The factorial-type products

$$\frac{1}{2}\cdot\frac{3}{2}\cdots\left(\ell+\frac{1}{2}\right) \tag{7.72}$$

can be expressed through the so-called Gamma-function[4] defined through

$$\Gamma(z) \;=\; \int_0^\infty dt\, t^{z-1}\, e^{-t}\,. \tag{7.73}$$

This function has the following properties[5]

$$\begin{aligned} \Gamma(z+1) \quad &= \quad z\,\Gamma(z) && (7.74)\\ \Gamma(n+1) \quad &= \quad n! \quad \text{for } n\in\mathbb{N} && (7.75)\\ \Gamma(\tfrac{1}{2}) \quad &= \quad \sqrt{\pi} && (7.76)\\ \Gamma(z)\,\Gamma(1-z) \quad &= \quad \frac{\pi}{\sin\pi z}\,. && (7.77) \end{aligned}$$

from which one can deduce readily

$$\Gamma(\ell+\tfrac{1}{2}) \;=\; \sqrt{\pi}\cdot\frac{1}{2}\cdot\frac{3}{2}\cdots\left(\ell+\frac{1}{2}\right)\,. \tag{7.78}$$

One can write then

$$j_\ell(z) \;=\; \frac{\sqrt{\pi}}{2}\,\left(\frac{z}{2}\right)^\ell \sum_{n=0}^\infty \frac{\left(\frac{iz}{2}\right)^{2n}}{n!\,\Gamma(n+1+\ell+\frac{1}{2})}\,. \tag{7.79}$$

---

[4]For further details see *Handbook of Mathematical Functions* by M. Abramowitz and I.A. Stegun (Dover Publications, New York)

[5]The proof of (7.74–7.76) is elementary; a derivation of (7.77) can be found in *Special Functions of Mathematical Physics* by A.F. Nikiforov and V.B. Uvarov, Birkhäuser, Boston, 1988)

Similarly, one can express $n_\ell(z)$ as given in (7.70)

$$n_\ell(z) \;=\; -\frac{2^\ell}{\sqrt{\pi}z^{\ell+1}}\,\Gamma(\ell+\frac{1}{2}) \;\left[1 \;+\; \frac{\left(\frac{iz}{2}\right)^2}{1!(\frac{1}{2}-\ell)} \;+\; \frac{\left(\frac{iz}{2}\right)^4}{2!(\frac{1}{2}-\ell)(\frac{3}{2}-\ell)} \;+\cdots\right]. \tag{7.80}$$

Using (7.77) for $z = \ell+\frac{1}{2}$, i.e.,

$$\Gamma(\ell+\tfrac{1}{2}) \;=\; (-1)^\ell\,\frac{\pi}{\Gamma(\frac{1}{2}-\ell)} \tag{7.81}$$

yields

$$\begin{aligned} n_\ell(z) \quad &= \quad (-1)^{\ell+1}\,\sqrt{\pi}\,\frac{2^\ell}{z^{\ell+1}} \;\left[\frac{1}{\Gamma(\frac{1}{2}-\ell)} \;+\; \frac{\left(\frac{iz}{2}\right)^2}{1!\,\Gamma(\frac{1}{2}-\ell)(\frac{1}{2}-\ell)}\right. \\ &\qquad \left. +\,\frac{\left(\frac{iz}{2}\right)^4}{2!\,\Gamma(\frac{1}{2}-\ell)(\frac{1}{2}-\ell)(\frac{3}{2}-\ell)} \;+\cdots\right]. \end{aligned} \tag{7.82}$$

or

$$n_\ell(z) \;=\; (-1)^{\ell+1}\,\frac{\sqrt{\pi}}{2}\,\left(\frac{2}{z}\right)^{\ell+1} \;\sum_{n=0}^{\infty}\,\frac{\left(\frac{iz}{2}\right)^{2n}}{n!\,\Gamma(n+1-\ell-\frac{1}{2})}\,. \tag{7.83}$$

Linear independence of the Regular and Irregular spherical Bessel Functions

We want to demonstrate now that the solutions (7.69) and (7.69) of (7.53) are linearly independent. For this purpose we need to demonstrate that the Wronskian

$$W(j_\ell, n_\ell) \;=\; j_\ell(z)\,\frac{d}{dz}\,n_\ell \;-\; \frac{d}{dz}\,j_\ell(z)\,n_\ell \tag{7.84}$$

does not vanish. Let f_1, f_2 be solutions of (7.53), or equivalently, of (7.58). Using

$$\frac{d^2}{dz^2}\,f_{1,2} \;=\; -\,\frac{2}{z}\,\frac{d}{dz}\,f_{1,2} \;+\; \frac{\ell(\ell+1)}{z^2}\,f_{1,2} \;-\; f_{1,2} \tag{7.85}$$

one can demonstrate the identity

$$\frac{d}{dz}\,W(f_1,f_2) \;=\; -\frac{2}{z}\,W(f_1,f_2) \tag{7.86}$$

This equation is equivalent to

$$\frac{d}{dz}\ln W \;=\; \frac{d}{dz}\ln\frac{1}{z^2} \tag{7.87}$$

the solution of which is

$$\ln W \;=\; \frac{c}{z^2} \tag{7.88}$$

for some constant c. For the case of $f_1 = j_\ell$ and $f_2 = n_\ell$ this constant can be determined using the expansions (7.69, 7.70) keeping only the leading terms. One obtains $c = 1$ and, hence,

$$W(j_\ell, n_\ell) \;=\; \frac{1}{z^2}\,. \tag{7.89}$$

The Wronskian (7.89) doesn not vanish and, therefore, the regular and irregular Bessel functions are linearly independent.

Relationship to Bessel functions

The differential equation (7.58) for the spherical Bessel functions $g_\ell(z)$ can be simplified by seeking the corresponding equation for $G_{\ell+\frac{1}{2}}(z)$ defined through

$$g_\ell(z) \;=\; \frac{1}{\sqrt{z}}\, G_{\ell+\frac{1}{2}}(z). \tag{7.90}$$

Using

$$\frac{d}{dz}\, g_\ell(z) \;=\; \frac{1}{\sqrt{z}}\,\frac{d}{dz}\, G_{\ell+\frac{1}{2}}(z) \;-\; \frac{1}{2z\sqrt{z}}\, G_{\ell+\frac{1}{2}}(z) \tag{7.91}$$

$$\frac{d^2}{dz^2}\, g_\ell(z) \;=\; \frac{1}{\sqrt{z}}\,\frac{d^2}{dz^2}\, G_{\ell+\frac{1}{2}}(z) \;-\; \frac{1}{z\sqrt{z}}\,\frac{d}{dz}\, G_{\ell+\frac{1}{2}}(z) \;+\; \frac{3}{4z^2\sqrt{z}}\, G_{\ell+\frac{1}{2}}(z) \tag{7.92}$$

one is lead to *Bessel's equation*

$$\left[\, \frac{d^2}{dz^2} \;+\; \frac{1}{z}\,\frac{d}{dz} \;-\; \frac{\nu^2}{z^2} \;+\; 1 \,\right] G_\nu(z) \;=\; 0 \tag{7.93}$$

where $\nu = \ell + \frac{1}{2}$. The regular solution of this equation is called the *regular Bessel function*. Its power expansion, using the conventional normalization, is given by [c.f. (7.79)]

$$J_\nu(z) \;=\; \left(\frac{z}{2}\right)^\nu \sum_{n=0}^{\infty} \frac{(iz/2)^{2n}}{n!\,\Gamma(\nu+n+1)}\,. \tag{7.94}$$

One can show that $J_{-\nu}(z)$, defined through (7.94), is also a solution of (7.93). This follows from the fact that ν appears in (7.93) only in the form ν^2. In the present case we consider solely the case $\nu = \ell + \frac{1}{2}$. In this case $J_{-\nu}(z)$ is linearly independent of $J_\nu(z)$ since the Wronskian

$$W(J_{\ell+\frac{1}{2}},\, J_{-\ell-\frac{1}{2}}) \;=\; (-1)^\ell\, \frac{2}{\pi z} \tag{7.95}$$

is non-vanishing. One can relate $J_{\ell+\frac{1}{2}}$ and $J_{-\ell-\frac{1}{2}}$ to the regular and irregular spherical Bessel functions. Comparision with (7.79) and (7.83) shows

$$j_\ell(z) \;=\; \sqrt{\frac{\pi}{2z}}\, J_{\ell+\frac{1}{2}}(z) \tag{7.96}$$

$$n_\ell(z) \;=\; (-1)^{\ell+1} \sqrt{\frac{\pi}{2z}}\, J_{-\ell-\frac{1}{2}}(z)\,. \tag{7.97}$$

These relationships are employed in case that since numerical algorithms provide the Bessel functions $J_\nu(z)$, but not directly the spherical Bessel functions $j_\ell(z)$ and $n_\ell(z)$.

Exercise 7.2.2: Demonstrate that expansion (7.94) is indeed a regular solution of (7.93). Adopt the procedures employed for the function $j_\ell(z)$.
Exercise 7.2.3: Prove the identity (7.95).

Generating Function of Spherical Bessel Functions

The stationary Schrödinger equation of free particles (7.43) has two solutions, namely, one given by (7.44, 7.45) and one given by (7.50). One can expand the former solution in terms of solutions (7.50). For example, in case of a free particle moving along the x_3-axis one expands

$$e^{ik_3x_3} \;=\; \sum_{\ell,m} a_{\ell m}\, j_\ell(kr)\, Y_{\ell m}(\theta,\phi)\;. \tag{7.98}$$

The l.h.s. can be written $\exp(ikr\cos\theta)$, i.e., the wave function does not depend on ϕ. In this case the expansion on the r.h.s. of (7.98) does not involve any non-vanishing m-values since $Y_{\ell m}(\theta,\phi)$ for non-vanishing m has a non-trivial ϕ-dependence as described by (5.106). Since the spherical harmonics $Y_{\ell 0}(\theta,\phi)$, according to (5.178) are given in terms of Legendre polynomials $P_\ell(\cos\theta)$ one can replace the expansion in (7.98) by

$$e^{ik_3x_3} \;=\; \sum_{\ell=0}^{\infty} b_\ell\, j_\ell(kr)\, P_\ell(\cos\theta)\;. \tag{7.99}$$

We want to determine the expansion coefficients b_ℓ.
The orthogonality properties (5.179) yield from (7.99)

$$\int_{-1}^{+1} d\cos\theta\; e^{ikr\cos\theta}\; P_\ell(\cos\theta) \;=\; b_\ell\, j_\ell(kr)\;\frac{2}{2\ell\,+\,1}\;. \tag{7.100}$$

Defining $x \,=\, \cos\theta$, $z \,=\, kr$, and using the Rodrigues formula for Legendre polynomials (5.150) one obtains

$$\int_{-1}^{+1} dx\; e^{izx}\; P_\ell(x) \;=\; \int_{-1}^{+1} dx\; e^{izx}\;\frac{1}{2^\ell\ell!}\,\frac{\partial^\ell}{\partial x^\ell}\,(x^2-1)^\ell\;. \tag{7.101}$$

Integration by parts yields

$$\begin{aligned}\int_{-1}^{+1} dx\; e^{izx}\; P_\ell(x) \quad &=\quad \frac{1}{2^\ell\ell!}\left[e^{izx}\,\frac{d^{\ell-1}}{dx^{\ell-1}}(x^2\,-\,1)^\ell\right]_{-1}^{+1} \\ &\quad -\,\frac{1}{2^\ell\ell!}\int_{-1}^{+1} dx\left(\frac{d}{dx}\,e^{izx}\right)\,\frac{d^{\ell-1}}{dx^{\ell-1}}\,(x^2\,-\,1)^\ell\;.\end{aligned} \tag{7.102}$$

One can show

$$\frac{d^{\ell-1}}{dx^{\ell-1}}(x^2\,-\,1)^\ell \;\sim\; (x^2\,-\,1)\;\times\;\text{polynomial in } x \tag{7.103}$$

and, hence, the surface term $\sim [\cdots]_{-1}^{+1}$ vanishes. This holds for ℓ consecutive integrations by part and one can conclude

$$\begin{aligned}\int_{-1}^{+1} dx\; e^{izx}\; P_\ell(x) \quad &=\quad \frac{(-1)^\ell}{2^\ell\ell!}\int_{-1}^{+1} dx\;(x^2\,-\,1)^\ell\,\frac{d^\ell}{dx^\ell}\,e^{izx} \\ &=\; \frac{(iz)^\ell}{2^\ell\ell!}\int_{-1}^{+1} dx\;(1\,-\,x^2)^\ell e^{izx}\;.\end{aligned} \tag{7.104}$$

Comparision with (7.100) gives

$$b_\ell \, j_\ell(kr) \, \frac{2}{2\ell \, + \, 1} \; = \; \frac{(iz)^\ell}{2^\ell \ell!} \int_{-1}^{+1} dx \, (1 \, - \, x^2)^\ell e^{izx} \, . \tag{7.105}$$

This expression allows one to determine the expansion coefficients b_ℓ. The identity (7.105) must hold for all powers of z, in particular, for the leading power x^ℓ [c.f. (7.69)]

$$b_\ell \, \frac{z^\ell}{1 \cdot 3 \cdot 5 \cdots (2\ell + 1)} \, \frac{2}{2\ell \, + \, 1} \; = \; \frac{(iz)^\ell}{2^\ell \ell!} \int_{-1}^{+1} dx \, (1 \, - \, x^2)^\ell \, . \tag{7.106}$$

Employing (5.117) one can write the r.h.s.

$$z^\ell \, i^\ell \, \frac{1}{2^\ell \ell!} \, \frac{(2\ell)!}{[1 \cdot 3 \cdot 5 \cdots (2\ell - 1)]^2} \, \frac{2}{2\ell + 1} \tag{7.107}$$

or

$$i^\ell \, (2\ell + 1) \, z^\ell \, \frac{1}{1 \cdot 3 \cdot 5 \cdots (2\ell + 1)} \, \frac{2}{2\ell + 1} \, \frac{(2\ell)!}{1 \cdot 2 \cdot 3 \cdot 4 \cdots (2\ell - 1) \cdot 2\ell} \tag{7.108}$$

where the last factor is equal to unity. Comparision with the l.h.s. of (7.106) yields finally

$$b_\ell \; = \; i^\ell \, (2\ell + 1) \tag{7.109}$$

or, after insertion into (7.99),

$$e^{ikr \cos\theta} \; = \; \sum_{\ell=0}^{\infty} i^\ell \, (2\ell + 1) \, j_\ell(kr) \, P_\ell(\cos\theta) \, . \tag{7.110}$$

One refers to the l.h.s. as the *generating function of the spherical Bessel functions.*

Integral Representation of Bessel Functions

Combining (7.105) and (7.109) results in the integral representation of $j_\ell(z)$

$$j_\ell(z) \; = \; \frac{(z)^\ell}{2^{\ell+1} \ell!} \int_{-1}^{+1} dx \, (1 \, - \, x^2)^\ell \, e^{izx} \, . \tag{7.111}$$

Employing (7.96) one can express this, using $\nu \, = \, \ell + \frac{1}{2}$,

$$J_\nu(z) \; = \; \frac{1}{\sqrt{\pi}\Gamma(\nu + \frac{1}{2})} \, \left(\frac{z}{2}\right)^\nu \int_{-1}^{+1} dx \, (1 \, - \, x^2)^{\nu - \frac{1}{2}} \, e^{izx} \, . \tag{7.112}$$

We want to consider the expression

$$G_\nu(z) \; = \; a_\nu z^\nu f_\nu(z) \tag{7.113}$$

where we define

$$f_\nu(z) \; = \; \int_C dt \, (1 \, - \, t^2)^{\nu - \frac{1}{2}} \; e^{izt} \, . \tag{7.114}$$

Here C is an integration path in the complex plane with endpoints t_1, t_2. $G_\nu(z)$, for properly chosen endpoints t_1, t_2 of the integration paths C, obeys Bessel's equation (7.93) for arbitrary ν. To prove this we note

$$f'_\nu(z) \;=\; i \int_C dt\, (1 \,-\, t^2)^{\nu-\frac{1}{2}}\; t\, e^{izt}\,. \tag{7.115}$$

Integration by part yields

$$f'_\nu(z) \;=\; -\frac{i}{2\nu+1}\,\Big[(1-t^2)^{\nu+\frac{1}{2}}\, e^{izt}\Big]_{t_1}^{t_2} \;-\; \frac{z}{2\nu+1}\int_C dt\, (1 \,-\, t^2)^{\nu+\frac{1}{2}}\, e^{izt}\,. \tag{7.116}$$

In case that the endpoints of the integration path C satisfy

$$\Big[(1-t^2)^{\nu+\frac{1}{2}}\, e^{izt}\Big]_{t_1}^{t_2} \;=\; 0 \tag{7.117}$$

one can write (7.116)

$$f'_\nu(z) \;=\; -\,\frac{z}{2\nu+1}\int_C dt\, (1 \,-\, t^2)^{\nu-\frac{1}{2}}\, e^{izt} \;+\; \int_C dt\, (1 \,-\, t^2)^{\nu-\frac{1}{2}}\; t^2\, e^{izt} \tag{7.118}$$

or

$$f'_\nu(z) \;=\; -\,\frac{z}{2\nu+1}\,\big[\, f_\nu(z) \;+\; f''_\nu(z)\,\big]\,. \tag{7.119}$$

From this we can conclude

$$f''_\nu(z) \;+\; \frac{2\nu+1}{z}\, f'_\nu(z) \;+\; f_\nu(z) \;=\; 0\,. \tag{7.120}$$

We note that equations (7.114, 7.120) imply also the property

$$(2\nu+1)\, f'_\nu(z) \;+\; z\, f_{\nu+1}(z) \;=\; 0\,. \tag{7.121}$$

Exercise 7.2.4: Prove (7.121).

We can now demonstrate that $G_\nu(z)$ defined in (7.113) obeys the Bessel equation (7.93) as long as the integration path in (7.113) satisfies (7.117). In fact, it holds for the derivatives of $G_\nu(z)$

$$G'_\nu(z) \;=\; \frac{\nu}{z}\, a_\nu z^\nu f_\nu(z) \;+\; a_\nu z^\nu f'_\nu(z) \tag{7.122}$$

$$G''(z) \;=\; \frac{\nu(\nu-1)}{z^2}\, a_\nu z^\nu f_\nu(z) \;+\; \frac{2\nu}{z}\, a_\nu z^\nu f'_\nu(z) \;+\; a_\nu z^\nu f''_\nu(z) \tag{7.123}$$

Insertion of these identities into Bessel's equation leads to a differential equation for $f_\nu(z)$ which is identical to (7.120) such that we can conclude that $G_\nu(z)$ for proper integration paths is a solution of (7.93).
We consider now the functions

$$u^{(j)}(z) \;=\; \frac{1}{\sqrt{\pi}\Gamma(\nu+\frac{1}{2})}\,\Big(\frac{z}{2}\Big)^\nu \int_{C_j} dx\, (1 \,-\, x^2)^{\nu-\frac{1}{2}}\, e^{izx}\,, \quad j \,=\, 1,2,3,4. \tag{7.124}$$

for the four integration paths in the complex plane parametrized as follows through a real path length s

$$C_1(t_1 = -1 \rightarrow t_2 = +1): \qquad t \;=\; s \qquad -1 \leq s \leq 1 \tag{7.125}$$

$$C_2(t_1 = 1 \rightarrow t_2 = 1 + i\,\infty): \qquad t \;=\; 1 + is \qquad 0 \leq s < \infty \tag{7.126}$$

$$C_3(t_1 = 1 + i\,\infty \rightarrow t_2 = -1 + i\,\infty): \qquad t \;=\; 1 + is \qquad -1 \leq s \leq 1 \tag{7.127}$$

$$C_4(t_1 = -1 + i\,\infty \rightarrow t_1 = -1): \qquad t \;=\; -1 + is \qquad 0 \leq s < \infty \tag{7.128}$$

$C \;=\; C_1 \cup C_2 \cup C_3 \cup C_4$ is a closed path. Since the integrand in (7.124) is analytical in the part of the complex plane surrounded by the path C we conclude

$$\sum_{j=1}^{4} u^{(j)}(z) \;\equiv\; 0 \tag{7.129}$$

The integrand in (7.124) vanishes along the whole path C_3 and, therefore, $u^{(3)}(z) \equiv 0$. Comparision with (7.112) shows $u^{(1)}(z) \;=\; J_\nu(z)$. Accordingly, one can state

$$J_\nu(z) \;=\; -\left[\, u^{(2)}(z) \;+\; u^{(4)}(z) \right] . \tag{7.130}$$

We note that the endpoints of the integration paths C_2 and C_4, for $\mathrm{Re} z \;>\; 0$ and $\nu \in \mathbb{R}$ obey (7.117) and, hence, $u^{(2)}(z)$ and $u^{(2)}(z)$ are both solutions of Bessel's equation (7.93).
Following convention, we introduce the so-called *Hankel functions*

$$H_\nu^{(1)}(z) \;=\; -2\,u^{(2)}(z)\,, \quad H_\nu^{(2)}(z) \;=\; -2\,u^{(4)}(z)\,. \tag{7.131}$$

According to (7.130) holds

$$J_\nu(z) \;=\; \frac{1}{2}\,\left[\, H_\nu^{(1)}(z) \;+\; H_\nu^{(2)}(z) \right] . \tag{7.132}$$

For $H_\nu^{(1)}(z)$ one derives, using $t \;=\; 1 \;+\; is$, $dt \;=\; i\,ds$, and

$$\begin{aligned} (1-t^2)^{\nu-\frac{1}{2}} e^{izt} \;&=\; [(1-t)(1+t)]^{\nu-\frac{1}{2}} e^{izt} \\ &=\; [-is(2+is)]^{\nu-\frac{1}{2}} e^{iz} e^{-zs} \\ =\; e^{i(z-\pi\nu/2+\pi/4)} 2^{\nu-\frac{1}{2}}\, &[s(1+is/2)]^{\nu-\frac{1}{2}}\, e^{-zs}\,, \end{aligned} \tag{7.133}$$

the integral expression

$$H_\nu^{(1)}(z) \;=\; e^{i(z-\pi\nu/2-\pi/4)}\, \frac{\sqrt{2}\,z^\nu}{\sqrt{\pi}\,\Gamma(\nu+\frac{1}{2})} \int_0^\infty ds\,[s(1+is/2)]^{\nu-\frac{1}{2}}\, e^{-zs}\,. \tag{7.134}$$

Similarly, one can derive

$$H_\nu^{(2)}(z) \;=\; e^{-i(z-\pi\nu/2-\pi/4)}\, \frac{\sqrt{2}\,z^\nu}{\sqrt{\pi}\,\Gamma(\nu+\frac{1}{2})} \int_0^\infty ds\,[s(1-is/2)]^{\nu-\frac{1}{2}}\, e^{-zs}\,. \tag{7.135}$$

(7.134) and (7.135) are known as the Poisson integrals of the Bessel functions.

Asymptotic Behaviour of Bessel Functions

We want to obtain now expansions of the Hankel functions $H^{(1,2)}_{\ell+\frac{1}{2}}(z)$ in terms of z^{-1} such that the expansions converge fast asymptotically, i.e., converge fast for $|z| \to \infty$. We employ for this purpose the Poisson integrals (7.134) and (7.135) which read for $\nu = \ell + \frac{1}{2}$

$$H^{(\frac{1}{2})}_{\ell+\frac{1}{2}}(z) \;=\; e^{\pm i[z-(\ell+1)\frac{\pi}{2}]}\,\sqrt{\frac{2z}{\pi}}\,\frac{z^\ell}{\ell!}\,f^{(\frac{1}{2})}_\ell(z) \tag{7.136}$$

where

$$f^{(\frac{1}{2})}_\ell(z) \;=\; \int_0^\infty ds\,[s(s\,\pm\,is/2)]^\ell\,e^{-zs}\;. \tag{7.137}$$

The binomial formula yields

$$f^{(\frac{1}{2})}_\ell(z) \;=\; \sum_{r=0}^{\ell}\binom{\ell}{r}\left(\pm\frac{i}{2}\right)^r\int_0^\infty ds\,s^{\ell+r}\,e^{-zs}\;. \tag{7.138}$$

The formula for the Laplace transform of s^n leads to

$$f^{(\frac{1}{2})}_\ell(z) \;=\; \sum_{r=0}^{\ell}\frac{(\ell+r)!\,\ell!}{r!(\ell-r)!}\,\left(\pm\frac{i}{2}\right)^r\left(\frac{1}{z}\right)^{\ell+r+1} \tag{7.139}$$

and, hence, we obtain

$$H^{(\frac{1}{2})}_{\ell+\frac{1}{2}}(z) \;=\; \sqrt{\frac{2}{\pi z}}\,e^{\pm i[z-(\ell+1)\frac{\pi}{2}]}\,\sum_{r=0}^{\ell}\frac{(\ell+r)!}{r!(\ell-r)!}\,\left(\pm\frac{i}{2z}\right)^r \tag{7.140}$$

Bessel Functions with Negative Index

Since ν enters the Bessel equation (7.93) only as ν^2, $H^{(1)}_\nu(z)$ as well as $H^{(1)}_{-\nu}(z)$ are solutions of this equation. As a second order differential equation the Bessel equation has two linearly independent solutions. For such solutions $g(z)$, $h(z)$ to be linearly independent, the Wronskian $W(g,h)$ must be a non-vanishing function.
For the Wronskian connected with the Bessel equation (7.93) holds the identity

$$W' \;=\; -\frac{1}{z}\,W \tag{7.141}$$

the derivation of which follows the derivation on page 192 for the Wronskian of the radial Schrödinger equation. The general solution of (7.141) is

$$W(z) \;=\; -\frac{c}{z}\;. \tag{7.142}$$

In case of $g(z) = H^{(1)}_{\ell+\frac{1}{2}}(z)$ and $h(z) = H^{(2)}_{\ell+\frac{1}{2}}(z)$ one can identify the constant c by using the leading terms in the expansions (7.140). One obtains

$$W(z) \;=\; -\frac{4i}{\pi z}\;, \tag{7.143}$$

i.e., $H^{(1,2)}_{\ell+\frac{1}{2}}(z)$ are, in fact, linearly independent.
One can then expand

$$H^{(1)}_{-\nu}(z) \;=\; A\, H^{(1)}_{\ell+\frac{1}{2}}(z) \;+\; B\, H^{(2)}_{\ell+\frac{1}{2}}(z)\,. \tag{7.144}$$

The expansion coefficients A, B can be obtained from the asymptotic expansion (7.140). For $|z| \to \infty$ the leading terms yield

$$\frac{1}{\sqrt{z}}\, e^{i(z+\frac{\ell\pi}{2})} \;=\; \frac{1}{\sqrt{z}}\, A\, e^{i(z-\frac{\ell\pi}{2}-\frac{\pi}{2})} \;+\; \frac{1}{\sqrt{z}}\, B\, e^{-i(z-\frac{\ell\pi}{2}-\frac{\pi}{2})} \quad |z| \;\to\; \infty\,. \tag{7.145}$$

This equation can hold only for $B = 0$ and $A = \exp[i(\ell+\frac{1}{2})\pi]$. We conclude

$$H^{(1)}_{-(\ell+\frac{1}{2})}(z) \;=\; i\,(-1)^{\ell}\, H^{(1)}_{\ell+\frac{1}{2}}(z)\,. \tag{7.146}$$

Simarly, one can show

$$H^{(2)}_{-(\ell+\frac{1}{2})}(z) \;=\; -i\,(-1)^{\ell}\, H^{(2)}_{\ell+\frac{1}{2}}(z)\,. \tag{7.147}$$

Spherical Hankel Functions

In analogy to equations (7.90, 7.97) one defines the spherical Hankel functions

$$h^{(1,2)}_{\ell}(z) \;=\; \sqrt{\frac{\pi}{2z}}\; H^{(1,2)}_{\ell+\frac{1}{2}}(z)\,. \tag{7.148}$$

Following the arguments provided above (see page 193) the functions $h^{(1,2)}_{\ell}(z)$ are solutions of the radial Schrödinger equation of free particles (7.53). According to (7.96, 7.132, 7.148) holds for the regular spherical Bessel function

$$j_\ell(z) \;=\; \frac{1}{2}\,[\,h^{(1)}_{\ell}(z) \;+\; h^{(2)}_{\ell}(z)\,]\,. \tag{7.149}$$

We want to establish also the relationship between $h^{(1,2)}_{\ell}(z)$ and the irregular spherical Bessel function $n_\ell(z)$ defined in (7.97). From (7.97, 7.132) follows

$$n_\ell(z) \;=\; \frac{1}{2}\sqrt{\frac{\pi}{2z}}\,(-1)^{\ell+1}\,\left[\,H^{(1)}_{-(\ell+\frac{1}{2})}(z) \;+\; H^{(2)}_{-(\ell+\frac{1}{2})}(z)\right]\,. \tag{7.150}$$

According to (7.146, 7.147) this can be written

$$n_\ell(z) \;=\; \frac{1}{2i}\,\left[\,h^{(1)}_{\ell}(z) \;-\; h^{(2)}_{\ell}(z)\right]\,. \tag{7.151}$$

Equations (7.149, 7.151) are equivalent to

$$h^{(1)}_{\ell}(z) \;=\; j_\ell(z) \;+\; i\,n_\ell(z) \tag{7.152}$$

$$h^{(2)}_{\ell}(z) \;=\; j_\ell(z) \;-\; i\,n_\ell(z)\,. \tag{7.153}$$

Asymptotic Behaviour of Spherical Bessel Functions

We want to derive now the asymptotic behaviour of the spherical Bessel functions $h_\ell^{(1,2)}(z)$, $j_\ell(z)$ and $n_\ell(z)$. From (7.140) and (7.148) one obtains readily

$$h_\ell^{(\frac{1}{2})}(z) \;=\; \frac{(\mp i)^{\ell+1}}{z}\, e^{\pm iz} \sum_{r=0}^{\ell} \frac{(\ell+r)!}{r!(\ell-r)!} \left(\pm\frac{i}{2z}\right)^r \tag{7.154}$$

The leading term in this expansion, at large $|z|$, is

$$h_\ell^{(\frac{1}{2})}(z) \;=\; \frac{(\mp i)^{\ell+1}}{z}\, e^{\pm iz}\,. \tag{7.155}$$

To determine $j_\ell(z)$ and $n_\ell(z)$ we note that for $z \in \mathbb{R}$ the spherical Hankel functions $h_\ell^{(1)}(z)$ and $h_\ell^{(2)}(z)$, as given by (7.140, 7.148), are complex conjugates. Hence, it follows

$$z \in \mathbb{R}: \quad j_\ell(z) \;=\; \mathrm{Re}[h_\ell^{(1)}(z)]\,, \quad n_e ll(z) \;=\; \mathrm{Im}[h_\ell^{(1)}(z)]\,. \tag{7.156}$$

Using

$$p_\ell(z) \;=\; \mathrm{Re}\sum_{r=0}^{\ell} \frac{(\ell+r)!}{r!(\ell-r)!} \left(\frac{i}{2z}\right)^r \;=\; \sum_{r=0}^{[\ell/2]} \frac{(\ell+2r)!}{(2r)!(\ell-2r)!} \left(\frac{-1}{4z^2}\right)^r \tag{7.157}$$

$$\begin{aligned} \frac{i}{2z}\, q_\ell(z) \;&=\; i\,\mathrm{Im}\sum_{r=0}^{\ell} \frac{(\ell+r)!}{r!(\ell-r)!} \left(\frac{i}{2z}\right)^r \\ &=\; \begin{cases} \frac{i}{2z} \sum_{r=0}^{[\ell-1/2]} \frac{(\ell+2r+1)!}{(2r+1)!(\ell-2r-1)!} \left(\frac{-1}{4z^2}\right)^r & \ell \geq 1 \\ 0 & \ell = 0 \end{cases} \end{aligned} \tag{7.158}$$

one can derive then from (7.154) the identities

$$j_\ell(z) \;=\; \frac{\cos[z - (\ell+1)\frac{\pi}{2}]}{z}\, p_\ell(z) \;-\; \frac{\sin[z - (\ell+1)\frac{\pi}{2}]}{2z^2}\, q_\ell(z) \tag{7.159}$$

$$n_\ell(z) \;=\; \frac{\cos[z - (\ell+1)\frac{\pi}{2}]}{2z^2}\, q_\ell(z) \;+\; \frac{\cos(z - \ell\pi/2)}{z}\, p_\ell(z)\,. \tag{7.160}$$

Employing $\cos[z - (\ell+1)\frac{\pi}{2}]\,. = \sin(z - \ell\pi/2)$ and $\sin[z - (\ell+1)\frac{\pi}{2}]\,. = -\cos(z - \ell\pi/2)$ results in the alternative expressions

$$j_\ell(z) \;=\; \frac{\sin(z - \ell\pi/2)}{z}\, p_\ell(z) \;+\; \frac{\cos(z - \ell\pi/2)}{2z^2}\, q_\ell(z) \tag{7.161}$$

$$n_\ell(z) \;=\; \frac{\sin(z - \ell\pi/2)}{2z^2}\, q_\ell(z) \;-\; \frac{\cos(z - \ell\pi/2)}{z}\, p_\ell(z)\,. \tag{7.162}$$

The leading terms in these expansions, at large $|z|$, are

$$j_\ell(z) \;=\; \frac{\sin(z - \ell\pi/2)}{z} \tag{7.163}$$

$$n_\ell(z) \;=\; -\,\frac{\cos(z - \ell\pi/2)}{z}\,. \tag{7.164}$$

Expressions for the Spherical bessel Functions $j_\ell(z)$ and $n_\ell(z)$

The identities (7.161, 7.162) allow one to provide explicit expressions for $j_\ell(z)$ and $n_\ell(z)$. One obtains for $\ell = 0, 1, 2$

$$j_0(z) = \frac{\sin z}{z} \tag{7.165}$$

$$n_0(z) = -\frac{\cos z}{z} \tag{7.166}$$

$$j_1(z) = \frac{\sin z}{z^2} - \frac{\cos z}{z} \tag{7.167}$$

$$n_1(z) = -\frac{\cos z}{z^2} - \frac{\sin z}{z} \tag{7.168}$$

$$j_2(z) = \left(\frac{3}{z^3} - \frac{1}{z}\right)\sin z - \frac{3}{z^2}\cos z \tag{7.169}$$

$$n_2(z) = \left(-\frac{3}{z^3} + \frac{1}{z}\right)\cos z - \frac{3}{z^2}\sin z \tag{7.170}$$

Recursion Formulas of Spherical Bessel Functions

The spherical Bessel functions obey the recursion relationships

$$g_{\ell+1}(z) = \frac{\ell}{z}\, g_\ell(z) - g'_\ell(z) \tag{7.171}$$

$$g_{\ell+1}(z) = \frac{2\ell+1}{z}\, g_\ell(z) - g_{\ell-1}(z) \tag{7.172}$$

where $g_\ell(z)$ is either of the functions $h_\ell^{(1,2)}(z)$, $j_\ell(z)$ and $n_\ell(z)$. One can combine (7.171, 7.172) to obtain the recursion relationship

$$\begin{pmatrix} g_{\ell+1}(z) \\ g'_{\ell+1}(z) \end{pmatrix} = \mathbf{A}_\ell(z) \begin{pmatrix} g_\ell(z) \\ g'_\ell(z) \end{pmatrix} \tag{7.173}$$

$$\mathbf{A}_\ell(z) = \begin{pmatrix} \frac{\ell}{z} & -1 \\ 1 - \frac{\ell(\ell+1)}{z^2} & \frac{\ell+2}{z} \end{pmatrix} \tag{7.174}$$

We want to prove these relationships. For this purpose we need to demonstrate only that the relationships hold for $g_\ell(z) = h_\ell^{(1,2)}(z)$. From the linearity of the relationships (7.171–7.174) and from (7.149, 7.151) follows then that the relatiosnhips hold also for $g_\ell(z) = j_\ell(z)$ and $g_\ell(z) = n_\ell(z)$. To demonstrate that (7.171) holds for $g_\ell(z) = h_\ell^{(1,2)}(z)$ we employ (7.124, 7.131, 7.148) and express

$$g_\ell(z) = h_\ell^{(1,2)}(z) = -2\sqrt{\frac{\pi}{2z}}\, \frac{1}{\sqrt{\pi}\Gamma(\nu+\frac{1}{2})} \left(\frac{z}{2}\right)^\nu \int_{C_{2,4}} dx\, (1-x^2)^{\nu-\frac{1}{2}}\, e^{izx}\,. \tag{7.175}$$

Using $\Gamma(\ell+1) = \ell!$, defining $a = -1$ and employing $f_\nu(z)$ as defined in (7.114) we can write

$$g_\ell(z) = a\, \frac{z^\ell}{2^\ell \ell!}\, f_{\ell+\frac{1}{2}}(z)\,. \tag{7.176}$$

The derivative of this expression is

$$g'_\ell(z) \;=\; \frac{\ell}{z}\, g_\ell(z) \;+\; a\, \frac{z^\ell}{2^\ell \ell!}\, f'_{\ell+\frac{1}{2}}(z)\,. \tag{7.177}$$

Employing (7.121), i.e.,

$$f'_{\ell+\frac{1}{2}}(z) \;=\; -\,\frac{z}{2\ell+2}\, f_{\ell+\frac{3}{2}}(z)\,, \tag{7.178}$$

yields, together with (7.176)

$$g'_\ell(z) \;=\; \frac{\ell}{z}\, g_\ell(z) \;-\; g_{\ell+1}(z) \tag{7.179}$$

from which follows (7.171).
In order to prove (7.172) we differentiate (7.179)

$$g''_\ell(z) \;=\; -\,\frac{\ell}{z^2}\, g_\ell(z) \;+\; \frac{\ell}{z}\, g'_\ell(z) \;-\; g'_{\ell+1}(z)\,. \tag{7.180}$$

Since $g_\ell(z)$ is a solution of the radial Schrödinger equation (7.58) it holds

$$g''_\ell(z) \;=\; -\frac{2}{z}\, g'_\ell(z) \;+\; \frac{\ell(\ell+1)}{z^2}\, g_\ell(z) \;-\; g_\ell(z)\,. \tag{7.181}$$

Using this identity to replace the second derivative in (7.180) yields

$$g'_{\ell+1}(z) \;=\; g_\ell(z) \;-\; \frac{\ell(\ell+2)}{z^2}\, g_\ell(z) \;+\; \frac{\ell+2}{z}\, g'_\ell(z)\,. \tag{7.182}$$

Replacing all first derivatives employing (7.179) leads to (7.172).
To prove (7.173, 7.174) we start from (7.171). The first component of (7.173), in fact, is equivalent to (7.171). The second component of (7.173) is equivalent to (7.182).

Exercise 7.2.5: Provide a detailed derivation of (7.172).
Exercise 7.2.6: Employ the recursion relationship (7.173, 7.174) to determine (a) $j_1(z)$, $j_2(z)$ from $j_0(z)$, $j'_0(z)$ using (7.165), and (b) $n_1(z)$, $n_2(z)$ from $n_0(z)$, $n'_0(z)$ using (7.166).

Chapter 8

Interaction of Charged Particles with Electromagnetic Radiation

In this Section we want to describe how a quantum mechanical particle, e.g., an electron in a hydrogen atom, is affected by electromagnetic fields. For this purpose we need to establish a suitable description of this field, then state the Hamiltonian which describes the resulting interaction.

It turns out that the proper description of the electromagnetic field requires a little bit of effort. We will describe the electromagnetic field classically. Such description should be sufficient for high quantum numbers, i.e., for situations in which the photons absorbed or emitted by the quantum system do not alter the energy content of the field. We will later introduce a simple rule which allows one to account to some limited degree for the quantum nature of the electromagnetic field, i.e., for the existence of discrete photons.

8.1 Description of the Classical Electromagnetic Field / Separation of Longitudinal and Transverse Components

The aim of the following derivation is to provide a description of the electromagnetic field which is most suitable for deriving later a perturbation expansion which yields the effect of electromagnetic radiation on a bound charged particle, e.g., on an electron in a hydrogen atom. The problem is that the latter electron, or other charged particles, are affected by the Coulomb interaction $V(\vec{r})$ which is part of the forces which produce the bound state, and are affected by the external electromagnetic field. However, both the Coulomb interaction due to charges contributing to binding the particle, e.g., the attractive Coulomb force between proton and electron in case of the hydrogen atom, and the external electromagnetic field are of electromagnetic origin and, hence, must be described consistently. This is achieved in the following derivation.

The classical electromagnetic field is governed by the Maxwell equations stated already in (1.27–1.29). We assume that the system considered is in vacuum in which charge and current sources described by the densities $\rho(\vec{r},t)$ and $\vec{J}(\vec{r},t)$ are present. These sources enter the two inhomogeneous

Maxwell equations[1]

$$\nabla \cdot \vec{E}(\vec{r},t) \;=\; 4\,\pi\,\rho(\vec{r},t) \tag{8.1}$$

$$\nabla \times \vec{B}(\vec{r},t) \;-\; \partial_t \vec{E}(\vec{r},t) \;=\; 4\,\pi\,\vec{J}(\vec{r},t)\;. \tag{8.2}$$

In addition, the two homogeneous Maxwell equations hold

$$\nabla \times \vec{E}(\vec{r},t) \;+\; \partial_t \vec{B}(\vec{r},t) \;=\; 0 \tag{8.3}$$

$$\nabla \cdot \vec{B}(\vec{r},t) \;=\; 0\;. \tag{8.4}$$

Lorentz Force A classical particle with charge q moving in the electromagnetic field experiences the so-called Lorentz force $q[\vec{E}(\vec{r},t) \,+\, \vec{v}\times\vec{B}(\vec{r},t)]$ and, accordingly, obeys the equation of motion

$$\frac{d}{dt}\vec{p} \;=\; q\,\Big\{\,\vec{E}[\vec{r}_o(t),t] \;+\; \vec{v}\times\vec{B}[\vec{r}_o(t),t]\,\Big\} \tag{8.5}$$

where $\vec{p}$ is the momentum of the particle and $\vec{r}_o(t)$ it's position at time t. The particle, in turn, contributes to the charge density $\rho(\vec{r},t)$ in (8.1) the term $q\delta(\vec{r}-\vec{r}_o(t))$ and to the current density $\vec{J}(\vec{r},t)$ in (8.2) the term $q\dot{\vec{r}}_o\delta(\vec{r}-\vec{r}_o(t))$. In the non-relativistic limit holds $\vec{p}\approx m\dot{\vec{r}}$ and (8.5) above agrees with the equation of motion as given in (1.25).

Scalar and Vector Potential Setting

$$\vec{B}(\vec{r},t) \;=\; \nabla\times\vec{A}(\vec{r},t) \tag{8.6}$$

for some vector-valued function $\vec{A}(\vec{r},t)$, called the *vector potential*, solves implicitly (8.4). Equation (8.3) reads then

$$\nabla\times\Big(\,\vec{E}(\vec{r},t) \;+\; \partial_t\vec{A}(\vec{r},t)\,\Big) \;=\; 0 \tag{8.7}$$

which is solved by

$$\vec{E}(\vec{r},t) \;+\; \partial_t\vec{A}(\vec{r},t) \;=\; -\nabla V(\vec{r},t) \tag{8.8}$$

where $V(\vec{r},t)$ is a scalar function, called the *scalar potential*. From this follows

$$\vec{E}(\vec{r},t) \;=\; -\nabla V(\vec{r},t) \;-\; \partial_t\vec{A}(\vec{r},t)\;. \tag{8.9}$$

Gauge Transformations We have expressed now the electric and magnetic fields $\vec{E}(\vec{r},t)$ and $\vec{B}(\vec{r},t)$ through the scalar and vector potentials $V(\vec{r},t)$ and $\vec{A}(\vec{r},t)$. As is well known, the relationship between fields and potentials is not unique. The following substitutions, called gauge transformations, alter the potentials, but leave the fields unaltered:

$$\vec{A}(\vec{r},t) \;\longrightarrow\; \vec{A}(\vec{r},t) \;+\; \nabla\chi(\vec{r},t) \tag{8.10}$$

$$V(\vec{r},t) \;\longrightarrow\; V(\vec{r},t) \;-\; \partial_t\chi(\vec{r},t)\;. \tag{8.11}$$

[1] We assume so-called Gaussian units. The reader is referred to the well-known textbook "Classical Electrodynamics", 2nd Edition, by J. D. Jackson (John Wiley & Sons, New York, 1975) for a discussion of these and other conventional units.

This gauge freedom will be exploited now to introduce potentials which are most suitable for the purpose of separating the electromagnetic field into a component arising from the Coulomb potential connected with the charge distribution $\rho(\vec{r},t)$ and the current due to moving net charges, and a component due to the remaining currents. In fact, the gauge freedom allows us to impose on the vector potential $\vec{A}(\vec{r},t)$ the condition

$$\nabla \cdot \vec{A}(\vec{r},t) \;=\; 0\,. \tag{8.12}$$

The corresponding gauge is referred to as the *Coulomb gauge*, a name which is due to the form of the resulting scalar potential $V(\vec{r},t)$. In fact, this potential results from inserting (8.9) into (8.1)

$$\nabla \cdot \Big(-\nabla V(\vec{r},t) \;-\; \partial_t \vec{A}(\vec{r},t) \Big) \;=\; 4\,\pi\,\rho(\vec{r},t)\,. \tag{8.13}$$

Using $\nabla \cdot \partial_t \vec{A}(\vec{r},t) \;=\; \partial_t \nabla \cdot \vec{A}(\vec{r},t)$ together with (8.12) yields then the Poisson equation

$$\nabla^2 V(\vec{r},t) \;=\; -\,4\,\pi\,\rho(\vec{r},t)\,. \tag{8.14}$$

In case of the boundary condition

$$V(\vec{r},t) \;=\; 0 \quad \text{for } \vec{r} \,\in\, \partial\Omega_\infty \tag{8.15}$$

the solution is given by the Coulomb integral

$$V(\vec{r},t) \;=\; \int_{\Omega_\infty} d^3r'\,\frac{\rho(\vec{r}\,',t)}{|\vec{r} \,-\, \vec{r}\,'|} \tag{8.16}$$

This is the potential commonly employed in quantum mechanical calculations for the description of Coulomb interactions between charged particles.
The vector potential $\vec{A}(\vec{r},t)$ can be obtained employing (8.2), the second inhomogeneous Maxwell equation. Using the expressions (8.6) and (8.9) for the fields results in

$$\nabla \times \Big(\nabla \times \vec{A}(\vec{r},t) \Big) \;+\; \partial_t \Big(\nabla V(\vec{r},t) \;+\; \partial_t \vec{A}(\vec{r},t \Big) \;=\; 4\,\pi\,\vec{J}(\vec{r},t)\,. \tag{8.17}$$

The identity

$$\nabla \times \Big(\nabla \times \vec{A}(\vec{r},t) \Big) \;=\; \nabla \Big(\nabla \cdot \vec{A}(\vec{r},t) \Big) \;-\; \nabla^2 \vec{A}(\vec{r},t) \tag{8.18}$$

together with condition (8.12) leads us to

$$\nabla^2 \vec{A}(\vec{r},t) \;-\; \partial_t^2 \vec{A}(\vec{r},t) \;-\; \partial_t \nabla V(\vec{r},t) \;=\; -\,4\,\pi\,\vec{J}(\vec{r},t)\,. \tag{8.19}$$

Unfortunately, equation (8.19) couples the vector potential $\vec{A}(\vec{r},t)$ and $V(\vec{r},t)$. One would prefer a description in which the Coulomb potential (8.16) and the vector potential are uncoupled, such that the latter describes the electromagnetic radiation, and the former the Coulomb interactions in the unperturbed bound particle system. Such description can, in fact, be achieved. For this purpose we examine the offending term $\partial_t \nabla V(\vec{r},t)$ in (8.19) and define

$$\vec{J}_\ell(\vec{r},t) \;=\; \frac{1}{4\pi}\,\partial_t \nabla V(\vec{r},t)\,. \tag{8.20}$$

For the curl of $\vec{J}_\ell$ holds

$$\nabla \times \vec{J}_\ell(\vec{r},t) \;=\; 0\,. \tag{8.21}$$

For the divergence of $\vec{J}_\ell(\vec{r},t)$ holds, using $\partial_t\nabla \;=\; \nabla\partial_t$ and the Poisson equation (8.14),

$$\nabla \cdot \vec{J}_\ell(\vec{r},t) \;=\; \frac{1}{4\pi}\,\partial_t\nabla^2 V(\vec{r},t) \;=\; -\,\partial_t\rho(\vec{r},t) \tag{8.22}$$

or

$$\nabla \cdot \vec{J}_\ell(\vec{r},t) \;+\; \partial_t\rho(\vec{r},t) \;=\; 0\,. \tag{8.23}$$

This continuity equation identifies $\vec{J}_\ell(\vec{r},t)$ as the current due to the time-dependence of the charge distribution $\rho(\vec{r},t)$. Let $\vec{J}(\vec{r},t)$ be the total current of the system under investigation and let $\vec{J}_t \;=\; \vec{J} \;-\; \vec{J}_\ell$. For $\vec{J}$ also holds the continuity equation

$$\nabla \cdot \vec{J}(\vec{r},t) \;+\; \partial_t\rho(\vec{r},t) \;=\; 0 \tag{8.24}$$

and from this follows

$$\nabla \cdot \vec{J}_t(\vec{r},t) \;=\; 0\,. \tag{8.25}$$

Because of properties (8.21) and (8.25) one refers to $\vec{J}_\ell$ and $\vec{J}_t$ as the *longitudinal* and the *transverse* currents, respectively.
The definitions of $\vec{J}_\ell$ and $\vec{J}_t$ applied to (8.19) yield

$$\nabla^2\vec{A}(\vec{r},t) \;-\; \partial_t^2\vec{A}(\vec{r},t) \;=\; -\,4\,\pi\,\vec{J}_t(\vec{r},t)\,. \tag{8.26}$$

This equation does not couple anymore scalar and vector potentials. The vector potential determined through (8.26) and (8.12) and the Coulomb potential (8.16) yield finally the electric and magnetic fields. $V(\vec{r},t)$ contributes solely an electric field component

$$\vec{E}_\ell(\vec{r},t) \;=\; -\,\nabla V(\vec{r},t) \tag{8.27}$$

which is obviously curl-free ($\nabla \times \vec{E}_\ell(\vec{r},t) \;=\; 0$), hence, the name *longitudinal electric field*. $\vec{A}(\vec{r},t)$ contributes an electrical field component as well as the total magnetic field

$$\vec{E}_t(\vec{r},t) \;=\; -\,\partial_t\vec{A}(\vec{r},t) \tag{8.28}$$

$$\vec{B}_t(\vec{r},t) \;=\; \nabla \times \vec{A}(\vec{r},t)\,. \tag{8.29}$$

These fields are obviously divergence -free (e.g., $\nabla \cdot \vec{E}_t(\vec{r},t) \;=\; 0$), hence, the name *transverse fields*.

8.2 Planar Electromagnetic Waves

The current density $\vec{J}_t$ describes ring-type currents in the space under consideration; such current densities exist, for example, in a ring-shaped antenna which exhibits no net charge, yet a current. Presently, we want to assume that no ring-type currents, i.e., no divergence-free currents, exist in the space considered. In this case (8.26) turns into the well-known wave equation

$$\nabla^2\vec{A}(\vec{r},t) \;-\; \partial_t^2\vec{A}(\vec{r},t) \;=\; 0 \tag{8.30}$$

which describes electromagnetic fields in vacuum. A complete set of solutions is given by the so-called plane waves

$$\vec{A}(\vec{r},t) \;=\; A_o\,\hat{u}\exp\Big[\,i(\vec{k}\cdot\vec{r}\,\mp\,\omega t)\Big] \tag{8.31}$$

where the dispersion relationship

$$|\vec{k}| \;=\; \omega \tag{8.32}$$

holds. Note that in the units chosen the velocity of light is $c = 1$. Here the "-" sign corresponds to so-called *incoming waves* and the "+" sign to *outgoing waves*[2], the constant $\vec{k}$ is referred to as the *wave vector*. The Coulomb gauge condition (8.12) yields

$$\hat{u}\cdot\vec{k} \;=\; 0\,. \tag{8.33}$$

$\hat{u}$ is a unit vector ($|\hat{u}| = 1$) which, obviously, is orthogonal to $\vec{k}$; accordingly, there exist two linearly independent orientations for $\hat{u}$ corresponding to two independent planes of polarization.
We want to characterize now the radiation field connected with the plane wave solutions (8.31). The corresponding electric and magnetic fields, according to (8.28, 8.29), are

$$\vec{E}_t(\vec{r},t) \;=\; \pm i\,\omega\,\vec{A}(\vec{r},t) \tag{8.34}$$

$$\vec{B}_t(\vec{r},t) \;=\; i\,\vec{k}\times\vec{A}(\vec{r},t)\,. \tag{8.35}$$

The vector potential in (8.31) and the resulting fields (8.34, 8.35) are complex-valued quantities. In applying the potential and fields to physical observables and processes we will only employ the real parts.
Obviously, $\vec{E}_t(\vec{r},t)$ and $\vec{B}_t(\vec{r},t)$ in (8.34, 8.35), at each point $\vec{r}$ and moment t, are orthogonal to each other and are both orthogonal to the wave vector $\vec{k}$. The latter vector describes the direction of propagation of the energy flux connected with the plane wave electromagnetic radiation. This flux is given by

$$\vec{S}(\vec{r},t) \;=\; \frac{1}{4\pi}\,\mathrm{Re}\,\vec{E}_t(\vec{r},t)\times\mathrm{Re}\,\vec{B}(\vec{r},t)\,. \tag{8.36}$$

Using the identity $\vec{a}\times(\vec{b}\times\vec{c}) = \vec{b}\,(\vec{a}\cdot\vec{c}) - \vec{c}\,(\vec{a}\cdot\vec{b})$ and (8.31, 8.32, 8.34, 8.35) one obtains

$$\vec{S}(\vec{r},t) \;=\; \pm\frac{\omega^2}{4\pi}\,|A_o|^2\,\hat{k}\,\sin^2(\,\vec{k}\cdot\vec{r}-\omega t\,) \tag{8.37}$$

where $\hat{k}$ is the unit vector $\hat{k} = \vec{k}/|\vec{k}|$. Time average over one period $2\pi/\omega$ yields

$$\langle\,\vec{S}(\vec{r},t)\,\rangle \;=\; \pm\frac{\omega^2}{8\pi}\,|A_o|^2\,\hat{k}\,. \tag{8.38}$$

In this expression for the energy flux one can interprete $\hat{k}$ as the propagation velocity (note $c = 1$) and, hence,

$$\langle\epsilon\rangle \;=\; \frac{\omega^2}{8\pi}\,|A_o|^2 \tag{8.39}$$

[2]The definition *incoming waves* and *outgoing waves* is rationalized below in the discussion following Eq. (8.158); see also the comment below Eqs. (8.38, 8.39).

as the energy density. The sign in (8.38) implies that for *incoming waves*, defined below Eqs. (8.31,8.32), the energy of the plane wave is transported in the direction of $-\vec{k}$, whereas in the case of *outgoing* waves the energy is transported in the direction of $\vec{k}$.
A correct description of the electromagnetic field requires that the field be quantized. A 'poor man's' quantization of the field is possible at this point by expressing the energy density (8.39) through the density of photons connected with the planar waves (8.31). These photons each carry the energy $\hbar\omega$. If we consider a volume $\mathcal{V}$ with a number of photons $\mathcal{N}_\omega$ the energy density is obviously

$$\langle \epsilon \rangle \;=\; \frac{\mathcal{N}_\omega \hbar\omega}{\mathcal{V}} \; . \tag{8.40}$$

It should be pointed out that $\mathcal{N}_\omega$ represents the number of photons for a specific frequency ω, a specific $\hat{k}$ and a specific $\hat{u}$. Comparision of (8.39) and (8.40) allows one to express then the field amplitudes

$$A_o \;=\; \sqrt{\frac{8\pi\mathcal{N}_\omega\hbar}{\omega\mathcal{V}}} \; . \tag{8.41}$$

Inserting this into (8.31) allows one finally to state for the planar wave vector potential

$$\vec{A}(\vec{r},t) \;=\; \sqrt{\frac{8\pi\mathcal{N}_\omega\hbar}{\omega\mathcal{V}}}\;\hat{u}\exp\Big[\,i(\vec{k}\cdot\vec{r}-\omega t)\Big] \quad , \; |\vec{k}| \;=\; \omega \; , \; \hat{u}\cdot\vec{k} \;=\; 0 \; . \tag{8.42}$$

8.3 Hamilton Operator

The classical Hamiltonian for a particle of charge q in a scalar and vector potential $V(\vec{r})$ and $\vec{A}(\vec{r},t)$, respectively, is

$$\begin{aligned} H \;=\; & \frac{\Big[\,\vec{p}\,-\,q\vec{A}(\vec{r},t)\,\Big]^2}{2\,m} \;+\; qV(\vec{r}) \\ & + \frac{1}{8\pi}\int_{\Omega_\infty} d^3r'\,E_\ell^2 \;+\; \frac{1}{16\pi}\int_{\Omega_\infty} d^3r\,\left(\,|E_t|^2\,+\,|B_t|^2\,\right) \; . \end{aligned} \tag{8.43}$$

Here the fields are defined through Eqs. (8.27, 8.28, 8.29) together with the potentials (8.16, 8.31). The integrals express the integration over the energy density of the fields. Note that $\vec{E}_\ell(\vec{r},t)$ is real and that $\vec{E}_t(\vec{r},t)$, $\vec{B}_t(\vec{r},t)$ are complex leading to the difference of a factor $\frac{1}{2}$ in the energy densities of the lontitudinal and transverse components of the fields.
We assume that the energy content of the fields is not altered significantly in the processes described and, hence, we will neglect the respective terms in the Hamiltonian (8.43). We are left with a classical Hamiltonian function which has an obvious quantum mechanical analogue

$$\hat{H} \;=\; \frac{\Big[\,\hat{\vec{p}}\,-\,q\vec{A}(\vec{r},t)\,\Big]^2}{2\,m} \;+\; qV(\vec{r}) \; . \tag{8.44}$$

replacing the classical momentum $\vec{p}$ by the differential operator $\hat{\vec{p}} = \frac{\hbar}{i}\nabla$. The wave function $\Psi(\vec{r},t)$ of the particle is then described by the Schrödinger equation

$$i\,\hbar\,\partial_t\,\Psi(\vec{r},t) \;=\; \hat{H}\,\Psi(\vec{r},t) \; . \tag{8.45}$$

Gauge Transformations It is interesting to note that in the quantum mechanical description of a charged particle the potentials $V(\vec{r},t)$ and $\vec{A}(\vec{r},t)$ enter whereas in the classical equations of motion

$$m\ddot{\vec{r}} \;=\; q\,\vec{E}(\vec{r},t) \;+\; q\,\dot{\vec{r}} \times \vec{B}(\vec{r},t) \tag{8.46}$$

the fields enter. This leads to the question in how far the gauge transformations (8.10, 8.11) affect the quantum mechanical description. In the classical case such question is mute since the gauge transformations do not alter the fields and, hence, have no effect on the motion of the particle described by (8.46).
Applying the gauge transformations (8.10, 8.11) to (8.44, 8.45) leads to the Schrödinger equation

$$i\hbar\partial_t\Psi(\vec{r},t) \;=\; \left[\frac{\left[\hat{\vec{p}} \,-\, q\vec{A} \,-\, q((\nabla\chi))\right]^2}{2\,m} \;+\; qV \;-\; q((\partial_t\chi))\right]\Psi(\vec{r},t) \tag{8.47}$$

where $((\cdots))$ denotes derivatives in $((\nabla\chi))$ and $((\partial_t\chi))$ which are confined to the function $\chi(\vec{r},t)$ inside the double brackets. One can show that (8.47) is equivalent to

$$i\hbar\partial_t e^{iq\chi(\vec{r},t)/\hbar}\Psi(\vec{r},t) \;=\; \left[\frac{\left[\hat{\vec{p}} \,-\, q\vec{A}\right]^2}{2\,m} \;+\; qV\right] e^{iq\chi(\vec{r},t)/\hbar}\Psi(\vec{r},t)\;. \tag{8.48}$$

For this purpose one notes

$$i\hbar\partial_t\, e^{iq\chi(\vec{r},t)/\hbar}\Psi(\vec{r},t) \;=\; e^{iq\chi(\vec{r},t)/\hbar}\,\left[\,i\hbar\partial_t \;-\; q((\partial_t\chi))\,\right]\Psi(\vec{r},t) \tag{8.49}$$

$$\hat{\vec{p}}\,e^{iq\chi(\vec{r},t)/\hbar}\Psi(\vec{r},t) \;=\; e^{iq\chi(\vec{r},t)/\hbar}\left[\hat{\vec{p}} \;+\; q((\nabla\chi))\right]\Psi(\vec{r},t)\;. \tag{8.50}$$

The equivalence of (8.47, 8.48) implies that the gauge transformation (8.10, 8.11) of the potentials is equivalent to multiplying the wave function $\Psi(\vec{r},t)$ by a local and time-dependent phase factor $e^{iq\chi(\vec{r},t)/\hbar}$. Obviously, such phase factor does not change the probability density $|\Psi(\vec{r},t)|^2$ and, hence, does not change expectation values which contain the probability densities[3].
An important conceptual step of modern physics has been to turn the derivation given around and to state that introduction of a local phase factor $e^{iq\chi(\vec{r},t)/\hbar}$ should not affect a system and that, accordingly, in the Schrödinger equation

$$i\hbar\partial_t\Psi(\vec{r},t) \;=\; \left[\frac{\left[\hat{\vec{p}} \,-\, q\vec{A}\right]^2}{2\,m} \;+\; qV\right]\Psi(\vec{r},t)\;. \tag{8.51}$$

the potentials $\vec{A}(\vec{r},t)$ and $V(\vec{r},t)$ are necessary to compensate terms which arise through the phase factor. It should be noted, however, that this principle applies only to fundamental interactions, not to phenomenological interactions like the molecular van der Waals interaction.
The idea just stated can be generalized by noting that multiplication by a phase factor $e^{iq\chi(\vec{r},t)/\hbar}$ constitutes a unitary transformation of a scalar quantity, i.e., an element of the group U(1). Elementary constituents of matter which are governed by other symmetry groups, e.g., by the group

[3]The effect on other expectation values is not discussed here.

SU(2), likewise can demand the existence of fields which compensate local transformations described by $e^{i\vec{\sigma}\cdot\vec{\chi}(\vec{r},t)}$ where $\vec{\sigma}$ is the vector of Pauli matrices, the generators of SU(2). The resulting fields are called Yang-Mills fields.
The Hamiltonian (8.44) can be expanded

$$H \;=\; \frac{\hat{\vec{p}}^{\,2}}{2m} \;-\; \frac{q}{2m}\left(\hat{\vec{p}}\cdot\vec{A} \;+\; \vec{A}\cdot\hat{\vec{p}}\right) \;+\; \frac{q^2}{2m}A^2 \;+\; qV \tag{8.52}$$

For any function $f(\vec{r})$ holds

$$\left(\hat{\vec{p}}\cdot\vec{A} \;-\; \vec{A}\cdot\hat{\vec{p}}\right) f(\vec{r}) \;=\; \frac{\hbar}{i}\left(\vec{A}\cdot\nabla f \;+\; f\,\nabla\cdot\vec{A} \;-\; \vec{A}\cdot\nabla f\right) \;=\; \frac{\hbar}{i}\,f\,\nabla\cdot\vec{A}\,. \tag{8.53}$$

This expression vanishes in the present case since since $\nabla\cdot A \;=\; 0$ [cf. (8.12)]. Accordingly, holds

$$\hat{\vec{p}}\cdot A\,f \;=\; \vec{A}\cdot\hat{\vec{p}}\,f \tag{8.54}$$

and, consequently,

$$H \;=\; \frac{\hat{\vec{p}}^{\,2}}{2m} \;-\; \frac{q}{m}\,\hat{\vec{p}}\cdot\vec{A} \;+\; \frac{q^2}{2m}A^2 \;+\; qV\,. \tag{8.55}$$

8.4 Electron in a Stationary Homogeneous Magnetic Field

We consider now the motion of an electron with charge $q \;=\; -e$ and mass $m \;=\; m_e$ in a homogeneous magnetic field as described by the Schrödinger equation (8.45) with Hamiltonian (8.55). In this case holds $V(\vec{r},t) \;\equiv\; 0$. The stationary homogeneous magnetic field

$$\vec{B}(\vec{r},t) \;=\; \vec{B}_o\,, \tag{8.56}$$

due to the gauge freedom, can be described by various vector potentials. The choice of a vector potential affects the form of the wave functions describing the eigenstates and, thereby, affects the complexity of the mathematical derivation of the wave functions.

Solution for Landau Gauge A particularly convenient form for the Hamiltonian results for a choice of a so-called Landau gauge for the vector potential $\vec{A}(\vec{r},t)$. In case of a homogeneous potential pointing in the x_3-direction, e.g., for $\vec{B}_o \;=\; B_o\,\hat{e}_3$ in (8.56), the so-called Landau gauge associates the vector potential

$$\vec{A}_L(\vec{r}) \;=\; B_o\,x_1\,\hat{e}_2 \tag{8.57}$$

with a homogeneous magnetic field $\vec{B}_o$. The vector potential (8.57) satisfies $\nabla\cdot\vec{A} \;=\; 0$ and, therefore, one can employ the Hamiltonian (8.55). Using Cartesian coordinates this yields

$$H \;=\; -\frac{\hbar^2}{2m_e}\left(\partial_1^2 \;+\; \partial_2^2 \;+\; \partial_3^2\right) \;+\; \frac{eB_o\hbar}{i\,m_e}\,x_1\,\partial_2 \;+\; \frac{e^2B_o^2}{2m_e}\,x_1^2 \tag{8.58}$$

where $\partial_j \;=\; (\partial/\partial x_j)$, $j = 1,2,3$.
We want to describe the stationary states corresponding to the Hamiltonian (8.58). For this purpose we use the wave function in the form

$$\Psi(E,k_2,k_3;x_1,x_2,x_3) \;=\; \exp(ik_2x_2 \;+\; ik_3x_3)\,\phi_E(x_1)\,. \tag{8.59}$$

This results in a stationary Schrödinger equation

$$\left(-\frac{\hbar^2}{2m_e}\,\partial_1^2 \;+\; \frac{\hbar^2 k_2^2}{2m_e} \;+\; \frac{\hbar^2 k_3^2}{2m_e} \;+\; \frac{eB_o\hbar k_2}{m_e}\,x_1 \;+\; \frac{e^2B_o^2}{2m_e}\,x_1^2 \right) \phi_E(x_1) \\ =\; E\,\phi_E(x_1)\;. \tag{8.60}$$

Completing the square

$$\frac{e^2B_o^2}{2m_e}\,x_1^2 \;+\; \frac{eB_o\hbar k_2}{m_e}\,x_1 \;=\; \frac{e^2B_o^2}{2m_e}\left(x \;+\; \frac{\hbar k_2}{eB_o} \right)^2 \;-\; \frac{\hbar^2 k_2^2}{2m_e} \tag{8.61}$$

leads to

$$\left[-\frac{\hbar^2}{2m_e}\,\partial_1^2 \;+\; \frac{1}{2}\,m_e\omega^2\,(\,x_1 \;+\; x_{1o}\,)^2 + \;\frac{\hbar^2 k_3^2}{2m_e} \right] \phi_E(x_1) \;=\; E\,\phi_E(x_1)\;. \tag{8.62}$$

where

$$x_{1o} \;=\; \frac{\hbar k_2}{eB_o} \tag{8.63}$$

and where

$$\omega \;=\; \frac{eB_o}{m_e} \tag{8.64}$$

is the classical Larmor frequency ($c = 1$). It is important to note that the completion of the square absorbs the kinetic energy term of the motion in the x_2-direction described by the factor $\exp(ik_2x_2)$ of wave function (8.59).
The stationary Schrödinger equation (8.62) is that of a displaced (by x_{1o}) harmonic oscillator with shifted (by $\hbar^2k_3^2/2m_e$) energies. From this observation one can immediately conclude that the wave function of the system, according to (8.59), is

$$\Psi(n, k_2, k_3; x_1, x_2, x_3) \;=\; \exp(ik_2x_2 \;+\; ik_3x_3) \;\times \\ \tfrac{1}{\sqrt{2^n\,n!}}\,\left[\tfrac{m_e\omega}{\pi\hbar}\right]^{\frac{1}{4}} \exp\left[-\tfrac{m_e\omega(x_1+x_{1o})^2}{2\hbar} \right]\, H_n\left(\sqrt{\tfrac{m\omega}{\hbar}}(x_1 + x_{1o})\right) \tag{8.65}$$

where we replaced the parameter E by the integer n, the familiar harmonic oscillator quantum number. The energies corresponding to these states are

$$E(n, k_2, k_3) \;=\; \hbar\omega(n \;+\; \frac{1}{2}) \;+\; \frac{\hbar^2 k_3^2}{2m_e}\;. \tag{8.66}$$

Obviously, the states are degenerate in the quantum number k_2 describing displacement along the x_2 coordinate. Without affecting the energy one can form wave packets in terms of the solutions (8.65) which localize the electrons. However, according to (8.63) this induces a spread of the wave function in the x_1 direction.

Solution for Symmetric Gauge The solution obtained above has the advantage that the derivation is comparatively simple. Unfortunately, the wave function (8.65), like the corresponding gauge (8.57), is not symmetric in the x_1- and x_2-coordinates. We want to employ , therefore, the so-called symmetric gauge which expresses the homogeneous potential (8.56) through the vector potential

$$\vec{A}(\vec{r}) \;=\; \frac{1}{2}\,\vec{B}_o \times \vec{r}\;. \tag{8.67}$$

One can readily verify that this vector potential satisfies the condition (8.12) for the Coulomb gauge.
For the vector potential (8.67) one can write

$$\hat{\vec{p}} \cdot \vec{A} \;=\; \frac{\hbar}{2i} \nabla \cdot \vec{B}_o \times \vec{r} \,. \tag{8.68}$$

Using $\nabla \cdot (\vec{u} \times \vec{v}) \;=\; -\,\vec{u} \cdot \nabla \times \vec{v} \;+\; \vec{v} \cdot \nabla \times \vec{u}$ yields, in the present case of constant $\vec{B}_o$, for any function $f(\vec{r})$

$$\hat{\vec{p}} \cdot \vec{A}\, f \;=\; -\,\vec{B}_o \cdot \hat{\vec{p}} \times \vec{r}\, f \,. \tag{8.69}$$

The latter can be rewritten, using $\nabla \times (\vec{u} f) \;=\; -\vec{u} \times \nabla f \;+\; f \nabla \times \vec{u}$ and $\nabla \times \vec{r} \;=\; 0$,

$$\vec{B}_o \cdot \hat{\vec{p}} \times \vec{r}\, f \;=\; \vec{B}_o \cdot \left(\vec{r} \times \hat{\vec{p}} \right) \, f \,. \tag{8.70}$$

Identifying $\vec{r} \times \hat{\vec{p}}$ with the angular momentum operator $\vec{L}$, the Hamiltonian (8.52) becomes

$$H \;=\; \frac{\hat{\vec{p}}^{\,2}}{2m_e} \;+\; \frac{e}{2m_e}\, \vec{B}_o \cdot \vec{L} \;+\; \frac{e^2}{8m_e} \, \left(\vec{B}_o \times \vec{r} \right)^2 \,. \tag{8.71}$$

Of particular interest is the contribution

$$V_{\rm mag} \;=\; \frac{e}{2m_e} \vec{L} \cdot \vec{B}_o \tag{8.72}$$

to Hamiltonian (8.71). The theory of classical electromagnetism predicts an analogue energy contribution , namely,

$$V_{\rm mag} \;=\; -\vec{\mu}_{\rm class} \cdot \vec{B}_o \tag{8.73}$$

where $\vec{\mu}_{\rm class}$ is the magnetic moment connected with a current density $\vec{j}$

$$\vec{\mu}_{\rm class} \;=\; \frac{1}{2} \int \vec{r} \times \vec{j}(\vec{r}) \; d\vec{r} \tag{8.74}$$

We consider a simple case to relate (8.72) and (8.73, 8.74), namely, an electron moving in the x, y-plane with constant velocity v on a ring of radius r. In this case the current density measures $-e\, v$ oriented tangentially to the ring. Accordingly, the magnetic moment (8.74) is in the present case

$$\vec{\mu}_{\rm class} \;=\; -\,\frac{1}{2}\, e\, r\, v\; \hat{e}_3 \,. \tag{8.75}$$

The latter can be related to the angular momentum $\vec{\ell}_{\rm class} \;=\; r\, m_e v\, \hat{e}_3$ of the electron

$$\vec{\mu}_{\rm class} \;=\; -\,\frac{e}{2m_e}\, \vec{\ell}_{\rm class} \tag{8.76}$$

and, accordingly,

$$V_{\rm mag} \;=\; \frac{e}{2m_e}\, \vec{\ell}_{\rm class} \cdot \vec{B}_o \,. \tag{8.77}$$

Comparision with (8.72) allows one to interpret

$$\vec{\mu} \;=\; -\,\frac{e}{2m_e}\, \vec{L} \tag{8.78}$$

as the quantum mechanical magnetic moment operator for the electron (charge $-e$).
We will demonstrate in Sect. 10 that the spin of the electron, described by the operator $\vec{S}$, likewise, gives rise to an energy contribution (8.72) with an associated magnetic moment $-\,g\,\frac{e}{2m_e}\,\vec{S}$ where $g \approx 2$. A derivation of his property and the value of g, the so-called gyromagnetic ratio of the electron, requires a Lorentz-invariant quantum mechanical description as provided in Sect. 10.
For a magnetic field (8.56) pointing in the x_3-direction the symmetric gauge (8.67) yields a more symmetric solution which decays to zero along both the $\pm x_1$- and the $\pm x_2$-direction. In this case, i.e., for $\vec{B}_o = B_o\,\hat{e}_3$, the Hamiltonian (8.71) is

$$\hat{H} \;=\; \frac{\hat{\vec{p}}^{\,2}}{2m_e} \;+\; \frac{e^2B_o^2}{8m_e}\,\left(x_1^2 \,+\, x_2^2\right) + \;\frac{eB_o}{2m_e}\,L_3 \;\;. \tag{8.79}$$

To obtain the stationary states, i.e, the solutions of

$$\hat{H}\,\Psi_E(x_1,x_2,x_3) \;=\; E\,\Psi_E(x_1,x_2,x_3)\;, \tag{8.80}$$

we separate the variable x_1, x_2 from x_3 setting

$$\Psi_E(x_1,x_2,x_3) \;=\; \exp(ik_3x_3)\,\psi(x_1,x_2)\;. \tag{8.81}$$

The functions $\psi(x_1,x_2)$ obey then

$$\hat{H}_o\,\psi(x_1,x_2) \;=\; E'\psi(x_1,x_2) \tag{8.82}$$

where

$$\hat{H}_o \;=\; -\frac{\hbar^2}{2m_e}\left(\partial_1^2 \,+\, \partial_2^2\right) \;+\; \frac{1}{2}m_e\omega^2\left(x_1^2 \,+\, x_2^2\right) \;+\; \hbar\omega\frac{1}{i}\left(x_1\partial_2 \,-\, x_2\partial_1\right) \tag{8.83}$$

$$E' \;=\; E \;-\; \frac{\hbar^2k_3^2}{2m_e}\;. \tag{8.84}$$

We have used here the expression for the angular momentum operator

$$\hat{L}_3 \;=\; (\hbar/i)(x_1\partial_2 - x_2\partial_1)\;. \tag{8.85}$$

The Hamiltonian (8.83) describes two identical oscillators along the x_1-and x_2-directions which are coupled through the angular momentum operator $\hat{L}_3$. Accordingly, we seek stationary states which are simultaneous eigenstates of the Hamiltonian of the two-dimensional isotropic harmonic oscillator

$$\hat{H}_{osc} \;=\; -\frac{\hbar^2}{2m_e}\left(\partial_1^2 \,+\, \partial_2^2\right) \;+\; \frac{1}{2}m_e\omega^2\left(x_1^2 \,+\, x_2^2\right) \tag{8.86}$$

as well as of the angular momentum operator $\hat{L}_3$. To obtain these eigenstates we introduce the customary dimensionless variables of the harmonic oscillator

$$X_j \;=\; \sqrt{\frac{m_e\omega}{\hbar}}\,x_j \quad,\; j \,=\, 1,2\;. \tag{8.87}$$

(8.83) can then be expressed

$$\frac{1}{\hbar\omega}\,\hat{H}_o \;=\; -\frac{1}{2}\left(\frac{\partial^2}{\partial X_1^2} \,+\, \frac{\partial^2}{\partial X_2^2}\right) \;+\; \frac{1}{2}\left(X_1^2 \,+\, X_2^2\right) \;+\; \frac{1}{i}\left(X_1\frac{\partial}{\partial X_2} \,-\, X_2\frac{\partial}{\partial X_1}\right)\;. \tag{8.88}$$

Employing the creation and annihilation operators

$$a_j^\dagger \;=\; \frac{1}{\sqrt{2}}\left(X_j \,-\, \frac{\partial}{\partial X_j}\right) \quad ;\; a_j \;=\; \frac{1}{\sqrt{2}}\left(X_j \,+\, \frac{\partial}{\partial X_j}\right) \qquad ;\; j \;=\; 1,2 \tag{8.89}$$

and the identity

$$\omega\,\hat{L}_3 \;=\; \frac{1}{i}\left(a_1^\dagger a_2 \;-\; a_2^\dagger a_1\right)\,, \tag{8.90}$$

which can readily be proven, one obtains

$$\frac{1}{\hbar\omega}\,\hat{H} \;=\; a_1^\dagger a_1 \;+\; a_2^\dagger a_2 \;+\; \mathbb{1} \;+\; \frac{1}{i}\left(a_1^\dagger a_2 \,-\, a_2^\dagger a_1\right)\,. \tag{8.91}$$

We note that the operator $a_1^\dagger a_2 \,-\, a_2^\dagger a_1$ leaves the total number of vibrational quanta invariant, since one phonon is annihilated and one created. We, therefore, attempt to express eigenstates in terms of vibrational wave functions

$$\Psi(j,m;x_1,x_2) \;=\; \frac{\left(a_1^\dagger\right)^{j+m}}{\sqrt{(j+m)!}}\,\frac{\left(a_1^\dagger\right)^{j-m}}{\sqrt{(j-m)!}}\,\Psi(0,0;x_1,x_2) \tag{8.92}$$

where $\Psi(0,0;x_1,x_2)$ is the wave function for the state with zero vibrational quanta for the x_1- as well as for the x_2-oscillator. (8.92) represents a state with $j+m$ quanta in the x_1-oscillator and $j-m$ quanta in the x_2-oscillator, the total vibrational energy being $\hbar\omega(2j+1)$. In order to cover all posible vibrational quantum numbers one needs to choose j, m as follows:

$$j \;=\; 0,\,\frac{1}{2},\,1,\,\frac{3}{2},\ldots \quad ,\; m \;=\; -j,\,-j+1,\,\ldots,+j\,. \tag{8.93}$$

The states (8.92) are not eigenstates of $\hat{L}_3$. Such eigenstates can be expressed, however, through a combination of states

$$\Psi'(j,m';x_1,x_2) \;=\; \sum_{m=-j}^{j} \alpha_{mm'}^{(j)}\,\Psi(j,m;x_1,x_2)\,. \tag{8.94}$$

Since this state is a linear combination of states which all have vibrational energy $(2j+1)\hbar\omega$, (8.94) is an eigenstate of the vibrational Hamiltonian, i.e., it holds

$$\left(a_1^\dagger a_1 \;+\; a_2^\dagger a_2 \;+\; \mathbb{1}\right)\,\Psi'(j,m';x_1,x_2) \;=\; (\,2j \;+\; 1\,)\,\Psi'(j,m';x_1,x_2)\,. \tag{8.95}$$

We want to choose the coefficients $\alpha_{mm'}^{(j)}$ such that (8.94) is also an eigenstate of $\hat{L}_3$, i.e., such that

$$\frac{1}{i}\left(a_1^\dagger a_2 \,-\, a_2^\dagger a_1\right)\,\Psi'(j,m';x_1,x_2) \;=\; 2m'\,\Psi'(j,m';x_1,x_2) \tag{8.96}$$

holds. If this property is, in fact, obeyed, (8.94) is an eigenstate of $\hat{H}_o$

$$\hat{H}_o\,\Psi'(j,m';x_1,x_2) \;=\; \hbar\omega\,\left(\,2j \;+\; 2m' \;+\; 1\,\right)\,\Psi'(j,m';x_1,x_2)\,. \tag{8.97}$$

In order to obtain coefficients $\alpha^{(j)}_{mm'}$ we can profitably employ the construction of angular momentum states in terms of spin–$\frac{1}{2}$ states as presented in Sects. 5.9,5.10,5.11. If we identify

$$\underbrace{a_1^\dagger,\, a_1,\, a_2^\dagger, a_2}_{\text{present notation}} \quad \longleftrightarrow \quad \underbrace{b_+^\dagger,\, b_+,\, b_-^\dagger,\, b_-}_{\text{notation in Sects. 5.9,5.10,5.11}} \tag{8.98}$$

then the states $\Psi(j,m;x_1,x_2)$ defined in (8.92) correspond to the eigenstates $|\Psi(j,m)\rangle$ in Sect. 5.9. According to the derivation given there, the states are eigenstates of the operator [we use for the operator the notation of Sect. 5.10, cf. Eq.(5.288)]

$$\hat{J}_3 \;=\; \frac{1}{2}\left(\, a_1^\dagger a_1 \;-\; a_2^\dagger a_2 \,\right) \tag{8.99}$$

with eigenvalue m. The connection with the present problem arises due to the fact that the operator J_2 in Sect. 5.10, which corresponds there to the angular momentum in the x_2–direction, is in the notation of the present section

$$\hat{J}_2 \;=\; \frac{1}{2i}\left(\, a_1^\dagger a_2 \;-\; a_2^\dagger a_1 \,\right)\;, \tag{8.100}$$

i.e., except for a factor $\frac{1}{2}$, is identical to the operator $\hat{L}_3$ introduced in (8.84) above. This implies that we can obtain eigenstates of $\hat{L}_3$ by rotation of the states $\Psi(j,m;x_1,x_2)$. The required rotation must transform the x_3–axis into the x_2–axis. According to Sect. 5.11 such transformation is provided through

$$\Psi'(j,m';x_1,x_2) \;=\; D^{(j)}_{mm'}(\frac{\pi}{2},\frac{\pi}{2},0)\,\Psi(j,m;x_1,x_2) \tag{8.101}$$

where $D^{(j)}_{mm'}(\frac{\pi}{2},\frac{\pi}{2},0)$ is a rotation matrix which describes the rotation around the x_3–axis by $\frac{\pi}{2}$ and then around the new x_2–axis by $\frac{\pi}{2}$, i.e., a transformation moving the x_3–axis into the x_2–axis. The first rotation contributes a factor $\exp(-im\frac{\pi}{2})$, the second rotation a factor $d^{(j)}_{mm'}(\frac{\pi}{2})$, the latter representing the Wigner rotation matrix of Sect. 5.11. Using the explicit form of the Wigner rotation matrix as given in (5.309) yields finally

$$\begin{aligned}\Psi'(j,m';x_1,x_2) \;=\; &\left(\tfrac{1}{2}\right)^{2j}\,\textstyle\sum_{m=-j}^{j}\sum_{t=0}^{j-m'}\sqrt{\frac{(j+m)!(j-m)!}{(j+m')!(j-m')!}}\\ &\begin{pmatrix} j+m' \\ m+m'-t \end{pmatrix}\begin{pmatrix} j-m' \\ t \end{pmatrix}(-1)^{j-m'-t}\,(-i)^m\,\Psi(j,m;x_1,x_2)\;.\end{aligned} \tag{8.102}$$

We have identified, thus, the eigenstates of (8.83) and confirmed the eigenvalues stated in (8.97).

8.5 Time-Dependent Perturbation Theory

We want to consider now a quantum system involving a charged particle in a bound state perturbed by an external radiation field described through the Hamiltonian (8.55). We assume that the scalar potential V in (8.55) confines the particle to stationary bound states; an example is the Coulomb potential $V(\vec{r},t) = 1/4\pi r$ confining an electron with energy $E < 0$ to move in the well known orbitals of the hydrogen atom. The external radiation field is accounted for by the vector potential $\vec{A}(\vec{r},t)$ introduced above. In the simplest case the radiation field consists of a single planar electromagnetic wave described through the potential (8.31). Other radiation fields can

be expanded through Fourier analysis in terms of such plane waves. We will see below that the perturbation resulting from a 'pure' plane wave radiation field will serve us to describe also the perturbation resulting from a radiation field made up of a superposition of many planar waves.
The Hamiltonian of the particle in the radiation field is then described through the Hamiltonian

$$H = H_o + \mathcal{V}_S \tag{8.103}$$

$$H_o = \frac{\hat{\vec{p}}^2}{2m} + qV \tag{8.104}$$

$$\mathcal{V}_S = -\frac{q}{m}\hat{\vec{p}}\cdot\vec{A}(\vec{r},t) + \frac{q^2}{2m}A^2(\vec{r},t) \tag{8.105}$$

where $\vec{A}(\vec{r},t)$ is given by (8.42). Here the so-called unperturbed system is governed by the Hamiltonian H_o with stationary states defined through the eigenvalue problem

$$H_o|n\rangle = \epsilon_n|n\rangle \quad , n = 0, 1, 2 \ldots \tag{8.106}$$

where we adopted the Dirac notation for the states of the quantum system. The states $|n\rangle$ are thought to form a complete, orthonormal basis, i.e., we assume

$$\langle n|m\rangle = \delta_{nm} \tag{8.107}$$

and for the identity

$$\mathbb{1} = \sum_{n=0}^{\infty}|n\rangle\langle n| \,. \tag{8.108}$$

We assume for the sake of simplicity that the eigenstates of H_o can be labeled through integers, i.e., we discount the possibility of a continuum of eigenstates. However, this assumption can be waved as our results below will not depend on it.

Estimate of the Magnitude of $\mathcal{V}_S$

We want to demonstrate now that the interaction $\mathcal{V}_S(t)$, as given in (8.105) for the case of radiation-induced transitions in atomic systems, can be considered a weak perturbation. In fact, one can estimate that the perturbation, in this case, is much smaller than the eigenvalue differences near typical atomic bound states, and that the first term in (8.105), i.e., the term $\sim \hat{\vec{p}}\cdot\vec{A}(\vec{r},t)$, is much larger than the second term, i.e., the term $\sim A^2(\vec{r},t)$. This result will allow us to neglect the second term in (8.105) in further calculations and to expand the wave function in terms of powers of $\mathcal{V}_S(t)$ in a perturbation calculation.
For an electron charge $q = -e$ and an electron mass $m = m_e$ one can provide the estimate for the first term of (8.105) as follows[4]. We first note, using (8.41)

$$\left|\frac{e}{m_e}\hat{\vec{p}}\cdot\vec{A}\right| \sim \frac{e}{m_e}\left|2m_e\frac{p^2}{2m_e}\right|^{\frac{1}{2}}\sqrt{\frac{8\pi\mathcal{N}_\omega\hbar}{\omega\,\mathcal{V}}} \,. \tag{8.109}$$

[4]The reader should note that the estimates are very crude since we are establishing an order of magnitude estimate only.

The virial theorem for the Coulomb problem provides the estimate for the case of a hydrogen atom

$$\left|\frac{p^2}{2m_e}\right| \sim \frac{1}{2}\frac{e^2}{a_o} \tag{8.110}$$

where a_o is the Bohr radius. Assuming a single photon, i.e., $\mathcal{N}_\omega = 1$, a volume $\mathcal{V} = \lambda^3$ where λ is the wave length corresponding to a plane wave with frequency ω, i.e., $\lambda = 2\pi c/\lambda$, one obtains for (8.109) using $\mathcal{V} = \lambda\, 4\pi^2 c^2/\omega^2$

$$\left|\frac{e}{m_e}\,\hat{\vec{p}}\cdot\vec{A}\right| \sim \frac{e^2}{4\pi a_o}\,\left|\frac{2}{\pi}\,\frac{a_o}{\lambda}\,\frac{\hbar\omega}{m_e c^2}\right|^{\frac{1}{2}} \tag{8.111}$$

For $\hbar\omega = 3$ eV and a corresponding $\lambda = 4000$ Å one obtains, with $a_o \approx 0.5$ Å, and $m_e c^2 \approx 500$ keV

$$\left|\frac{2}{\pi}\,\frac{a_o}{\lambda}\,\frac{\hbar\omega}{m_e c^2}\right| \approx 10^{-8} \tag{8.112}$$

and with $e^2/a_o \approx 27$ eV, altogether,

$$\left|\frac{e}{m_e}\,\hat{\vec{p}}\cdot\vec{A}\right| \sim 10\,\text{eV}\cdot 10^{-4} = 10^{-3}\,\text{eV}\,. \tag{8.113}$$

This magnitude is much less than the differences of the typical eigenvalues of the lowest states of the hydrogen atom which are of the order of 1 eV. Hence, the first term in (8.105) for radiation fields can be considered a small perturbation.
We want to estimate now the second term in (8.105). Using again (8.41) one can state

$$\left|\frac{e^2}{2m_e}\,A^2\right| \sim \frac{e^2}{2m_e}\,\frac{1}{\omega^2}\,\frac{8\pi\mathcal{N}_\omega\hbar\omega}{\mathcal{V}} \tag{8.114}$$

For the same assumptions as above one obtains

$$\left|\frac{e^2}{2m_e}\,A^2\right| \sim \frac{e^2}{8\pi a_o}\cdot\left(\frac{a_o}{\lambda}\,\frac{4\hbar\omega}{m_e c^2}\right)\,. \tag{8.115}$$

Employing for the second factor the estimate as stated in (8.112) yields

$$\left|\frac{e^2}{2m_e}\,A^2\right| \sim 10\,\text{eV}\cdot 10^{-8} = 10^{-7}\,\text{eV}\,. \tag{8.116}$$

This term is obviously much smaller than the first term. Consequently, one can neglect this term as long as the first term gives non-vanishing contributions, and as long as the photon densities $\mathcal{N}_\omega/\mathcal{V}$ are small. We can, hence, replace the perturbation (8.105) due to a radiation field by

$$\mathcal{V}_S = -\frac{q}{m}\,\hat{\vec{p}}\cdot\vec{A}(\vec{r},t)\,. \tag{8.117}$$

In case that such perturbation acts on an electron and is due to superpositions of planar waves described through the vector potential (8.42) it holds

$$\mathcal{V}_S \approx \frac{e}{m}\sum_{\vec{k},\hat{u}}\sqrt{\frac{4\pi\mathcal{N}_k\hbar}{k\mathcal{V}}}\,\alpha(\vec{k},\hat{u})\,\hat{\vec{p}}\cdot\hat{u}\,\exp\left[\,i(\vec{k}\cdot\vec{r}-\omega t)\right]\,. \tag{8.118}$$

where we have replaced ω in (8.42) through $k = |\vec{k}| = \omega$. The sum runs over all possible $\vec{k}$ vectors and might actually be an integral, the sum over $\hat{u}$ involves the two possible polarizations of planar electromagnetic waves. A factor $\alpha(\vec{k}, \hat{u})$ has been added to describe eliptically or circularly polarized waves. Equation (8.118) is the form of the perturbation which, under ordinary circumstances, describes the effect of a radiation field on an electron system and which will be assumed below to describe radiative transitions.

Perturbation Expansion

The generic situation we attempt to describe entails a particle at time $t = t_o$ in a state $|0\rangle$ and a radiation field beginning to act at $t = t_o$ on the particle promoting it into some of the other states $|n\rangle$, $n = 1, 2, \ldots$. The states $|0\rangle$, $|n\rangle$ are defined in (8.106–8.108) as the eigenstates of the unperturbed Hamiltonian H_o. One seeks to predict the probability to observe the particle in one of the states $|n\rangle$, $n \neq 0$ at some later time $t \geq t_o$. For this purpose one needs to determine the state $|\Psi_S(t)\rangle$ of the particle. This state obeys the Schrödinger equation

$$i\hbar\,\partial_t|\Psi_S(t)\rangle \;=\; [\,H_o \;+\; \mathcal{V}_S(t)\,]\,|\Psi_S(t)\rangle \tag{8.119}$$

subject to the initial condition

$$|\Psi_S(t_o)\rangle \;=\; |0\rangle\;. \tag{8.120}$$

The probability to find the particle in the state $|n\rangle$ at time t is then

$$p_{0\to n}(t) \;=\; |\langle n|\Psi_S(t)\rangle|^2\;. \tag{8.121}$$

In order to determine the wave function $\Psi_S(t)\rangle$ we choose the so-called Dirac representation defined through

$$|\Psi_S(t)\rangle \;=\; \exp\left[-\frac{i}{\hbar}\,H_o(t\,-\,t_o)\right]\,|\Psi_D(t)\rangle \tag{8.122}$$

where

$$|\Psi_D(t_o)\rangle \;=\; |0\rangle\;. \tag{8.123}$$

Using

$$i\hbar\,\partial_t\exp\left[-\frac{i}{\hbar}\,H_o(t\,-\,t_o)\right] \;=\; H_o\exp\left[-\frac{i}{\hbar}\,H_o(t\,-\,t_o)\right] \tag{8.124}$$

and (8.119) one obtains

$$\begin{aligned}&\exp\left[-\frac{i}{\hbar}\,H_o(t\,-\,t_o)\right]\,(\,H_o\;+\;i\hbar\,\partial_t\,)\,|\Psi_D(t)\rangle\\ &=\;[\,H_o\;+\;\mathcal{V}_S(t)\,]\,\exp\left[-\frac{i}{\hbar}\,H_o(t\,-\,t_o)\right]|\Psi_D(t)\rangle\end{aligned} \tag{8.125}$$

from which follows

$$\exp\left[-\frac{i}{\hbar}\,H_o(t\,-\,t_o)\right]\,i\hbar\,\partial_t|\Psi_D(t)\rangle \;=\; \mathcal{V}_S(t)\exp\left[-\frac{i}{\hbar}\,H_o(t\,-\,t_o)\right]|\Psi_D(t)\rangle\;. \tag{8.126}$$

Multiplying the latter equation by the operator $\exp\left[\frac{i}{\hbar}\,H_o(t\,-\,t_o)\right]$ yields finally

$$i\hbar\partial_t|\Psi_D(t)\rangle \;=\; \mathcal{V}_D(t)\,|\Psi_D(t)\rangle \quad\;,\; |\Psi(t_o)\rangle \;=\; |0\rangle \tag{8.127}$$

where

$$\mathcal{V}_D(t) \;=\; \exp\left[\frac{i}{\hbar}\,H_o(t\,-\,t_o)\right]\,\mathcal{V}_S(t)\exp\left[-\frac{i}{\hbar}\,H_o(t\,-\,t_o)\right]\;. \tag{8.128}$$

We note that the transition probability (8.121) expressed in terms of $\Psi_D(t)\rangle$ is

$$p_{0\to n}(t) \;=\; |\langle n|\exp\left[-\frac{i}{\hbar}\,H_o(t\,-\,t_o)\right]\,|\Psi_D(t)\rangle|^2\;. \tag{8.129}$$

Due to the Hermitean property of the Hamiltonian H_o holds $\langle n|H_o = \epsilon_n\langle n|$ and, consequently,

$$\langle n|\exp\left[-\frac{i}{\hbar}\,H_o(t\,-\,t_o)\right] \;=\; \exp\left[-\frac{i}{\hbar}\,\epsilon_n(t\,-\,t_o)\right]\,\langle n| \tag{8.130}$$

from which we conclude, using $|\exp[-\frac{i}{\hbar}\,\epsilon_n(t\,-\,t_o)]| = 1$,

$$p_{0\to n}(t) \;=\; |\langle n\,|\Psi_D(t)\rangle|^2\;. \tag{8.131}$$

In order to determine $|\Psi_D(t)\rangle$ described through (8.127) we assume the expansion

$$|\Psi_D(t)\rangle \;=\; \sum_{n=0}^{\infty}|\Psi_D^{(n)}(t)\rangle \tag{8.132}$$

where $|\Psi_D^{(n)}(t)\rangle$ accounts for the contribution due to n-fold products of $\mathcal{V}_D(t)$ to $|\Psi_D(t)\rangle$. Accordingly, we define $|\Psi_D^{(n)}(t)\rangle$ through the evolution equations

$$\begin{aligned}
i\hbar\partial_t|\Psi_D^{(0)}(t)\rangle &= 0 &(8.133)\\
i\hbar\partial_t|\Psi_D^{(1)}(t)\rangle &= \mathcal{V}_D(t)\,|\Psi_D^{(0)}(t)\rangle &(8.134)\\
i\hbar\partial_t|\Psi_D^{(2)}(t)\rangle &= \mathcal{V}_D(t)\,|\Psi_D^{(1)}(t)\rangle &(8.135)\\
&\vdots\\
i\hbar\partial_t|\Psi_D^{(n)}(t)\rangle &= \mathcal{V}_D(t)\,|\Psi_D^{(n-1)}(t)\rangle &(8.136)\\
&\vdots
\end{aligned}$$

together with the initial conditions

$$|\psi_D(t_o)\rangle \;=\; \begin{cases} |0\rangle & \text{for } n\,=\,0\\ 0 & \text{for } n\,=\,1,\,2.\,\ldots\end{cases} \tag{8.137}$$

One can readily verify that (8.132–8.137) are consistent with (8.127, 8.128).
Equations (8.133–8.137) can be solved recursively. We will consider here only the two leading contributions to $|\Psi_D(t)\rangle$. From (8.133, 8.137) follows

$$|\Psi_D^{(0)}(t)\rangle \;=\; |0\rangle\;. \tag{8.138}$$

Employing this result one obtains for (8.134, 8.137)

$$|\Psi_D^{(1)}(t)\rangle \;=\; \frac{1}{i\hbar}\int_{t_o}^{t}dt'\,\mathcal{V}_D(t')\,|0\rangle\;. \tag{8.139}$$

This result, in turn, yields for (8.135, 8.137)

$$|\Psi_D^{(2)}(t)\rangle \;=\; \left(\frac{1}{i\hbar}\right)^2 \int_{t_o}^{t} dt' \int_{t_o}^{t'} dt'' \mathcal{V}_D(t')\,\mathcal{V}_D(t'')\,|0\rangle\;. \tag{8.140}$$

Altogether we have provided the formal expansion for the transition amplitude

$$\begin{aligned}\langle n|\Psi_D(t)\rangle \;&=\; \langle n|0\rangle \;+\; \frac{1}{i\hbar}\int_{t_o}^{t} dt'\,\langle n|\mathcal{V}_D(t')\,|0\rangle \\ &\quad+\sum_{m=0}^{\infty}\left(\frac{1}{i\hbar}\right)^2\int_{t_o}^{t}dt'\int_{t_o}^{t'}dt''\langle n|\mathcal{V}_D(t')|m\rangle\langle m|\mathcal{V}_D(t'')\,|0\rangle \;+\;\ldots\end{aligned} \tag{8.141}$$

8.6 Perturbations due to Electromagnetic Radiation

We had identified in Eq. (8.118) above that the effect of a radiation field on an electronic system is accounted for by perturbations with a so-called harmonic time dependence $\sim \exp(-i\omega t)$. We want to apply now the perturbation expansion derived to such perturbations. For the sake of including the effect of superpositions of plane waves we will assume, however, that two planar waves simulataneously interact with an electronic system, such that the combined radiation field is decribed by the vector potential

$$\begin{aligned}\vec{A}(\vec{r},t) \;&=\; A_1\,\hat{u}_1\exp\left[i\,(\vec{k}_1\cdot\vec{r}\;-\;\omega_1\,t)\right] \quad \text{incoming wave} \\ &+\; A_2\,\hat{u}_2\exp\left[i\,(\vec{k}_2\cdot\vec{r}\;\mp\;\omega_2\,t)\right] \quad \text{incoming or outgoing wave}\end{aligned} \tag{8.142}$$

combining an incoming and an incoming or outgoing wave. The coefficients A_1, A_2 are defined through (8.41).
The resulting perturbation on an electron system, according to (8.118), is

$$\mathcal{V}_S \;=\; \left[\hat{V}_1\,\exp(-i\omega_1 t)\;+\;\hat{V}_2\,\exp(\mp i\omega_2 t)\right]\,e^{\lambda t}\;,\;\lambda\;\to\;0+\;,\;t_o\;\to\;-\infty \tag{8.143}$$

where $\hat{V}_1$ and $\hat{V}_2$ are time-independent operators defined as

$$\hat{V}_j \;=\; \underbrace{\frac{e}{m}\sqrt{\frac{8\pi\mathcal{N}_j\hbar}{\omega_j\mathcal{V}}}}_{\text{I}}\;\underbrace{\hat{\vec{p}}\cdot\hat{u}_j}_{\text{II}}\;\underbrace{e^{i\vec{k}\cdot\vec{r}}}_{\text{III}}\;. \tag{8.144}$$

Here the factor I describes the strength of the radiation field (for the specified planar wave) as determined through the photon density $\mathcal{N}_j/\mathcal{V}$ and the factor II describes the polarization of the planar wave; note that $\hat{u}_j$, according to (8.34, 8.142), defines the direction of the $\vec{E}$-field of the radiation. The factor III in (8.144) describes the propagation of the planar wave, the direction of the propagation being determined by $\hat{k} = \vec{k}/|\vec{k}|$. We will demonstrate below that the the sign of $\mp i\omega t$ determines if the energy of the planar wave is absorbed ("-" sign) or emitted ("+" sign) by the quantum system. In (8.144) $\vec{r}$ is the position of the electron and $\hat{\vec{p}} = (\hbar/i)\nabla$ is the momentum operator of the electron. A factor $\exp(\lambda t)$, $\lambda \to 0+$ has been introduced which describes that at time $t_o \to -\infty$ the perturbation is turned on gradually. This factor will serve mainly the purpose of keeping in the following derivation all mathematical quantities properly behaved, i.e., non-singular.

1st Order Processes

We employ now the perturbation (8.143) to the expansion (8.141). For the 1st order contribution to the transition amplitude

$$\langle n|\Psi_D^{(1)}(t)\rangle \;=\; \frac{1}{i\hbar}\int_{t_o}^{t} dt'\,\langle n|\mathcal{V}_D(t')\,|0\rangle \tag{8.145}$$

we obtain then, using (8.128), (8.130) and (for $m = 0$)

$$\exp\left[-\frac{i}{\hbar}\,H_o(t\,-\,t_o)\right]\,|m\rangle \;=\; \exp\left[-\frac{i}{\hbar}\,\epsilon_m(t\,-\,t_o)\right]\,|m\rangle\;, \tag{8.146}$$

for (8.145)

$$\begin{aligned}\langle n|\Psi_D^{(1)}(t)\rangle \quad &= \quad \lim_{\lambda\to 0+}\lim_{t\to-\infty}\frac{1}{i\hbar}\int_{t_o}^{t} dt'\exp\left[\frac{i}{\hbar}(\epsilon_n\,-\,\epsilon_o\,-\,i\hbar\lambda)\,t'\right]\,\times \\ &\qquad \times\,\left(\langle n|\hat{V}_1|0\rangle\,e^{-i\omega_1 t'}\;+\;\langle n|\hat{V}_2|0\rangle\,e^{\mp i\omega_2 t'}\right)\,.\end{aligned} \tag{8.147}$$

Carrying out the time integration and taking the limit $\lim_{t\to-\infty}$ yields

$$\begin{aligned}\langle n|\Psi_D^{(1)}(t)\rangle \quad &= \quad \lim_{\lambda\to 0+} e^{\lambda t}\,\left[\langle n|\hat{V}_1|0\rangle\,\frac{\exp\left[\frac{i}{\hbar}(\epsilon_n\,-\,\epsilon_o\,-\,\hbar\omega_1)\,t\right]}{\epsilon_o\,+\,\hbar\omega_1\,-\,\epsilon_n\,+\,i\lambda\hbar}\,+\right. \\ &\qquad \left.+\;\langle n|\hat{V}_2|0\rangle\,\frac{\exp\left[\frac{i}{\hbar}(\epsilon_n\,-\,\epsilon_o\,\mp\,\hbar\omega_2)\,t\right]}{\epsilon_o\,\pm\,\hbar\omega_2\,-\,\epsilon_n\,+\,i\lambda\hbar}\right]\,.\end{aligned} \tag{8.148}$$

2nd Order Processes

We consider now the 2nd order contribution to the transition amplitude. According to (8.140, 8.141) this is

$$\langle n|\Psi_D^{(2)}(t)\rangle \;=\; -\,\frac{1}{\hbar^2}\sum_{m=0}^{\infty}\int_{t_o}^{t} dt'\int_{t_o}^{t'} dt''\,\langle n|\mathcal{V}_D(t')\,|m\rangle\,\langle m|\mathcal{V}_D(t'')\,|0\rangle\;. \tag{8.149}$$

Using the definition of $\mathcal{V}_D$ stated in (8.128) one obtains

$$\langle k|\mathcal{V}_D(t)\,|\ell\rangle \;=\; \langle k|\mathcal{V}_S(t)|\ell\rangle\,\exp\left[\frac{i}{\hbar}\,(\epsilon_k\,-\,\epsilon_\ell)\right] \tag{8.150}$$

and, employing the perturbation (8.143), yields

$$\begin{aligned}&\langle n|\Psi_D^{(2)}(t)\rangle \;=\; -\,\frac{1}{\hbar^2}\,\lim_{\lambda\to 0+}\lim_{t\to-\infty}\sum_{m=0}^{\infty}\int_{t_o}^{t} dt'\int_{t_o}^{t'} dt'' \qquad (8.151)\\ &\left\{\;\langle n|\hat{V}_1\,|m\rangle\,\langle m|\hat{V}_1\,|0\rangle\,\exp\left[\frac{i}{\hbar}\,(\epsilon_n-\epsilon_m-\hbar\omega_1-i\hbar\lambda)t'\right]\exp\left[\frac{i}{\hbar}\,(\epsilon_m-\epsilon_o-\hbar\omega_1-i\hbar\lambda)t''\right]\right.\\ &\left.+\;\langle n|\hat{V}_2\,|m\rangle\,\langle m|\hat{V}_2\,|0\rangle\,\exp\left[\frac{i}{\hbar}\,(\epsilon_n-\epsilon_m\mp\hbar\omega_2-i\hbar\lambda)t'\right]\exp\left[\frac{i}{\hbar}\,(\epsilon_m-\epsilon_o\mp\hbar\omega_2-i\hbar\lambda)t''\right]\right.\end{aligned}$$

$$+ \langle n|\hat{V}_1\,|m\rangle\,\langle m|\hat{V}_2\,|0\rangle\,\exp\left[\frac{i}{\hbar}\,(\epsilon_n - \epsilon_m - \hbar\omega_1 - i\hbar\lambda)t'\right]\exp\left[\frac{i}{\hbar}\,(\epsilon_m - \epsilon_o \mp \hbar\omega_2 - i\hbar\lambda)t''\right]$$
$$+ \langle n|\hat{V}_2\,|m\rangle\,\langle m|\hat{V}_1\,|0\rangle\,\exp\left[\frac{i}{\hbar}\,(\epsilon_n - \epsilon_m \mp \hbar\omega_2 - i\hbar\lambda)t'\right]\exp\left[\frac{i}{\hbar}\,(\epsilon_m - \epsilon_o - \hbar\omega_1 - i\hbar\lambda)t''\right]\Bigg\}$$

Carrying out the integrations and the limit $\lim_{t\to-\infty}$ provides the result

$$\langle n|\Psi_D^{(2)}(t)\rangle \;=\; -\frac{1}{\hbar^2}\lim_{\lambda\to 0+}\sum_{m=0}^{\infty} \tag{8.152}$$
$$\Bigg\{\frac{\langle n|\hat{V}_1\,|m\rangle\,\langle m|\hat{V}_1\,|0\rangle}{\epsilon_m - \epsilon_o - \hbar\omega_1 - i\hbar\lambda}\;\frac{\exp\left[\frac{i}{\hbar}\,(\epsilon_n - \epsilon_o - 2\hbar\omega_1 - 2i\hbar\lambda)t\right]}{\epsilon_n - \epsilon_o - 2\hbar\omega_1 - 2i\hbar\lambda}$$
$$+\frac{\langle n|\hat{V}_2\,|m\rangle\,\langle m|\hat{V}_2\,|0\rangle}{\epsilon_m - \epsilon_o \mp \hbar\omega_2 - i\hbar\lambda}\;\frac{\exp\left[\frac{i}{\hbar}\,(\epsilon_n - \epsilon_o \mp 2\hbar\omega_2 - 2i\hbar\lambda)t\right]}{\epsilon_n - \epsilon_o \mp 2\hbar\omega_2 - 2i\hbar\lambda}$$
$$+\frac{\langle n|\hat{V}_1\,|m\rangle\,\langle m|\hat{V}_2\,|0\rangle}{\epsilon_m - \epsilon_o \mp \hbar\omega_2 - i\hbar\lambda}\;\frac{\exp\left[\frac{i}{\hbar}\,(\epsilon_n - \epsilon_o - \hbar\omega_1 \mp \hbar\omega_2 - 2i\hbar\lambda)t\right]}{\epsilon_n - \epsilon_o - \hbar\omega_1 \mp \hbar\omega_2 - 2i\hbar\lambda}$$
$$+\frac{\langle n|\hat{V}_2\,|m\rangle\,\langle m|\hat{V}_1\,|0\rangle}{\epsilon_m - \epsilon_o - \hbar\omega_1 - i\hbar\lambda}\;\frac{\exp\left[\frac{i}{\hbar}\,(\epsilon_n - \epsilon_o - \hbar\omega_1 \mp \hbar\omega_2 - 2i\hbar\lambda)t\right]}{\epsilon_n - \epsilon_o - \hbar\omega_1 \mp \hbar\omega_2 - 2i\hbar\lambda}\Bigg\}$$

1st Order Radiative Transitions

The 1st and 2nd order transition amplitudes (8.148) and (8.152), respectively, provide now the transition probability $p_{0\to n}(t)$ according to Eq. (8.131). We assume first that the first order transition amplitude $\langle n|\Psi_D^{(1)}(t)\rangle$ is non-zero, in which case one can expect that it is larger than the 2nd order contribution $\langle n|\Psi_D^{(2)}(t)\rangle$ which we will neglect. We also assume for the final state $n \neq 0$ such that $\langle n|0\rangle = 0$ and

$$p_{0\to n}(t) \;=\; |\,\langle n|\Psi_D^{(1)}(t)\rangle\,|^2 \tag{8.153}$$

holds. Using (8.148) and

$$|z_1 + z_2|^2 \;=\; |z_1|^2 \;+\; |z_2|^2 \;+\; 2\mathrm{Re}(z_1 z_2^*) \tag{8.154}$$

yields

$$\begin{aligned} p_{0\to n}(t) \quad = \quad & \lim_{\lambda\to 0+} e^{2\lambda t}\Bigg\{\frac{\langle n|\hat{V}_1|0\rangle|^2}{(\epsilon_o + \hbar\omega_1 - \epsilon_n)^2 \,+\, (\lambda\hbar)^2} \\ & \qquad + \frac{\langle n|\hat{V}_2|0\rangle|^2}{(\epsilon_o \pm \hbar\omega_1 - \epsilon_n)^2 \,+\, (\lambda\hbar)^2} \\ & \qquad + 2\,\mathrm{Re}\,\frac{\langle n|\hat{V}_1|0\rangle\langle 0|\hat{V}_2|n\rangle\,\exp\left[\frac{i}{\hbar}(\pm\hbar\omega_2 \,-\, \hbar\omega_1)\,t\right]}{(\epsilon_o \,+\, \hbar\omega_1 \,-\, \epsilon_n \,+\, i\lambda\hbar)\,(\epsilon_o \,\pm\, \hbar\omega_2 \,-\, \epsilon_n \,-\, i\lambda\hbar)}\Bigg\} \end{aligned} \tag{8.155}$$

We are actually interested in the transition rate, i.e., the time derivative of $p_{0\to n}(t)$. For this rate holds

$$\frac{d}{dt}\,p_{0\to n}(t) \;=\; \lim_{\lambda\to 0+} e^{2\lambda t}\,\Bigg\{ \frac{2\,\lambda\,\langle n|\hat{V}_1|0\rangle|^2}{(\epsilon_o+\hbar\omega_1-\epsilon_n)^2\,+\,(\lambda\hbar)^2} \tag{8.156}$$
$$+\,\frac{2\,\lambda\,\langle n|\hat{V}_2|0\rangle|^2}{(\epsilon_o\pm\hbar\omega_1-\epsilon_n)^2\,+\,(\lambda\hbar)^2}\;+\;\left(2\,\lambda\;+\;\frac{d}{dt}\right)\times$$
$$\times\,2\,\mathrm{Re}\,\frac{\langle n|\hat{V}_1|0\rangle\langle 0|\hat{V}_2|n\rangle\,\exp\left[\frac{i}{\hbar}(\pm\hbar\omega_2\,-\,\hbar\omega_1)\,t\,\right]}{(\epsilon_o\,+\,\hbar\omega_1\,-\,\epsilon_n\,+\,i\lambda\hbar)\,(\epsilon_o\,\pm\,\hbar\omega_2\,-\,\epsilon_n\,-\,i\lambda\hbar)}\Bigg\}$$

The period of electromagnetic radiation absorbed by electronic systems in atoms is of the order 10^{-17} s, i.e., is much shorter than could be resolved in any observation; in fact, any attempt to do so, due to the uncertainty relationship would introduce a considerable perturbation to the system. The time average will be denoted by $\langle\,\cdots\,\rangle_t$. Hence, one should average the rate over many periods of the radiation. The result of such average is, however, to cancel the third term in (8.156) such that the 1st order contributions of the two planar waves of the perturbation simply add. For the resulting expression the limit $\lim_{\lambda\to 0+}$ can be taken. Using

$$\lim_{\delta\to 0+}\frac{\epsilon}{x^2\,+\,\epsilon^2}\;=\;\pi\,\delta(x) \tag{8.157}$$

one can conclude for the average transition rate

$$k\;=\;\langle\,\frac{d}{dt}\,p_{0\to n}(t)\,\rangle_t \;=\; \frac{2\pi}{\hbar}\,\Big[\;|\langle n|\hat{V}_1|0\rangle|^2\,\delta(\epsilon_n\,-\,\epsilon_o\,-\,\hbar\omega_1) \tag{8.158}$$
$$+\quad|\langle n|\hat{V}_2|0\rangle|^2\,\delta(\epsilon_n\,-\,\epsilon_o\,\mp\,\hbar\omega_2)\;\Big]$$

Obviously, the two terms apearing on the rhs. of this expression describe the individual effects of the two planar wave contributions of the perturbation (8.142–8.144). The δ-functions appearing in this expression reflect energy conservation: the incoming plane wave contribution of (8.143, 8.144), due to the vector potential

$$A_1\,\hat{u}_1\exp\Big[i\,(\vec{k}_1\cdot\vec{r}\,-\,\omega_1\,t)\Big]\;, \tag{8.159}$$

leads to final states $|n\rangle$ with energy $\epsilon_n\,=\,\epsilon_o\,+\,\hbar\omega_1$. The second contribution to (8.158), describing either an incoming or an outgoing plane wave due to the vector potential

$$A_2\,\hat{u}_2\exp\Big[i\,(\vec{k}_1\cdot\vec{r}\,\mp\,\omega_2\,t)\Big]\;, \tag{8.160}$$

leads to final states $|n\rangle$ with energy $\epsilon_n\,=\,\epsilon_o\,\pm\,\hbar\omega_2$. The result supports our definition of *incoming* and *outgoing* waves in (8.31) and (8.142)

The matrix elements $\langle n|\hat{V}_1|0\rangle$ and $\langle n|\hat{V}_2|0\rangle$ in (8.158) play an essential role for the transition rates of radiative transitions. First, these matrix elements determine the so-called *selection rules* for the transition: the matrix elements vanish for many states $|n\rangle$ and $|0\rangle$ on the ground of symmetry and geometrical properties. In case the matrix elements are non-zero, the matrix elements can vary strongly for different states $|n\rangle$ of the system, a property, which is observed through the so-called spectral intensities of transitions $|0\rangle\;\to\;|n\rangle$.

2nd Order Radiative Transitions

We now consider situations where the first order transition amplitude in (8.153) vanishes such that the leading contribution to the transition probability $p_{0\to n}(t)$ arises from the 2nd order amplitude (8.152), i.e., it holds

$$p_{0\to n}(t) \;=\; |\,\langle n|\Psi_D^{(2)}(t)\rangle\,|^2\;. \tag{8.161}$$

To determine the transition rate we proceed again, as we did in the the case of 1st order transitions, i.e., in Eqs. (8.153–8.158). We define

$$\begin{aligned} z_1 \;&=\; \left(\sum_{m=0}^{\infty}\frac{\langle n|\hat{V}_1\,|m\rangle\,\langle m|\hat{V}_1\,|0\rangle}{\epsilon_m-\epsilon_o-\hbar\omega_1-i\hbar\lambda}\right)\times \\ &\qquad\times\;\frac{\exp\big[\frac{i}{\hbar}\,(\epsilon_n-\epsilon_o-2\hbar\omega_1-2i\hbar\lambda)t\big]}{\epsilon_n-\epsilon_o-2\hbar\omega_1-2i\hbar\lambda} \end{aligned} \tag{8.162}$$

and, similarly,

$$\begin{aligned} z_2 \;&=\; \left(\sum_{m=0}^{\infty}\frac{\langle n|\hat{V}_2\,|m\rangle\,\langle m|\hat{V}_2\,|0\rangle}{\epsilon_m-\epsilon_o\mp\hbar\omega_2-i\hbar\lambda}\right)\times \\ &\qquad\times\;\frac{\exp\big[\frac{i}{\hbar}\,(\epsilon_n-\epsilon_o\mp 2\hbar\omega_2-2i\hbar\lambda)t\big]}{\epsilon_n-\epsilon_o\mp 2\hbar\omega_2-2i\hbar\lambda} \end{aligned} \tag{8.163}$$

$$z_3 \;=\; \left[\sum_{m=0}^{\infty}\left(\frac{\langle n|\hat{V}_2\,|m\rangle\langle m|\hat{V}_1\,|0\rangle}{\epsilon_m-\epsilon_o-\hbar\omega_1-i\hbar\lambda}\;+\;\frac{\langle n|\hat{V}_2\,|m\rangle\langle m|\hat{V}_1\,|0\rangle}{\epsilon_m-\epsilon_o\mp\hbar\omega_2-i\hbar\lambda}\right)\right]\times \tag{8.164}$$

$$\qquad\times\;\frac{\exp\big[\frac{i}{\hbar}\,(\epsilon_n-\epsilon_o-\hbar\omega_1\mp\hbar\omega_2-2i\hbar\lambda)t\big]}{\epsilon_n-\epsilon_o-\hbar\omega_1\mp\hbar\omega_2-2i\hbar\lambda} \tag{8.165}$$

It holds

$$|z_1\,+\,z_2\,+\,z_3|^2 \;=\; |z_1|^2\,+\,|z_2|^2\,+\,|z_3|^2\,+\sum_{\substack{j,k=1\\ j\neq k}}^{3} z_j z_k^* \tag{8.166}$$

In this expression the terms $|z_j|^2$ exhibit only a time dependence through a factor $e^{2\lambda t}$ whereas the terms $z_j z^* k$ for $j\neq k$ have also time-dependent phase factors, e.g., $\exp[\frac{i}{\hbar}(\pm\omega_2-\omega_1)]$. Time average $\langle\,\cdots\,\rangle_t$ of expression (8.166) over many periods of the radiation yields $\langle\exp[\frac{i}{\hbar}(\pm\omega_2-\omega_1)]\rangle_t \;=\; 0$ and, hence,

$$\langle\,|z_1\,+\,z_2\,+\,z_3|^2\rangle_t \;=\; |z_1|^2\,+\,|z_2|^2\,+\,|z_3|^2 \tag{8.167}$$

Taking now the limit $\lim_{\lambda\to 0+}$ and using (8.157) yields, in analogy to (8.158),

$$\begin{aligned} k \;&=\; \langle\,\frac{d}{dt}\,p_{0\to n}(t)\,\rangle_t \\ &=\; \underbrace{\frac{2\pi}{\hbar}\,\left|\sum_{m=0}^{\infty}\frac{\langle n|\hat{V}_1\,|m\rangle\,\langle m|\hat{V}_1\,|0\rangle}{\epsilon_m-\epsilon_o-\hbar\omega_1-i\hbar\lambda}\right|^2\,\delta(\epsilon_m-\epsilon_o-2\hbar\omega_1)}_{\text{absorption of 2 photons }\hbar\omega_1} \end{aligned} \tag{8.168}$$

$$+ \underbrace{\frac{2\pi}{\hbar}\left|\sum_{m=0}^{\infty}\frac{\langle n|\hat{V}_2\,|m\rangle\,\langle m|\hat{V}_2\,|0\rangle}{\epsilon_m-\epsilon_o\mp\hbar\omega_2-i\hbar\lambda}\right|^2\delta(\epsilon_m-\epsilon_o\mp 2\hbar\omega_2)}_{\text{absorption/emission of 2 photons }\hbar\omega_2}$$

$$+ \underbrace{\frac{2\pi}{\hbar}\left|\sum_{m=0}^{\infty}\left(\frac{\langle n|\hat{V}_2\,|m\rangle\langle m|\hat{V}_1\,|0\rangle}{\epsilon_m-\epsilon_o-\hbar\omega_1-i\hbar\lambda}+\right.\right.}$$

$$\left.\left.+\frac{\langle n|\hat{V}_2\,|m\rangle\langle m|\hat{V}_1\,|0\rangle}{\epsilon_m-\epsilon_o\mp\hbar\omega_2-i\hbar\lambda}\right)\right|^2\delta(\epsilon_n-\epsilon_o-\hbar\omega_1\mp\hbar\omega_2)$$

$$\text{absorption of a photon }\hbar\omega_1\text{ and absorption/emission of a photon }\hbar\omega_2$$

This transition rate is to be interpreted as follows. The first term, according to its δ-function factor, describes processes which lead to final states $|n\rangle$ with energy $\epsilon_n = \epsilon_o + 2\hbar\omega_1$ and, accordingly, describe the absorption of two photons, each of energy $\hbar\omega_1$. Similarly, the second term describes the processes leading to final states $|n\rangle$ with energy $\epsilon_n = \epsilon_o \pm 2\hbar\omega_2$ and, accordingly, describe the absorption/emission of two photons, each of energy $\hbar\omega_2$. Similarly, the third term describes processes in which a photon of energy $\hbar\omega_1$ is absorbed and a second photon of energy $\hbar\omega_2$ is absorbed/emitted. The factors $|\cdots|^2$ in (8.168) describe the time sequence of the two photon absorption/ emission processes. In case of the first term in (8.168) the interpretation is

$$\sum_{m=0}^{\infty}\underbrace{\langle n|\hat{V}_1\,|m\rangle}_{\text{pert. }|n\rangle\leftarrow|m\rangle}\quad\underbrace{\frac{1}{\epsilon_m-\epsilon_o-\hbar\omega_1-i\hbar\lambda}}_{\text{virtually occupied state }|m\rangle}\quad\underbrace{\langle m|\hat{V}_1\,|0\rangle}_{\text{pert. }|m\rangle\leftarrow|0\rangle} \tag{8.169}$$

, i.e., the system is perturbed through absorption of a photon with energy $\hbar\omega_1$ from the initial state $|0\rangle$ into a state $|m\rangle$; this state is only virtually excited, i.e., there is no energy conservation necessary (in general, $\epsilon_m \neq \epsilon_o + \hbar\omega_1$) and the evolution of state $|m\rangle$ is described by a factor $1/(\epsilon_m-\epsilon_o-\hbar\omega_1-i\hbar\lambda)$; a second perturbation, through absorption of a photon, promotes the system then to the state $|n\rangle$, which is stationary and energy is conserved, i.e., it must hold $\epsilon_n = \epsilon_o + 2\hbar\omega_1$. The expression sums over all possible virtually occupied states $|m\rangle$ and takes the absolute value of this sum, i.e., interference between the contributions from all intermediate states $|m\rangle$ can arise. The remaining two contributions in (8.168) describe similar histories of the excitation process. Most remarkably, the third term in (8.168) describes two intermediate histories, namely absorption/ emission first of photon $\hbar\omega_2$ and then absorption of photon $\hbar\omega_1$ and, *vice versa*, first absorption of photon $\hbar\omega_1$ and then absorption/ emission of photon $\hbar\omega_2$.

8.7 One-Photon Absorption and Emission in Atoms

We finally can apply the results derived to describe transition processes which involve the absorption or emission of a single photon. For this purpose we will employ the transition rate as given in Eq. (8.158) which accounts for such transitions.

Absorption of a Plane Polarized Wave

We consider first the case of absorption of a monochromatic, plane polarized wave described through the complex vector potential

$$\vec{A}(\vec{r},t) \;=\; \sqrt{\frac{8\pi\mathcal{N}\hbar}{\omega\mathcal{V}}}\,\hat{u}\,\exp\left[\tfrac{1}{\hbar}(\vec{k}\cdot\vec{r}\,-\,\omega t)\right]\;. \tag{8.170}$$

We will employ only the real part of this potential, i.e., the vector potential actually assumed is

$$\vec{A}(\vec{r},t) \;=\; \sqrt{\frac{2\pi\mathcal{N}\hbar}{\omega\mathcal{V}}}\,\hat{u}\,\exp\left[\tfrac{1}{\hbar}(\vec{k}\cdot\vec{r}\,-\,\omega t)\right] \;+\; \sqrt{\frac{2\pi\mathcal{N}\hbar}{\omega\mathcal{V}}}\,\hat{u}\,\exp\left[\tfrac{1}{\hbar}(-\vec{k}\cdot\vec{r}\,+\,\omega t)\right]\;. \tag{8.171}$$

The perturbation on an atomic electron system is then according to (8.143, 8.144)

$$\mathcal{V}_S \;=\; \left[\,\hat{V}_1\,\exp(-i\omega t)\;+\;\hat{V}_2\,\exp(+i\omega t)\right]\,e^{\lambda t}\;,\;\lambda\,\to\,0+\;,\;t_o\,\to\,-\infty \tag{8.172}$$

where

$$\hat{V}_{1,2} \;=\; \frac{e}{m}\sqrt{\frac{2\pi\mathcal{N}\hbar}{\omega\mathcal{V}}}\,\hat{\vec{p}}\cdot\,\hat{u}\,e^{\pm i\vec{k}\cdot\vec{r}}\;. \tag{8.173}$$

Only the first term of (8.143) will contribute to the absorption process, the second term can be discounted in case of absorption. The absorption rate, according to (8.158), is then

$$k_{\rm abs} \;=\; \frac{2\pi}{\hbar}\,\frac{e^2}{m_e^2}\,\frac{2\pi\mathcal{N}\hbar}{\omega\mathcal{V}}\;\left|\,\hat{u}\cdot\langle n|\,\hat{\vec{p}}\,e^{i\vec{k}\cdot\vec{r}}\,|0\rangle\,\right|^2\;\delta(\epsilon_n\,-\,\epsilon_o\,-\,\hbar\omega) \tag{8.174}$$

Dipole Approximation We seek to evaluate the matrix element

$$\vec{M} \;=\; \langle n|\,\hat{\vec{p}}\,e^{i\vec{k}\cdot\vec{r}}\,|0\rangle\;. \tag{8.175}$$

The matrix element involves a spatial integral over the electronic wave functions associated with states $|n\rangle$ and $|0\rangle$. For example, in case of a radiative transition from the 1s state of hydrogen to one of its three 2p states, the wave functions are (n,ℓ,m denote the relevant quantum numbers)

$$\psi_{n=1,\ell=0,m=0}(r,\theta,\phi) \;\;=\;\; 2\sqrt{\frac{1}{a_o^3}}\,e^{-r/a_o}\,Y_{00}(\theta,\phi)\quad 1{\rm s} \tag{8.176}$$

$$\psi_{n=2,\ell=1,m}(r,\theta,\phi) \;\;=\;\; -\frac{1}{2}\sqrt{\frac{6}{a_o^3}}\,\frac{r}{a_o}\,e^{-r/2a_o}\,Y_{1m}(\theta,\phi)\quad 2{\rm p} \tag{8.177}$$

and the integral is

$$\begin{aligned}\vec{M} \;\;=\;\; \frac{\hbar\sqrt{6}}{ia_o^4}\int_0^\infty r^2dr\int_{-1}^1 d\cos\theta\int_0^{2\pi}d\phi\;r\,e^{-r/2a_o}\,Y_{1m}^*(\theta,\phi)\;\times\\ \times\nabla e^{i\vec{k}\cdot\vec{r}}e^{-r/a_o}\,Y_{00}(\theta,\phi)\end{aligned} \tag{8.178}$$

These wave functions make significant contributions to this integral only for r-values in the range $r\,<\,10\,a_o$. However, in this range one can expand

$$e^{i\vec{k}\cdot\vec{r}} \;\approx\; 1\;+\;i\vec{k}\cdot\vec{r}\;+\;\ldots \tag{8.179}$$

One can estimate that the absolute magnitude of the second term in (8.179) and other terms are never larger than $20\pi\, a_o/\lambda$. Using $|\vec{k}| \;=\; 2\pi/\lambda$, the value of the wave length for the 1s $\rightarrow$ 2p transition

$$\lambda \;=\; \frac{2\pi\hbar c}{\Delta E_{2p-1s}} \;=\; 1216 \text{ Å} \tag{8.180}$$

and $a_o \;=\; 0.529$ Å one concludes that in the significant integration range in (8.178) holds $e^{i\vec{k}\cdot\vec{r}} \approx 1 \,+\, \mathrm{O}(\frac{1}{50})$ such that one can approximate

$$e^{i\vec{k}\cdot\vec{r}} \;\approx\; 1\,. \tag{8.181}$$

One refers to this approximation as the *dipole approximation*.

Transition Dipole Moment A further simplification of the matrix element (8.175) can then be achieved and the differential operator $\hat{\vec{p}} \,=\, \frac{\hbar}{i}\nabla$ replaced by by the simpler multiplicative operator $\vec{r}$. This simplification results from the identity

$$\hat{\vec{p}} \;=\; \frac{m}{i\hbar}\,[\,\vec{r},\, H_o\,] \tag{8.182}$$

where H_o is the Hamiltonian given by (8.104) and, in case of the hydrogen atom, is

$$H_o \;=\; \frac{(\hat{\vec{p}})^2}{2m_e} \;+\; V(\vec{r}) \quad ,\, V(\vec{r}) \;=\, -\,\frac{e^2}{r}\,. \tag{8.183}$$

For the commutator in (8.182) one finds

$$\begin{aligned}
[\,\vec{r},\, H_o\,] \;&=\; [\,\vec{r},\, \frac{\hat{\vec{p}}^2}{2m_e}\,] \;+\; \underbrace{[\,\vec{r},\, V(\vec{r})\,]}_{=\,0} \\
&=\; \frac{1}{2m_e}\sum_{k=1}^{3}\hat{p}_k\,[\,\vec{r},\,\hat{p}_k\,] \;+\; \frac{1}{2m_e}\sum_{k=1}^{3}[\,\vec{r},\,\hat{p}_k\,]\,p_k
\end{aligned} \tag{8.184}$$

Using $\vec{r} \,=\, \sum_{j=1}^{3} x_j\hat{e}_j$ and the commutation property $[x_k,\, \hat{p}_j] \;=\; i\hbar\,\delta_{kj}$ one obtains

$$[\,\vec{r},\, H_o\,] \;=\; \frac{i\hbar}{m}\sum_{j,k=1}^{3} p_k\;\hat{e}_j\,\delta_{jk} \;=\; \frac{i\hbar}{m}\sum_{j,k=1}^{3} p_k\;\hat{e}_k \;=\; \frac{i\hbar}{m}\,\hat{\vec{p}} \tag{8.185}$$

from which follows (8.182).
We are now in a position to obtain an alternative expression for the matrix element (8.175). Using (8.181) and (8.182) one obtains

$$\vec{M} \;\approx\; \frac{m}{i\hbar}\,\langle n|\,[\vec{r},\, H_o]\,|0\rangle \;=\; \frac{m\,(\epsilon_o \,-\, \epsilon_n)}{i\hbar}\,\langle n|\vec{r}\,|0\rangle\,. \tag{8.186}$$

Insertion into (8.174) yields

$$k_{\text{abs}} \;=\; \frac{4\pi^2\,e^2\,\mathcal{N}\,\omega}{\mathcal{V}}\;\Big|\;\hat{u}\cdot\langle n|\,\hat{\vec{r}}\,|0\rangle\;\Big|^2\;\delta(\epsilon_n - \epsilon_o - \hbar\omega) \tag{8.187}$$

where we used the fact that due to the δ-function factor in (8.174) one can replace $\epsilon_n - \epsilon_o$ by $\hbar\omega$. The δ-function appearing in this expression, in practical situations, will actually be replaced by a distribution function which reflects (1) the finite life time of the states $|n\rangle$, $|0\rangle$, and (2) the fact that strictly monochromatic radiation cannot be prepared such that any radiation source provides radiation with a frequency distribution.

Absorption of Thermal Radiation

We want to assume now that the hydrogen atom is placed in an evironment which is sufficiently hot, i.e., a very hot oven, such that the thermal radiation present supplies a continuum of frequencies, directions, and all polarizations of the radiation. We have demonstrated in our derivation of the rate of one-photon processes (8.158) above that in first order the contributions of all components of the radiation field add. We can, hence, obtain the transition rate in the present case by adding the individual transition rates of all planar waves present in the oven. Instead of adding the components of all possible $\vec{k}$ values we integrate over all $\vec{k}$ using the following rule

$$\sum_{\vec{k}} \sum_{\hat{u}} \quad \Longrightarrow \quad \mathcal{V} \int_{-\infty}^{+\infty} \frac{k^2\, dk}{(2\pi)^3} \int d\hat{k} \sum_{\hat{u}} \tag{8.188}$$

Here $\int d\hat{k}$ is the integral over all orientations of $\vec{k}$. Integrating and summing accordingly over all contributions as given by (8.187) and using $k\,c = \omega$ results in the total absorption rate

$$k_{\rm abs}^{\rm (tot)} \;=\; \frac{e^2 \mathcal{N}_\omega\, \omega^3}{2\pi\, c^3 \hbar} \int d\hat{k} \sum_{\hat{u}} \Big|\, \hat{u} \cdot \langle n|\, \hat{\vec{r}}\, |0\rangle \,\Big|^2 \tag{8.189}$$

where the factor $1/\hbar$ arose from the integral over the δ-function.
In order to carry out the integral $\int d\hat{k}$ we note that $\hat{u}$ describes the possible polarizations of the planar waves as defined in (8.31–8.35). $\hat{k}$ and $\hat{u}$, according to (8.33) are orthogonal to each other. As a result, there are ony two linearly independent directions of $\hat{u}$ possible, say $\hat{u}_1$ and $\hat{u}_2$. The unit vectors $\hat{u}_1$, $\hat{u}_2$ and $\hat{k}$ can be chosen to point along the x_1, x_2, x_3-axes of a right-handed cartesian coordinate system. Let us assume that the wave functions describing states $|n\rangle$ and $|0\rangle$ have been chosen real such that $\vec{\rho} = \langle n|\vec{r}|0\rangle$ is a real, three-dimensional vector. The direction of this vector in the $\hat{u}_1$, $\hat{u}_2$, $\hat{k}$ frame is described by the angles ϑ, φ, the direction of $\hat{u}_1$ is described by the angles $\vartheta_1 = \pi/2, \varphi_1 = 0$ and of $\hat{u}_2$ by $\vartheta_2 = \pi/2, \varphi_2 = \pi/2$. For the two angles $\alpha = \angle(\hat{u}_1, \vec{\rho})$ and $\beta = \angle(\hat{u}_2, \vec{\rho})$ holds then

$$\cos\!\alpha \;=\; \cos\!\vartheta_1 \cos\!\vartheta \;+\; \sin\!\vartheta_1 \sin\!\vartheta \cos(\varphi_1 - \varphi) \;=\; \sin\!\vartheta \cos\!\varphi \tag{8.190}$$

and

$$\cos\!\beta \;=\; \cos\!\vartheta_2 \cos\!\vartheta \;+\; \sin\!\vartheta_2 \sin\!\vartheta \cos(\varphi_2 - \varphi) \;=\; \sin\!\vartheta \sin\!\varphi\,. \tag{8.191}$$

Accordingly, one can express

$$\sum_{\hat{u}} |\, \hat{u} \cdot \langle n|\, \vec{r}\, |0\rangle\, |^2 \;=\; |\rho|^2\, (\cos^2\!\alpha \;+\; \cos^2\!\beta\,) \;=\; \sin^2\!\theta\,. \tag{8.192}$$

and obtain

$$\int d\hat{k} \sum_{\hat{u}} \Big|\, \hat{u} \cdot \langle n|\, \hat{\vec{r}}\, |0\rangle \,\Big|^2 \;=\; |\vec{\rho}|^2 \int_0^{2\pi} \int_{-1}^{1} d\cos\!\vartheta\, (1 - \cos^2\!\vartheta) \;=\; \frac{8\pi}{3} \tag{8.193}$$

This geometrical average, finally, can be inserted into (8.189) to yield the total absorption rate

$$k_{\rm abs}^{\rm (tot)} \;=\; \mathcal{N}_\omega \;\frac{4\,e^2\,\omega^3}{3\,c^3\hbar}\;|\,\langle n|\vec{r}|0\rangle\,|^2 \quad ,\; \mathcal{N}_\omega \text{ photons } \textit{before} \text{ absorption.} \tag{8.194}$$

For absorption processes involving the electronic degrees of freedom of atoms and molecules this radiation rate is typicaly of the order of $10^9\,s^{-1}$. For practical evaluations we provide an expression which eliminates the physical constants and allows one to determine numerical values readily. For this purpose we use $\omega/c \,=\, 2\pi/\lambda$ and obtain

$$\frac{4\,e^2\,\omega^3}{3\,c^3\hbar} \;=\; \frac{32\pi^3}{3}\;\frac{e^2}{a_o\hbar}\frac{a_o}{\lambda^3} \;=\; 1.37\times 10^{19}\,\frac{1}{\rm s}\;\times\frac{a_o}{\lambda^3} \tag{8.195}$$

and

$$k_{\rm abs}^{\rm (tot)} \;=\; \mathcal{N}_\omega\; 1.37\times 10^{19}\,\frac{1}{\rm s}\;\times\frac{a_o}{\lambda}\;\frac{|\,\langle n|\vec{r}|0\rangle\,|^2}{\lambda^2} \quad , \tag{8.196}$$

where

$$\lambda \;=\; \frac{2\,\pi\,c\,\hbar}{\epsilon_n \,-\, \epsilon_o} \tag{8.197}$$

The last two factors in (8.194) combined are typically somewhat smaller than (1 Å/1000 Å)^3 = 10^{-9}. Accordingly, the absorption rate is of the order of $10^9\,{\rm s}^{-1}$ or 1/nanosecond.

Transition Dipole Moment The expression (8.194) for the absorption rate shows that the essential property of a molecule which determines the absorption rate is the so-called transition dipole moment $|\langle n|\,\vec{r}\,|0\rangle|$. The transition dipole moment can vanish for many transitions between stationary states of a quantum system, in particular, for atoms or symmetric molecules. The value of $|\langle n|\,\vec{r}\,|0\rangle|$ determines the strength of an optical transition. The most intensely absorbing molecules are long, linear molecules.

Emission of Radiation

We now consider the rate of emission of a photon. The radiation field is described, as for the absorption process, by planar waves with vector potential (8.171) and perturbation (8.172, 8.173). In case of emission only the second term $\hat{V}_2\,\exp(+i\omega t)$ in (8.173) contributes. Otherwise, the calculation of the emission rate proceeds as in the case of absorption. However, the resulting total rate of emission bears a different dependence on the number of photons present in the environment. This difference between emission and absorption is due to the quantum nature of the radiation field. The quantum nature of radiation manifests itself in that the number of photons $\mathcal{N}_\omega$ msut be an integer, i.e., $\mathcal{N}_\omega \,=\, 0,\,1,\,2\,,\ldots$. This poses, however, a problem in case of emission by quantum systems in complete darkness, i.e., for $\mathcal{N}_\omega \,=\, 0$. In case of a classical radiation field one would expect that emission cannot occur. However, a quantum mechanical treatment of the radiation field leads to a total emission rate which is proportional to $\mathcal{N}_\omega \,+\, 1$ where $\mathcal{N}_\omega$ is the number of photons *before* emission. This dependence predicts, in agreement with observations, that emission occurs even if no photon is present in the environment. The corresponding process is termed *spontaneous emission*. However, there is also a contribution to the emission rate which is proportional to $\mathcal{N}_\omega$

which is termed *induced emission* since it can be induced through radiation provided, e.g., in lasers. The total rate of emission, accordingly, is

$$\begin{aligned} k_{\rm em}^{\rm (tot)} &= \frac{4\,e^2\,\omega^3}{3\,c^3\hbar}\;|\,\langle n|\vec{r}|0\rangle\,|^2 \quad \text{(spontaneous emission)} \\ &\quad + \mathcal{N}_\omega\;\frac{4\,e^2\,\omega^3}{3\,c^3\hbar}\;|\,\langle n|\vec{r}|0\rangle\,|^2 \quad \text{(induced emission)} \\ &= (\,\mathcal{N}_\omega\,+\,1\,)\;\frac{4\,e^2\,\omega^3}{3\,c^3\hbar}\;|\,\langle n|\vec{r}|0\rangle\,|^2 \qquad (8.198) \end{aligned}$$

$$\mathcal{N}_\omega \text{ photons } \textit{before} \text{ emission.} \qquad (8.199)$$

Planck's Radiation Law

The postulate of the $\mathcal{N}_\omega + 1$ dependence of the rate of emission as given in (8.198) is consistent with Planck's radiation law which reflects the (boson) quantum nature of the radiation field. To demonstrate this property we apply the transition rates (8.195) and (8.198) to determine the stationary distribution of photons $\hbar\omega$ in an oven of temperature T. Let N_o and N_n denote the number of atoms in state $|0\rangle$ and $|n\rangle$, respectively. For these numbers holds

$$N_n\,/\,N_o \;=\; \exp[-(\epsilon_n - \epsilon_o)/k_B T] \qquad (8.200)$$

where k_B is the Boltzmann constant. We assume $\epsilon_n - \epsilon_o = \hbar\omega$. Under stationary conditions the number of hydrogen atoms undergoing an absorption process $|0\rangle \rightarrow |n\rangle$ must be the same as the number of atoms undergoing an emission process $|n\rangle \rightarrow |0\rangle$. Defining the rate of spontaneous emission

$$k_{sp} \;=\; \frac{4\,e^2\,\omega^3}{3\,c^3\hbar}\;|\,\langle n|\vec{r}|0\rangle\,|^2 \qquad (8.201)$$

the rates of absorption and emission are $\mathcal{N}_\omega k_{sp}$ and $(\mathcal{N}_\omega+1)k_{sp}$, respectively. The number of atoms undergoing absorption in unit time are $\mathcal{N}_\omega k_{sp} N_o$ and undergoing emission are $(\mathcal{N}_\omega+1)k_{sp}N_n$. Hence, it must hold

$$\mathcal{N}_\omega\;k_{sp}\;N_0 \;=\; (\mathcal{N}_\omega\,+\,1)\;k_{sp}\;N_n \qquad (8.202)$$

It follows, using (8.200),

$$\exp[-\hbar\omega/k_B T] \;=\; \frac{\mathcal{N}_\omega}{\mathcal{N}_\omega\,+\,1}\;. \qquad (8.203)$$

This equation yields

$$\mathcal{N}_\omega \;=\; \frac{1}{\exp[\hbar\omega/k_B T]\;-\;1}\;, \qquad (8.204)$$

i.e., the well-known Planck radiation formula.

8.8 Two-Photon Processes

In many important processes induced by interactions between radiation and matter two or more photons participate. Examples are radiative transitions in which two photons are absorbed or emitted or scattering of radiation by matter in which a photon is aborbed and another re-emitted. In the following we discuss several examples.

Two-Photon Absorption

The interaction of electrons with radiation, under ordinary circumstances, induce single photon absorption processes as described by the transition rate Eq. (8.187). The transition requires that the transition dipole moment $\langle n|\,\vec{r}\,|0\rangle$ does not vanish for two states $|0\rangle$ and $|n\rangle$. However, a transition between the states $|0\rangle$ and $|n\rangle$ may be possible, even if $\langle n|\,\vec{r}\,|0\rangle$ vanishes, but then requires the absorption of two photons. In this case one needs to choose the energy of the photons to obey

$$\epsilon_n \;=\; \epsilon_o \;+\; 2\,\hbar\omega\;. \tag{8.205}$$

The respective radiative transition is of 2nd order as described by the transition rate (8.168) where the first term describes the relevant contribution. The resulting rate of the transition depends on $\mathcal{N}_\omega^2$. The intense radiation fields of lasers allow one to increase transition rates to levels which can readily be observed in the laboratory.
The perturbation which accounts for the coupling of the electronic system and the radiation field is the same as in case of 1st order absorption processes and given by (8.172, 8.173); however, in case of absorption only $\hat{V}_1$ contributes. One obtains, dropping the index 1 characterizing the radiation,

$$\begin{aligned} k \;\;=\;\; & \frac{2\pi}{\hbar}\left(\frac{e^2}{m_e^2}\,\frac{2\pi\mathcal{N}_\omega\hbar}{\omega\mathcal{V}}\right)^2 \left|\sum_{m=0}^{\infty}\frac{\langle n|\hat{u}\cdot\hat{\vec{p}}\,e^{i\vec{k}\cdot\vec{r}}\,|m\rangle\,\langle m|\hat{u}\cdot\hat{\vec{p}}\,e^{i\vec{k}\cdot\vec{r}}\,|0\rangle}{\epsilon_m-\epsilon_o-\hbar\omega_1-i\hbar\lambda}\right|^2 \times \\ & \times\,\delta(\epsilon_m-\epsilon_o-2\hbar\omega)\;. \end{aligned} \tag{8.206}$$

Employing the dipole approximation (8.181) and using (8.182) yields, finally,

$$\begin{aligned} k \;\;=\;\; & \left(\frac{\mathcal{N}_\omega}{\mathcal{V}}\right)^2 \frac{8\pi^3e^4}{\hbar}\left|\sum_{m=0}^{\infty}\frac{(\epsilon_n-\epsilon_m)\,\hat{u}\cdot\langle n|\hat{\vec{r}}|m\rangle\,(\epsilon_m-\epsilon_o)\,\hat{u}\cdot\langle m|\hat{\vec{r}}|0\rangle}{\hbar\omega\,(\,\epsilon_m-\epsilon_o-\hbar\omega-i\hbar\lambda\,)}\right|^2 \\ & \times\,\delta(\epsilon_m-\epsilon_o-2\hbar\omega)\;. \end{aligned} \tag{8.207}$$

Expression (8.207) for the rate of 2-photon transitions shows that the transition $|0\rangle \rightarrow |n\rangle$ becomes possible through intermediate states $|m\rangle$ which become virtually excited through absorption of a single photon. In applying (8.207) one is, however, faced with the dilemma of having to sum over all intermediate states $|m\rangle$ of the system. If the sum in (8.207) does not converge rapidly, which is not necessarily the case, then expression (8.207) does not provide a suitable avenue of computing the rates of 2-photon transitions.

Scattering of Photons at Electrons – Kramers-Heisenberg Cross Section

We consider in the following the scattering of a photon at an electron governed by the Hamiltonian H_o as given in (8.104) with stationary states $|n\rangle$ defined through (8.106). We assume that a planar wave with wave vector $\vec{k}_1$ and polarization $\hat{u}_1$, as described through the vector potential

$$\vec{A}(\vec{r},t) \;=\; A_{o1}\,\hat{u}_1\cos(\vec{k}_1\cdot\vec{r}\,-\,\omega_1 t)\;, \tag{8.208}$$

has been prepared. The electron absorbs the radiation and emits immediately a second photon. We wish to describe an observation in which a detector is placed at a solid angle element $d\Omega_2 = \sin\theta_2\,d\theta_2\,d\phi_2$ with respect to the origin of the coordinate system in which the electron is described.

We assume that the experimental set-up also includes a polarizer which selects only radiation with a certain polarization $\hat{u}_2$. Let us assume for the present that the emitted photon has a wave vector $\vec{k}_2$ with cartesian components

$$\vec{k}_2 \;=\; k_2 \begin{pmatrix} \sin\theta_2 \cos\phi_2 \\ \sin\theta_2 \sin\phi_2 \\ \cos\theta_2 \end{pmatrix} \tag{8.209}$$

where the value of k_2 has been fixed; however, later we will allow the quantum system to select appropriate values. The vector potential describing the emitted plane wave is then

$$\vec{A}(\vec{r},t) \;=\; A_{o2}\,\hat{u}_2 \cos(\vec{k}_2 \cdot \vec{r} \,-\, \omega_2 t)\;. \tag{8.210}$$

The vector potential which describes both incoming wave and outgoing wave is a superposition of the potentials in (8.208, 8.210). We know already from our description in Section 8.6 above that the absorption of the radiation in (8.208) and the emission of the radiation in (8.210) is accounted for by the following contributions of (8.208, 8.210)

$$\vec{A}(\vec{r},t) \;=\; A^{+}_{o1}\,\hat{u}_1 \exp[\,i\,(\vec{k}_1 \cdot \vec{r} \,-\, \omega_1 t)\,] \;+\; A^{-}_{o2}\,\hat{u}_2 \exp[\,i\,(\vec{k}_2 \cdot \vec{r} \,-\, \omega_2 t)\,]\;. \tag{8.211}$$

The first term describes the absorption of a photon and, hence, the amplitude A^{+}_{o1} is given by

$$A^{+}_{o1} \;=\; \sqrt{\frac{8\pi \mathcal{N}_1 \hbar}{\omega_1 \mathcal{V}}} \tag{8.212}$$

where $\mathcal{N}_1/\mathcal{V}$ is the density of photons for the wave described by (8.208), i.e., the wave characterized through $\vec{k}_1$, $\hat{u}_1$. The second term in (8.211) accounts for the emitted wave and, according to the description of emission processes on page 229, the amplititude A^{-}_{o2} defined in (8.211) is

$$A^{-}_{o2} \;=\; \sqrt{\frac{8\pi\,(\mathcal{N}_2 + 1)\,\hbar}{\omega_1 \mathcal{V}}} \tag{8.213}$$

where $\mathcal{N}_2/\mathcal{V}$ is the density of photons characterized through $\vec{k}_2, \hat{u}_2$.
The perturbation which arises due to the vector potential (8.211) is stated in Eq. (8.105). In the present case we consider only scattering processes which absorb radiation corresponding to the vector potential (8.208) and emit radiation corresponding to the vector potential (8.210). The relevant terms of the perturbation (8.105) using the vector potantial (8.211) are given by

$$\begin{aligned} \mathcal{V}_S(t) \;=\; & \underbrace{\frac{e}{2m_e}\,\hat{\vec{p}}\cdot\left\{A^{+}_{o1}\hat{u}_1 \exp[i(\vec{k}_1\cdot\vec{r}-\omega_1 t)] \;+\; A^{-}_{o2}\hat{u}_2 \exp[-i(\vec{k}_2\cdot\vec{r}-\omega_2 t)]\right\}}_{\text{contributes in 2nd order}} \\ & + \underbrace{\frac{e^2}{4m_e}\,A^{+}_{o1}A^{-}_{o2}\,\hat{u}_1\cdot\hat{u}_2 \exp\{i[(\vec{k}_1-\vec{k}_2)\cdot\vec{r} \,-\, (\omega_1-\omega_2)\,t]\}}_{\text{contributes in 1st order}} \end{aligned} \tag{8.214}$$

The effect of the perturbation on the state of the electronic system is as stated in the perturbation expansion (8.141). This expansion yields, in the present case, for the components of the wave

function accounting for absorption and re-emission of a photon

$$\begin{aligned}
\langle n|\Psi_D(t)\rangle &= \langle n|0\rangle + \\
&+ \frac{1}{i\hbar}\frac{e^2}{4m_e} A^+_{o1}A^-_{o2}\,\hat{u}_1\cdot\hat{u}_2\,\langle n|0\rangle \int_{t_o}^{t} dt'\, e^{i(\epsilon_n-\epsilon_o-\hbar\omega_1+\hbar\omega_2+i\hbar\lambda)t'} \\
&+ \sum_{m=0}^{\infty}\left(\frac{1}{i\hbar}\right)^2 \frac{e^2}{4m_e^2} A^+_{o1}A^-_{o2}\ \times \\
&\quad \times \Big\{\ \hat{u}_1\cdot\langle n|\,\hat{\vec{p}}\,|m\rangle\ \hat{u}_2\cdot\langle m|\,\hat{\vec{p}}\,|0\rangle\ \times \\
&\qquad \times \int_{t_o}^{t} dt' \int_{t_o}^{t'} dt''\, e^{i(\epsilon_n-\epsilon_m-\hbar\omega_1+i\hbar\lambda)t'} e^{i(\epsilon_m-\epsilon_o+\hbar\omega_2+i\hbar\lambda)t''} \\
&\quad + \hat{u}_2\cdot\langle n|\,\hat{\vec{p}}\,|m\rangle\ \hat{u}_1\cdot\langle m|\,\hat{\vec{p}}\,|0\rangle\ \times \\
&\qquad \times \int_{t_o}^{t} dt' \int_{t_o}^{t'} dt''\, e^{i(\epsilon_n-\epsilon_m+\hbar\omega_2+i\hbar\lambda)t'} e^{i(\epsilon_m-\epsilon_o-\hbar\omega_1+i\hbar\lambda)t''} \Big\}
\end{aligned} \tag{8.215}$$

We have adopted the dipole approximation (8.181) in stating this result.
Only the second (1st order) and the third (2nd order) terms in (8.215) correspond to scattering processes in which the radiation field 'looses' a photon $\hbar\omega_1$ and 'gains' a photon $\hbar\omega_2$. Hence, only these two terms contribute to the scattering amplitude. Following closely the procedures adopted in evaluating the rates of 1st order and 2nd order radiative transitions on page 222–225, i.e., evaluating the time integrals in (8.215) and taking the limits $\lim_{t_o\to-\infty}$ and $\lim_{\lambda\to 0^+}$ yields the transition rate

$$\begin{aligned}
k &= \frac{2\pi}{\hbar}\,\delta(\epsilon_n-\epsilon_o-\hbar\omega_1+\hbar\omega_2)\ \Bigg|\ \frac{e^2}{4m_e^2}\,A^+_{o1}A^-_{o2}\,\hat{u}_1\cdot\hat{u}_2\,\langle n|0\rangle \\
&- \sum_m \frac{e^2}{4m_e} A^+_{o1}A^-_{o2}\left(\frac{\langle n|\hat{u}_1\cdot\hat{\vec{p}}\,|m\rangle\langle m|\hat{u}_2\cdot\hat{\vec{p}}\,|0\rangle}{\epsilon_m\,-\,\epsilon_o\,+\,\hbar\omega_2}\ +\ \frac{\langle n|\hat{u}_2\cdot\hat{\vec{p}}\,|m\rangle\langle m|\hat{u}_1\cdot\hat{\vec{p}}\,|0\rangle}{\epsilon_m\,-\,\epsilon_o\,-\,\hbar\omega_1}\right)\Bigg|^2
\end{aligned} \tag{8.216}$$

We now note that the quantum system has the freedom to interact with any component of the radiation field to produce the emitted photon $\hbar\omega_2$. Accordingly, one needs to integrate the rate as given by (8.216) over all available modes of the field, i.e., one needs to carry out the integration $\mathcal{V}(2\pi)^{-3}\int k_2^2 dk_2\cdots$. Inserting also the values (8.212, 8.213) for the amplitudes A^+_{o1} and A^-_{o2} results in the *Kramers-Heisenberg* formula for the scattering rate

$$\begin{aligned}
k &= \frac{\mathcal{N}_1 c}{\mathcal{V}}\, r_o^2\,\frac{\omega_2}{\omega_1}\,(\mathcal{N}_2+1)\,d\Omega_2\ \Big|\hat{u}_1\cdot\hat{u}_2\,\langle n|0\rangle \\
&- \frac{1}{m_e}\sum_m\left(\frac{\langle n|\hat{u}_1\cdot\hat{\vec{p}}\,|m\rangle\langle m|\hat{u}_2\cdot\hat{\vec{p}}\,|0\rangle}{\epsilon_m\,-\,\epsilon_o\,+\,\hbar\omega_2}\ +\ \frac{\langle n|\hat{u}_2\cdot\hat{\vec{p}}\,|m\rangle\langle m|\hat{u}_1\cdot\hat{\vec{p}}\,|0\rangle}{\epsilon_m\,-\,\epsilon_o\,-\,\hbar\omega_1}\right)\Bigg|^2
\end{aligned} \tag{8.217}$$

Here r_o denotes the classical electron radius

$$r_o\ =\ \frac{e^2}{m_e c^2}\ =\ 2.8\cdot 10^{-15}\,\mathrm{m}\,. \tag{8.218}$$

The factor $\mathcal{N}_1 c/\mathcal{V}$ can be interpreted as the flux of incoming photons. Accordingly, one can relate (8.217) to the scattering cross section defined through

$$d\sigma \;=\; \frac{\text{rate of photons arriving in the the solid angle element } d\Omega_2}{\text{flux of incoming photons}} \tag{8.219}$$

It holds then

$$\begin{aligned} d\sigma \;\;=\;\; & r_o^2\,\frac{\omega_2}{\omega_1}\,(\mathcal{N}_2+1)\,d\Omega_2\,\Big|\hat{u}_1\cdot\hat{u}_2\;\langle n|0\rangle \\ & -\frac{1}{m_e}\sum_m\Big(\frac{\langle n|\hat{u}_1\cdot\hat{\vec{p}}|m\rangle\langle m|\hat{u}_2\cdot\hat{\vec{p}}|0\rangle}{\epsilon_m\,-\,\epsilon_o\,+\,\hbar\omega_2}\;+\;\frac{\langle n|\hat{u}_2\cdot\hat{\vec{p}}|m\rangle\langle m|\hat{u}_1\cdot\hat{\vec{p}}|0\rangle}{\epsilon_m\,-\,\epsilon_o\,-\,\hbar\omega_1}\Big)\Big|^2 \end{aligned} \tag{8.220}$$

In the following we want to consider various applications of this formula.

Rayleigh Scattering

We turn first to an example of so-called elastic scattering, i.e., a process in which the electronic state remains unaltered after the scattering. Rayleigh scattering is defined as the limit in which the wave length of the scattered radiation is so long that none of the quantum states of the electronic system can be excited; in fact, one assumes the even stronger condition

$$\hbar\omega_1 \;<<\; |\epsilon_o\,-\,\epsilon_m| \quad \text{, for all states } |m\rangle \text{ of the electronic system} \tag{8.221}$$

Using $|n\rangle \;=\; |0\rangle$ and, consequently, $\omega_1 \;=\; \omega_2$, it follows

$$d\sigma \;=\; r_o^2\,(\mathcal{N}_2+1)\,d\Omega_2\;\;|\hat{u}_1\cdot\hat{u}_2\;-\;S(\hbar\omega)\,|^2 \tag{8.222}$$

where

$$S(\hbar\omega) \;=\; \frac{1}{m_e}\sum_m\Big(\frac{\langle 0|\hat{u}_1\cdot\hat{\vec{p}}|m\rangle\langle m|\hat{u}_2\cdot\hat{\vec{p}}|0\rangle}{\epsilon_m\,-\,\epsilon_o\,+\,\hbar\omega}\;+\;\frac{\langle 0|\hat{u}_2\cdot\hat{\vec{p}}|m\rangle\langle m|\hat{u}_1\cdot\hat{\vec{p}}|0\rangle}{\epsilon_m\,-\,\epsilon_o\,-\,\hbar\omega}\Big)\,. \tag{8.223}$$

Condition (8.221) suggests to expand $S(\hbar\omega)$

$$S(\hbar\omega) \;=\; S(0)\;+\;S'(0)\,\hbar\omega\;+\;\frac{1}{2}S''(0)(\hbar\omega)^2\;+\;\ldots \tag{8.224}$$

Using

$$\frac{1}{\epsilon_m\,-\,\epsilon_o\,\pm\,\hbar\omega} \;=\; \frac{1}{\epsilon_m\,-\,\epsilon_o}\;\mp\;\frac{\hbar\omega}{(\epsilon_m\,-\,\epsilon_o)^2}\;+\;\frac{(\hbar\omega)^2}{(\epsilon_m\,-\,\epsilon_o)^3}\;+\;\ldots \tag{8.225}$$

one can readily determine

$$S(0) \;\;=\;\; \sum_m\Big(\frac{\langle 0|\hat{u}_1\cdot\hat{\vec{p}}|m\rangle\langle m|\hat{u}_2\cdot\hat{\vec{p}}|0\rangle}{m_e\,(\epsilon_m\,-\,\epsilon_o)}\;+\;\frac{\langle 0|\hat{u}_2\cdot\hat{\vec{p}}|m\rangle\langle m|\hat{u}_1\cdot\hat{\vec{p}}|0\rangle}{m_e\,(\epsilon_m\,-\,\epsilon_o)}\Big) \tag{8.226}$$

$$S'(0) \;\;=\;\; \sum_m\Big(\frac{\langle 0|\hat{u}_2\cdot\hat{\vec{p}}|m\rangle\langle m|\hat{u}_1\cdot\hat{\vec{p}}|0\rangle}{m_e\,(\epsilon_m\,-\,\epsilon_o)^2}\;-\;\frac{\langle 0|\hat{u}_1\cdot\hat{\vec{p}}|m\rangle\langle m|\hat{u}_2\cdot\hat{\vec{p}}|0\rangle}{m_e\,(\epsilon_m\,-\,\epsilon_o)^2}\Big) \tag{8.227}$$

$$S''(0) \;\;=\;\; 2\sum_m\Big(\frac{\langle 0|\hat{u}_1\cdot\hat{\vec{p}}|m\rangle\langle m|\hat{u}_2\cdot\hat{\vec{p}}|0\rangle}{m_e\,(\epsilon_m\,-\,\epsilon_o)^3}\;+\;\frac{\langle 0|\hat{u}_2\cdot\hat{\vec{p}}|m\rangle\langle m|\hat{u}_1\cdot\hat{\vec{p}}|0\rangle}{m_e\,(\epsilon_m\,-\,\epsilon_o)^3}\Big) \tag{8.228}$$

These three expressions can be simplified using the expression (8.182) for $\hat{\vec{p}}$ and the expression (8.108) for the identity operator.
We want to simplify first (8.226). For this purpose we replace $\hat{\vec{p}}$ using (8.182)

$$\frac{\langle 0|\hat{u}_1\cdot\hat{\vec{p}}|m\rangle}{m_e\,(\epsilon_m\,-\,\epsilon_o)}\;=\;\frac{1}{i\hbar}\,\langle 0|\hat{u}_1\cdot\vec{r}|m\rangle\quad,\quad\frac{\langle m|\hat{u}_1\cdot\hat{\vec{p}}|0\rangle}{m_e\,(\epsilon_m\,-\,\epsilon_o)}\;=\;-\,\frac{1}{i\hbar}\,\langle m|\hat{u}_1\cdot\vec{r}|0\rangle \tag{8.229}$$

This transforms (8.226) into

$$S(0)\;=\;\frac{1}{i\hbar}\sum_m(\;\langle 0|\hat{u}_1\cdot\vec{r}|m\rangle\langle m|\hat{u}_2\cdot\hat{\vec{p}}|0\rangle\;-\;\langle 0|\hat{u}_2\cdot\hat{\vec{p}}|m\rangle\langle m|\hat{u}_1\cdot\vec{r}|0\rangle\;) \tag{8.230}$$

According to (8.108) this is

$$S(0)\;=\;\frac{1}{i\hbar}\,\langle 0|\hat{u}_1\cdot\vec{r}\,\hat{u}_2\cdot\hat{\vec{p}}\;-\;\hat{u}_2\cdot\hat{\vec{p}}\,\hat{u}_1\cdot\vec{r}|0\rangle\;. \tag{8.231}$$

The commutator property $[x_j,\,\hat{p}_k]\;=\;i\hbar\,\delta_{jk}$ yields finally

$$S(0)\;=\;\frac{1}{i\hbar}\sum_{j,k=1}^{3}(\hat{u}_1)_j\;(\hat{u}_2)_k\;\langle 0|[x_j,\,\hat{p}_k]|0\rangle\;=\;\sum_{j,k=1}^{3}(\hat{u}_1)_j(\hat{u}_2)_k\;\delta_{jk}\;=\;\hat{u}_1\cdot\hat{u}_2 \tag{8.232}$$

Obviously, this term cancels the $\hat{u}_1\cdot\hat{u}_2$ term in (8.222).
We want to prove now that expression (8.227) vanishes. For this purpose we apply (8.229) both to $\hat{u}_1\cdot\hat{\vec{p}}$ and to $\hat{u}_2\cdot\hat{\vec{p}}$ which results in

$$S'(0)\;=\;\frac{m_e}{\hbar^2}\sum_m(\;\langle 0|\hat{u}_2\cdot\vec{r}\,|m\rangle\langle m|\hat{u}_1\cdot\vec{r}\,|0\rangle\;-\;\langle 0|\hat{u}_1\cdot\vec{r}\,|m\rangle\langle m|\hat{u}_2\cdot\vec{r}\,|0\rangle\;)\;. \tag{8.233}$$

Employing again (8.108) yields

$$S'(0)\;=\;\frac{m_e}{\hbar^2}\,\langle 0|\,[\hat{u}_2\cdot\vec{r},\,\hat{u}_1\cdot\vec{r}\,]\,|0\rangle\;=\;0 \tag{8.234}$$

where we used for the second identity the fact that $\hat{u}_1\cdot\vec{r}$ and $\hat{u}_2\cdot\vec{r}$ commute.
$S''(0)$ given in (8.228) provides then the first non-vanishing contribution to the scattering cross section (8.222). Using again (8.229) both for the $\hat{u}_1\cdot\hat{\vec{p}}$ and the to $\hat{u}_2\cdot\hat{\vec{p}}$ terms in (8.228) we obtain

$$S''(0)\;=\;\frac{2m_e}{\hbar^2}\sum_m\left(\;\frac{\langle 0|\hat{u}_1\cdot\vec{r}\,|m\rangle\langle m|\hat{u}_2\cdot\vec{r}\,|0\rangle}{\epsilon_m\,-\,\epsilon_o}\;+\;\frac{\langle 0|\hat{u}_2\cdot\vec{r}\,|m\rangle\langle m|\hat{u}_1\cdot\vec{r}\,|0\rangle}{\epsilon_m\,-\,\epsilon_o}\right) \tag{8.235}$$

We can now combine eqs. (8.224, 8.232, 8.234, 8.235) and obtain the leading contribution to the expression (8.222) of the cross section for Rayleigh scattering

$$\begin{aligned}d\sigma\;\;&=\;\;r_o^2\,m_e^2\,\omega^4\,(\mathcal{N}_2+1)\,d\Omega_2\;\times\\&\quad\times\left|\sum_m\left(\;\frac{\langle 0|\hat{u}_1^*\cdot\vec{r}\,|m\rangle\langle m|\hat{u}_2\cdot\vec{r}\,|0\rangle}{\epsilon_m\,-\,\epsilon_o}\;+\;\frac{\langle 0|\hat{u}_2^*\cdot\vec{r}\,|m\rangle\langle m|\hat{u}_1\cdot\vec{r}\,|0\rangle}{\epsilon_m\,-\,\epsilon_o}\right)\right|^2\end{aligned} \tag{8.236}$$

We have applied here a modification which arises in case of complex polarization vectors $\hat{u}$ which describe circular and elliptical polarizaed light.
Expression (8.236) is of great practical importance. It explains, for example, the blue color of the sky and the polarization pattern in the sky which serves many animals, i.e., honey bees, as a compass.

Thomson Scattering

We consider again elastic scattering., i.e., $|n\rangle = |0\rangle$ and $\omega_1 = \omega_2 = \omega$ in (8.220), however, assume now that the scattered radiation has very short wave length such that

$$\omega >> |\epsilon_o - \epsilon_m| \quad , \text{ for all states } |m\rangle \text{ of the electronic system .} \tag{8.237}$$

The resulting scattering is called *Thomson scattering*. We want to assume, though, that the dipole approximation is still valid which restricts the applicability of the following derivation to

$$k_1 >> \frac{1}{a_o} \quad , \; a_o \text{ Bohr radius .} \tag{8.238}$$

One obtains immediately from (8.220)

$$d\sigma = r_o^2 \, (\mathcal{N}_2 + 1) \, d\Omega_2 \, |\hat{u}_1 \cdot \hat{u}_2|^2 \, . \tag{8.239}$$

We will show below that this expression decribes the non-relativistic limit of *Compton scattering*. To evaluate $|\hat{u}_1 \cdot \hat{u}_2|^2$ we assume that $\vec{k}_1$ is oriented along the x_3-axis and, hence, the emitted radiation is decribed by the wave vector

$$\vec{k}_2 = k_1 \begin{pmatrix} \sin\theta_2 \cos\phi_2 \\ \sin\theta_2 \sin\phi_2 \\ \cos\theta_2 \end{pmatrix} \tag{8.240}$$

We choose for the polarization of the incoming radiation the directions along the x_1- and the x_2-axes

$$\hat{u}_1^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad , \; \hat{u}_1^{(2)} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \tag{8.241}$$

Similarly, we choose for the polarization of the emitted radiation two perpendicular directions $\hat{u}_2^{(1)}$ and $\hat{u}_2^{(2)}$ which are also orthogonal to the direction of $\vec{k}_2$. The first choice is

$$\hat{u}_2^{(1)} = \frac{\vec{k}_2 \times \vec{k}_1}{|\vec{k}_2 \times \vec{k}_1|} = \begin{pmatrix} \sin\phi_2 \\ -\cos\phi_2 \\ 0 \end{pmatrix} \tag{8.242}$$

where the second identity follows readily from $\vec{k}_1 = \hat{e}_3$ and from (8.240). Since $\hat{u}_2^{(2)}$ needs to be orthogonal to $\vec{k}_2$ as well as to $\hat{u}_2^{(1)}$ the sole choice is

$$\hat{u}_2^{(2)} = \frac{\vec{k}_2 \times \hat{u}_2^{(1)}}{|\vec{k}_2 \times \hat{u}_2^{(1)}|} = \begin{pmatrix} \cos\theta_2 \cos\phi_2 \\ \cos\theta_2 \sin\phi_2 \\ -\sin\theta_2 \end{pmatrix} \tag{8.243}$$

The resulting scattering cross sections for the various choices of polarizations are

$$d\sigma = r_o^2 \, (\mathcal{N}_2 + 1) \, d\Omega_2 \times \begin{cases} \sin^2\phi_2 & \text{for } \hat{u}_1 = \hat{u}_1^{(1)} \, , \, \hat{u}_2 = \hat{u}_2^{(1)} \\ \cos^2\theta_2 \cos^2\phi_2 & \text{for } \hat{u}_1 = \hat{u}_1^{(1)} \, , \, \hat{u}_2 = \hat{u}_2^{(2)} \\ \cos^2\phi_2 & \text{for } \hat{u}_1 = \hat{u}_1^{(2)} \, , \, \hat{u}_2 = \hat{u}_2^{(1)} \\ \cos^2\theta_2 \sin^2\phi_2 & \text{for } \hat{u}_1 = \hat{u}_1^{(2)} \, , \, \hat{u}_2 = \hat{u}_2^{(2)} \end{cases} \tag{8.244}$$

In case that the incident radiation is not polarized the cross section needs to be averaged over the two polarization directions $\hat{u}_1^{(1)}$ and $\hat{u}_1^{(2)}$. One obtains then for the scattering cross section of unpolarized radiation

$$\overline{d\sigma} \;=\; r_o^2\,(\mathcal{N}_2+1)\,d\Omega_2 \;\times\; \begin{cases} \frac{1}{2} & \text{for } \hat{u}_2 \;=\; \hat{u}_2^{(1)} \\ \frac{1}{2}\cos^2\theta_2 & \text{for } \hat{u}_2 \;=\; \hat{u}_2^{(2)} \end{cases} \tag{8.245}$$

The result implies that even though the incident radiation is unpolarized, the scattered radiation is polarized to some degree. The radiation scattered at right angles is even completely polarized along the $\hat{u}_2^{(1)}$-direction.
In case that one measures the scattered radiation irrespective of its polarization, the resulting scattering cross section is

$$\overline{d\sigma}_{\rm tot} \;=\; \frac{r_o^2}{2}\,(\mathcal{N}_2+1)\,(\,1\;+\;\cos^2\theta_2\,)\,d\Omega_2\;. \tag{8.246}$$

This expression is the non-relativistic limit of the cross section of *Compton scattering*. The *Compton scattering* cross section which is derived from a model which treats photons and electrons as colliding relativistic particles is

$$\overline{d\sigma}_{\rm tot}^{\rm (rel)} \;=\; \frac{r_o^2}{2}\,(\mathcal{N}_2+1)\,\left(\frac{\omega_2}{\omega_1}\right)^2\,\left(\frac{\omega_1}{\omega_2}\;+\;\frac{\omega_2}{\omega_1}\;-\;\sin^2\theta_2\right)\,d\Omega_2 \tag{8.247}$$

where

$$\omega_2^{-1}\;-\;\omega_1^{-1}\;=\;\frac{\hbar}{m_e c^2}\,(\,1\;-\;\cos\theta_2\,) \tag{8.248}$$

One can readily show that in the non-relativistic limit, i.e., for $c\;\rightarrow\;\infty$ the Compton scattering cross section (8.247, 8.247) becomes identical with the Thomson scattering cross section (8.246).

Raman Scattering and Brillouin Scattering

We now consider ineleastic scattering described by the Kramers-Heisenberg formula. In the case of such scattering an electron system absorbs and re-emits radiation without ending up in the initial state. The energy deficit is used to excite the system. The excitation can be electronic, but most often involves other degrees of freedom. For electronic systems in molecules or crystals the degrees of freedom excited are nuclear motions, i.e., molecular vibrations or crystal vibrational modes. Such scattering is called *Raman scattering*. If energy is absorbed by the system, one speaks of *Stokes scattering*, if energy is released, one speaks of *anti-Stokes scattering*. In case that the nuclear degrees of freedom excited absorb very little energy, as in the case of excitations of accustical modes of crystals, or in case of translational motion of molecules in liquids, the scattering is termed *Brillouin scattering*.
In the case that the scattering excites other than electronic degrees of freedom, the states $|n\rangle$ etc. defined in (8.220) represent actually electronic as well as nuclear motions, e.g., in case of a diatomic molecule $|n\rangle\;=\;|\phi(elect.)_n,\,\phi(vibr.)_n\rangle$. Since the scattering is inelastic, the first term in (8.220) vanishes and one obtains in case of Raman scattering

$$d\sigma \;=\; r_o^2\,(\mathcal{N}_2+1)\,\frac{\omega_2}{\omega_1}\,d\Omega_2\,|\,\hat{u}_2\cdot\mathbf{R}\cdot\hat{u}_1\,|^2 \tag{8.249}$$

where $\mathbf{R}$ represents a 3×3-matrix with elements

$$R_{jk} = \frac{1}{m_e} \sum_m \left(\frac{\langle n| \hat{p}_j |m\rangle\langle m| \hat{p}_k |0\rangle}{\epsilon_m - \epsilon_o + \hbar\omega_2} + \frac{\langle n| \hat{p}_k |m\rangle\langle m| \hat{p}j |0\rangle}{\epsilon_m - \epsilon_o - \hbar\omega_1} \right) \tag{8.250}$$

$$\omega_2 = \omega_1 - (\epsilon_n - \epsilon_o)/\hbar \tag{8.251}$$

We define $\vec{x} \cdot \mathbf{R} \cdot \vec{y} = \sum_{j,k} x_j R_{jk} y_k$.
In case that the incoming photon energy $\hbar\omega_1$ is chosen to match one of the electronic excitations, e.g., $\hbar\omega_1 \approx \epsilon_m - \epsilon_o$ for a particular state $|m\rangle$, the Raman scattering cross section will be much enhanced, a case called *resonant Raman scattering*. Of course, no singlularity developes in such case due to the finite life time of the state $|m\rangle$. Nevertheless, the cross section for resonant Raman scattering can be several orders of magnitude larger than that of ordinary Raman scattering, a property which can be exploited to selectively probe suitable molecules of low concentration in bulk matter.

Chapter 9

Many–Particle Systems

In this chapter we develop the quantummechanical description of non-relativistic many–particle-systems. Systems to which this chapter applies appear in many disguises, as electrons in crystals, molecules and atoms, as photons in the electromagmetic field, as vibrations and combination of electrons and phonons in crystals, as protons and neutrons in nuclei and, finally, as quarks in mesons and baryons. The latter systems require, however, relativistic descriptions.
The interaction among many identical particles gives rise to a host of fascinating phenomena, e.g., collective excitations in nuclei, atoms and molecules, superconductivity, the quantum Hall effect. A description of the mentioned phenomena requires a sufficient account of the interactions among the particles and of the associated many–particle motions. In the following we will introduce the rudimentary tools, mainly those tools which are connected with effective single–particle descriptions.

9.1 Permutation Symmetry of Bosons and Fermions

We seek to determine the stationary states of systems of many *identical* particles as described through the stationary Schrödinger equation

$$H\,|\Psi\rangle \;=\; E\,|\Psi\rangle\;. \tag{9.1}$$

Here $|\Psi\rangle$ represents the state of the system which in the following will be considered in a representation defined by *space–spin–coordinates* $\vec{x}_j \;=\; (\vec{r}_j, \sigma_j)$. $\vec{r}_j$ and σ_j denote position and spin, respectively, of the j-th particle. The spin variable for electrons, for example, is $\sigma_j \;=\; \pm\frac{1}{2}$. We will assume systems composed of N particles, i.e., the particle index j runs from 1 to N. The wave function in the space–spin representation is then

$$\Psi(\vec{x}_1, \vec{x}_2, \ldots \vec{x}_N) \;=\; \langle \vec{x}_1, \vec{x}_2, \ldots \vec{x}_N\,|\Psi\rangle\;. \tag{9.2}$$

Permutations

The essential aspect of the systems described is that the particles are assumed to be identical. This has the consequence that one cannot distinguish between states which differ only in a permutation of particles. In fact, quantum theory dictates that all such states are identical, i.e., are not counted

repeatedly like degenerate states. However, not any state can be chosen as to represent the many–particle system, rather only states which obey a certain transformation property under permutations are acceptable. This property will be established now. In the following sections we will then introduce the *operators of second quantization* which provide a means to ascertain that manipulation of wave functions always leave the transformation property under permutations uncompromised.

The permutations we need to consider for a system of N particles are the elements of the group S_N, the set of all permutations of N objects. The elements $P \in S_N$ can be represented conveniently as $2 \times N$ matrices, the top row representing the numbers 1 to N labelling the N particles under consideration, and the second row showing the numbers 1 to N again, but in a different order, number k under the entry j of the first row indicating that particle j is switched with particle k. Examples for an eight particle system are

$$P_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 2 & 4 & 5 & 6 & 3 & 8 & 7 \end{pmatrix}; \; P_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 2 & 3 & 5 & 4 & 6 & 7 & 8 \end{pmatrix}. \tag{9.3}$$

P_2 which affects only two particles, i.e., particles 4 and 5, leaving all other particles unchanged, is called a transposition. Transpositions denoted by $T(j,k)$ are characterized by the two indices (j,k) of the two particles transposed, i.e., $(j,k) = (4,5)$ in case of P_2. We can then state $P_2 = T(4,5)$.

We will not discuss here at any depth the properties of the permutation groups S_N, even though these properties, in particular, the representations in the space of N–particle wave functions are extremely useful in dealing with N–particle systems. A reader interested in the quantum mechanical description of N particle systems is strongly encouraged to study these representations. For the following we will require only two properties of S_N, namely implicitly the group property, and explicitly the fact that any $P \in S_N$ can be given as a product of transpositions $T(j,k)$. The latter factorization has the essential property that the number of factors is either even or odd, i.e., a given P can either only be presented by even numbers of transpositions or by odd numbers of transpositions. One calls permutations of the first type *even permutations* and permutations of the second type *odd permutations*. The group S_N is then composed of two disjunct classes of even and of odd permutations. For the permutations P_1 and P_2 given above holds

$$P_1 \; = \; T(3,4)\,T(4,5)\,T(5,6)\,T(6,3)\,T(7,8)\;; \quad P_2 \; = \; T(4,5) \tag{9.4}$$

both permutations being obviously odd. The product $P_1 P_2$ of even and odd permutations P_1 and P_2 obey the following multiplication table:

$P_2 \setminus P_1$	even	odd
even	even	odd
odd	odd	even

We finally determine the action of permutations P on the wave functions (9.2). Denoting by $P(j)$ the image of j, i.e., in notation (9.3) the j-th index in the second row, one can state

$$P\,\Psi(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_j, \ldots, \vec{x}_N) \; = \; \Psi(\vec{x}_{P(1)}, \vec{x}_{P(2)}, \ldots, \ldots, \vec{x}_{P(j)}, \ldots, \vec{x}_{P(N)})\,. \tag{9.5}$$

Permutation Symmetry

A second consequence of the identical character of the particles under consideration is that the Hamiltonian must reflect the permutation symmetry. In fact, the Hamiltonian must be of the type

$$H \;=\; \sum_{j=1}^{N} F(\vec{x}_j) \;+\; \frac{1}{2}\sum_{j,k=1}^{N} G(\vec{x}_j, \vec{x}_k) \tag{9.6}$$

which describes one–particle interactions and two–particle interactions of the system. $G(\vec{x}_j, \vec{x}_k)$ is symmetric in its two arguments. The terms of the Hamiltonian will be discussed further below. Presently, it is essential that the functions $F(\,)$ and $G(\,,\,)$ are the same for all particles and for all pairs of particles, all particles being governed by the same interactions. From this property follows that H is independent of any specific assignment of indices to the particles (note that (9.5) implies that a permutation effectively changes the indices of the particles) and, hence, the Hamiltonian for permuted indices PHP^{-1} is identical to H. From the latter results

$$[H, P] \;=\; 0\,. \tag{9.7}$$

This property, in turn, implies that the permutation operators can be chosen in a diagonal representation. In fact, it is a postulate of quantum mechanics that for any description of many–particle systems the permutation operators must be diagonal. Since permutations do not necessarily commute, a diagonal representation can only berealized in a simpler group. In fact, the representation is either isomorphic to the group composed of only the '1', i.e., 1 together with multiplication, or the corresponding group of two elements $1, -1$. The first case applies to bosons, i.e. particles with integer spin values, the second group to fermions, i.e., particles with half–integer spin values. In the latter case all even permutations are represented by '1' and all odd permutations by '-1'. Obviously, both groups, i.e., {1} (trivial) and {1,-1}, provide a representation of the group structure represented by the multiplication table of even and odd permutations given above.
The boson and fermion property stated implies that for a system of N *bosons* holds

$$P\,|\Psi\rangle \;=\; |\Psi\rangle \quad \forall P \in S_N\,. \tag{9.8}$$

For *fermions* holds

$$P\,|\Psi\rangle \;=\; \epsilon_P\,|\Psi\rangle \quad \forall P \in S_N \tag{9.9}$$

where

$$\epsilon_P \;=\; \begin{cases} 1 & \text{for } P \text{ even} \\ -1 & \text{for } P \text{ odd} \end{cases}\,. \tag{9.10}$$

We will construct now wave functions which exhibit the appropriate properties.

Fock-Space

Our derivation will start by placing each particle in a set of S *orthonormal* single particle states $\langle\vec{x}|k\rangle \;=\; \phi_k(\vec{x})$, $k = 1, 2, \ldots S$ where S denotes the number of available single particle states, usually a very large number. The single particle states are assumed to be orthonormal, i.e., for the scalar product holds

$$\langle j|k\rangle \;=\; \delta_{j\,k}\,. \tag{9.11}$$

The scalar product involves spin as well as space coordinates and is explicitly given by

$$\langle j|k\rangle \;=\; \int d^3r\, \phi_j^*(\vec{r})\, \phi_k(\vec{r}) \;\cdot\; \langle \sigma_j|\sigma_k\rangle \tag{9.12}$$

where the second factor represents the scalar product between spin states. We first consider orthonormal many–particle wave functions, the so-called Fock-states, which do not yet obey the symmetries (9.8, 9.9). In a second step we form linear combinations of Fock states with the desired symmetries.
The Fock states represent N particles which are placed into S states $|\lambda_1\rangle, |\lambda_2\rangle, \ldots, |\lambda_N\rangle$, $\lambda_j \in \{1, 2, \ldots S\}$ of the type (9.11, 9.12), each particle j into a specific state $|\lambda_j\rangle$, i.e.,

$$\langle \vec{x}_1, \vec{x}_2, \ldots \vec{x}_N|\Psi^{\mathrm{Fock}}(\lambda_1, \lambda_2, \ldots, \lambda_N)\rangle \;=\; \prod_{j=1}^{N} \phi_{\lambda_j}(\vec{x}_j)\;. \tag{9.13}$$

These states form an orthonormal basis of N-particle states. It holds

$$\langle \Psi^{\mathrm{Fock}}(\lambda_1', \lambda_2', \ldots, \lambda_N')|\Psi^{\mathrm{Fock}}(\lambda_1, \lambda_2, \ldots, \lambda_N)\rangle \;=\; \prod_{j=1}^{N} \delta_{\lambda_j' \lambda_j}\;. \tag{9.14}$$

Obviously, the states $|\Psi^{\mathrm{Fock}}(\lambda_1, \lambda_2, \ldots, \lambda_N)\rangle$ do not exhibit the symmetries dictated by (9.8, 9.9).

Wave Functions with Boson Symmetry in the Occupation Number Representation

Wave functions with the proper symmetries can be obtained as linear combinations of Fock states. A wave function which obeys the boson symmetry (9.8) is

$$\langle \vec{x}_1, \vec{x}_2, \ldots \vec{x}_N|\Psi^{\mathrm{B}}(n_1, n_2, \ldots, n_S)\rangle \;= \tag{9.15}$$

$$\left[\frac{\prod_{j=1}^{S}(n_j!)}{N!}\right]^{\frac{1}{2}} \sum\nolimits_{P\in\Lambda(\lambda_1,\ldots\lambda_N)} P\,\langle \vec{x}_1, \vec{x}_2, \ldots \vec{x}_N|\Psi^{\mathrm{Fock}}(\lambda_1, \lambda_2, \ldots\; .\lambda_N)\rangle$$

Here the states $(\lambda_1, \lambda_2, \ldots\; .\lambda_N)$ on the rhs. is a particular choice of placing N particles consistent with the occupation numbers $(n_1, n_2, \ldots, n_S)$. $\Lambda(\lambda_1, \ldots, \lambda_N) \subset S_N$ is the set of all permutations involving only particles j with *different* λ_j, e.g., $T(j,k) \in \Lambda(\lambda_1, \ldots, \lambda_N)$ only if $\lambda_j \neq \lambda_k$. The integers n_j are equal to the number of λ_k in (9.15) with $\lambda_k = j$, i.e. the n_j specify how often a single particle state $|j\rangle$ is occupied. The numbers n_j, referred to as the occupation numbers, characterize the wave function (9.15) completely and, therefore, are essential. The reader is well advised to pause and grasp the definition of the n_j.
An important detail of the definition (9.15) is that the sum over permutations does not involve permutations among indices $j, k, \ldots$ with $\lambda_j = \lambda_k = \ldots$. This detail is connected with the fact that a two–particle state of the type $\Psi^{\mathrm{Fock}}(\lambda, \lambda)$ in which two particles occupy the *same* single–particle state $|\lambda\rangle$ does not allow transposition of particles; the reason is that such transposition duplicates the state which would, hence, appear severalfold in (9.15).

The wave functions (9.15) describe, for example, the photons of the electromagnetic field or the phonons in a crystal. These states form an orthonormal basis, i.e., for two sets of occupation numbers $\mathcal{N} = (n_1, n_2, \ldots, n_S)$ and $\mathcal{N}' = (n'_1, n'_2, \ldots, n'_S)$ holds

$$\langle \Psi^{\mathrm{B}}(\mathcal{N}') | \Psi^{\mathrm{B}}(\mathcal{N}) \rangle \; = \; \prod_{j=1}^{S} \delta_{n_j \, n'_j} \; . \tag{9.16}$$

Exercise 9.1.1:
(a) Show that the states (9.15) obey the boson symmetry (9.8).
(b) Show that for the states (9.15) holds (9.16).

Wave Functions with Fermion Symmetry in the Occupation Number Representation

One can construct similarly a wave function which obeys the fermion symmetry property (9.9). Such wave function is

$$\langle \vec{x}_1, \vec{x}_2, \ldots \vec{x}_N | \Psi^{\mathrm{F}}(n_1, n_2, \ldots, n_S) \rangle \; =$$

$$\tfrac{1}{\sqrt{N!}} \, \textstyle\sum_{P \in S_N} \, \epsilon_P \, P \, \langle \vec{x}_1, \vec{x}_2, \ldots \vec{x}_N | \Psi^{\mathrm{Fock}}(\lambda_1, \lambda_2, \ldots \, . \lambda_N) \rangle \tag{9.17}$$

Here the states $(\lambda_1, \lambda_2, \ldots \, . \lambda_N)$ on the rhs. correspond to a particular choice of placing N particles consistent with the occupation numbers $(n_1, n_2, \ldots, n_S)$. Using the identity governing the determinant of $N \times N$–matrices $\mathbf{A}$

$$det(\mathbf{A}) \; = \; \sum_{P \in S_N} \, \epsilon_P \, \prod_{j=1}^{N} A_{j \, P(j)} \tag{9.18}$$

this wave function can also be expressed through the so-called *Slater determinant*

$$\langle \vec{x}_1, \vec{x}_2, \ldots \vec{x}_N | \Psi^{\mathrm{F}}(n_1, n_2, \ldots, n_S) \rangle \; =$$

$$\tfrac{1}{\sqrt{N!}} \begin{vmatrix} \phi_{\lambda_1}(\vec{x}_1) & \phi_{\lambda_2}(\vec{x}_1) & \ldots & \phi_{\lambda_N}(\vec{x}_1) \\ \phi_{\lambda_1}(\vec{x}_2) & \phi_{\lambda_2}(\vec{x}_2) & \ldots & \phi_{\lambda_N}(\vec{x}_2) \\ \vdots & \vdots & \vdots & \vdots \\ \phi_{\lambda_1}(\vec{x}_N) & \phi_{\lambda_2}(\vec{x}_N) & \ldots & \phi_{\lambda_N}(\vec{x}_N) \end{vmatrix} \tag{9.19}$$

Here, the integers n_j denote the occupancy of the single–particle state $|j\rangle$. The important property holds that n_j can only assume the two values $n_j = 0$ or $n_j = 1$. For any value $n_j > 1$ two or more of the columns of the Slater matrix are identical and the wave function vanishes.
The fermion states (9.17) describe, for example, electrons in an atom, a molecule or a crystal. These states form an orthonormal basis, i.e., for two sets of occupation numbers $\mathcal{N} = (n_1, n_2, \ldots, n_S)$ and $\mathcal{N}' = (n'_1, n'_2, \ldots, n'_S)$ holds

$$\langle \Psi^{\mathrm{F}}(\mathcal{N}') | \Psi^{\mathrm{F}}(\mathcal{N}) \rangle \; = \; \prod_{j=1}^{S} \delta_{n_j \, n'_j} \; . \tag{9.20}$$

The representation of wave functions (9.15) and (9.17) is commonly referred to as the *occupation number representation* since the wave functions are uniquely specified by the set of occupation numbers $\mathcal{N} = (n_1, n_2, \ldots, n_S)$.

Exercise 9.1.2:
(a) Show that the states (9.17) obey the fermion symmetry (9.9).
(b) Show that the states (9.17) obey the orthonormality property (9.20).

9.2 Operators of 2nd Quantization

An important tool to describe mathematically systems of many bosons or fermions, guaranteeing the correct permutation properties of the many–particle states, are the so-called operators of 2nd quantization. These operators allow one to construct and manipulate many–particle wave functions while preserving permutation symmetry.

Creation and Annihilation Operators for Bosons

We consider the operator a_j defined through

$$a_j \, \Psi^{\mathrm{B}}(n_1, \ldots, n_j, \ldots, n_S) = \begin{cases} \sqrt{n_j} \, \Psi^{\mathrm{B}}(n_1, \ldots, n_j - 1, \ldots, n_S) & n_j \geq 1 \\ 0 & n_j = 0 \end{cases} . \tag{9.21}$$

The factor $\sqrt{n_j}$ appears here on the rhs. since both the N-particle wave function $\Psi^{\mathrm{B}}(n_1, \ldots, n_j, \ldots, n_S)$ and the $N-1$-particle wave function $\Psi^{\mathrm{B}}(n_1, \ldots, n_j - 1, \ldots, n_S)$ are normalized according to (9.15). For the adjoint operator $a_j^\dagger$ holds

$$a_j^\dagger \Psi^{\mathrm{B}}(n_1, \ldots, n_j, \ldots, n_S) \; = \; \sqrt{n_j + 1} \; \Psi^{\mathrm{B}}(n_1, \ldots, n_j + 1, \ldots, n_S) \, . \tag{9.22}$$

The operators a_j and $a_j^\dagger$ obey the commutation properties

$$[a_j, a_k] \; = \; 0 \, , \quad [a_j^\dagger, a_k^\dagger] \; = \; 0 \tag{9.23}$$

$$[a_j, a_k^\dagger] \; = \delta_{j\,k} \, . \tag{9.24}$$

The operators $a_j^\dagger$ and a_j are referred to as the creation and annihilation operators of bosons for the orthonormal single–particle states $|j\rangle$.
To prove that (9.22) follows from (9.21) we consider the matrix $\mathbb{A}_j$ corresponding to a_j in the basis of many–particle states (9.15). Employing the superindices $\mathcal{N}'$ and $\mathcal{N}$ introduced above and using the orthonormality property (9.16) one obtains

$$(\mathbb{A}_j)_{\mathcal{N}'\mathcal{N}} \; = \; \langle \Psi^{\mathrm{B}}(\mathcal{N}') | \, a_j \, | \Psi^{\mathrm{B}}(\mathcal{N}) \rangle \; = \; \sqrt{n_j} \, \delta_{n'_j \, n_j - 1} \prod_{\substack{k=1 \\ k \neq j}}^{S} \delta_{n'_k n_k} \, . \tag{9.25}$$

Let $\mathbb{A}_j^\dagger$ be the matrix adjoint to $\mathbb{A}_j$. One obtains for its matrix elements

$$\begin{aligned}\left(\mathbb{A}_j^\dagger\right)_{\mathcal{N}'\mathcal{N}} &= \langle\Psi^{\mathrm{B}}(\mathcal{N}')|\, a_j^\dagger\,|\Psi^{\mathrm{B}}(\mathcal{N})\rangle \;=\; \overline{(\mathbb{A}_j)_{\mathcal{N}\mathcal{N}'}} \\ \sqrt{n'_j}\,\delta_{n_j\,n'_j-1}\prod_{\substack{k=1\\k\neq j}}^{S}\delta_{n_k n'_k} &= \sqrt{n_j+1}\,\delta_{n'_j\,n_j+1}\prod_{\substack{k=1\\k\neq j}}^{S}\delta_{n'_k n_k}\,. \end{aligned} \tag{9.26}$$

From this result one can immediately conclude that (9.22) follows from (9.21) and *vice versa*.
We will prove the commutation properties (9.24), the properties (9.23) follow in an analogous way. It holds for $j = k$

$$(a_j a_j^\dagger \;-\; a_j^\dagger a_j)|\Psi(\mathcal{N})\rangle \;=\; (n_j + 1 \;-\; n_j)|\Psi(\mathcal{N})\rangle \;=\; |\Psi(\mathcal{N})\rangle\,. \tag{9.27}$$

Since this holds for any $|\Psi(\mathcal{N})\rangle$ it follows $[a_j, a_j^\dagger] = \mathbb{1}$. For $j \neq k$ follows similarly

$$\begin{aligned}(a_j a_k^\dagger \;-\; a_k^\dagger a_j)|\Psi(\mathcal{N})\rangle \;=\; \\ (\sqrt{n_j(n_k+1)} \;-\; \sqrt{(n_k+1)n_j})|\Psi(n_1,\ldots n_j-1\ldots n_k+1\ldots n_S)\rangle \;=\; 0\,.\end{aligned} \tag{9.28}$$

Equations (9.27, 9.29) yield the commutation relationship (9.24).
The creation operators $a_j^\dagger$ allow one to construct boson states $\Psi^{\mathrm{B}}(\mathcal{N})\rangle$ from the vacuum state

$$|0\rangle \;=\; |\Psi^{\mathrm{B}}(n_1 = 0, n_2 = 0, \ldots, n_S = 0)\rangle \tag{9.29}$$

as follows

$$|\Psi^{\mathrm{B}}(n_1, n_2, \ldots, n_S)\rangle \;=\; \prod_{j=1}^{S}\frac{\left(a_j^\dagger\right)^{n_j}}{\sqrt{n_j!}}\,|0\rangle\,. \tag{9.30}$$

Of particular interest is the operator

$$\hat{N}_j \;=\; a_j^\dagger a_j\,. \tag{9.31}$$

This operator is diagonal in the occupation number representation, i.e., for basis states $|\Psi^{\mathrm{B}}(n_1, n_2, \ldots, n_S)\rangle = |\Psi^{\mathrm{B}}(\mathcal{N})\rangle$. One can readily show using (9.21, 9.22)

$$\hat{N}_j\,|\Psi^{\mathrm{B}}(\mathcal{N})\rangle \;=\; n_j\,|\Psi^{\mathrm{B}}(\mathcal{N})\rangle\,, \tag{9.32}$$

i.e., the eigenvalues of $\hat{N}_j$ are the occupation numbers n_j. One refers to $\hat{N}_j$ as the *occupation number operator*. The related operator

$$\hat{N} \;=\; \sum_{j=1}^{S}\hat{N}_j \tag{9.33}$$

is called the *particle number operator* since, obviously,

$$\begin{aligned}\hat{N}\,|\Psi^{\mathrm{B}}(n_1, n_2, \ldots, n_S)\rangle \;&=\; \sum_{j=1}^{S} n_j\,|\Psi^{\mathrm{B}}(n_1, n_2, \ldots, n_S)\rangle \\ &=\; N\,|\Psi^{\mathrm{B}}(n_1, n_2, \ldots, n_S)\rangle\,.\end{aligned} \tag{9.34}$$

In using the boson creation and annihilation operators, in general, one needs to apply only the commutation propertes (9.23, 9.24) and the property $a_j|0\rangle = 0$. As long as one starts from a wave function with the proper boson symmetry, e.g., from the vacuum state $|0\rangle$ or states $|\Psi^{\rm B}(\mathcal{N})\rangle$, one does not need to worry ever about proper symmetries of wave functions, since they are induced through the algebra of $a_j^\dagger$ and a_j.

Exercise 9.2.1: The commutation relationships (9.23, 9.24) and $a_j|0\rangle = 0$ imply that the the properties (9.21, 9.22, 9.23) hold for the state defined through (9.29, 9.30). Prove this by induction, showing that the property holds for $n_j = 0$ and, in case it holds for n_j, it also holds for $n_j + 1$.

Exercise 9.2.2: Show that the boson operators $a_j^\dagger$ and a_j satisfy

$$\begin{aligned}[a_j, f(a_j^\dagger)] &= \frac{\partial f(a_j^\dagger)}{\partial a_j^\dagger} \\ [a_j^\dagger, f(a_j)] &= -\frac{\partial f(a_j)}{\partial a_j}\end{aligned}$$

where the operator function is assumed to have a convergent Taylor expansion

$$f(A) \;=\; \sum_{n=0}^{\infty} \frac{1}{n!}\, f^{(n)}(0)\, A^n \tag{9.35}$$

and where the derivative operation is defined through

$$\frac{\partial f(A)}{\partial A} \;=\; \sum_{n=1}^{\infty} \frac{1}{(n-1)!}\, f^{(n)}(0)\, A^{n-1}\,. \tag{9.36}$$

Creation and Annihilation Operators for Fermions

We want to consider now creation and annihilation operators for fermions, i.e., operators which can alter the occupancy of the wave function (9.17, 9.19) without affecting the fermion symmetry (9.9). Before proceeding in this respect we need to account for the following property of the fermion wave function which applies in the case $n_j, n_{j+1} = 1$, i.e., in case that the single particle states $|j\rangle$ and $|j+1\rangle$ are both occupied,

$$\langle \vec{x}_1, \vec{x}_2, \ldots \vec{x}_N | \Psi^{\rm F}(n_1, \ldots n_j, n_{j+1} \ldots n_S)\rangle \;= \tag{9.37}$$

$$-\langle \vec{x}_1, \vec{x}_2, \ldots \vec{x}_N | \Psi^{\rm F}(n_1, \ldots, n_{j+1}, n_j \ldots n_S)\rangle\,. \tag{9.38}$$

Obviously, the fermion wave function changes sign when one exchanges the order of the occupancy. To prove this property we notice that the l.h.s. of (9.38) corresponds to

$$\langle \vec{x}_1, \vec{x}_2, \ldots \vec{x}_N | \Psi^{\rm F}(n_1, \ldots n_j, n_{j+1} \ldots n_S)\rangle \;=$$

$$\frac{1}{\sqrt{N!}} \begin{vmatrix} \ldots & \phi_j(\vec{x}_1) & \phi_{j+1}(\vec{x}_1) & \ldots \\ \ldots & \phi_j(\vec{x}_2) & \phi_{j+1}(\vec{x}_2) & \ldots \\ \vdots & \vdots & \vdots & \vdots \\ \ldots & \phi_j(\vec{x}_N) & \phi_{j+1}(\vec{x}_N) & \ldots \end{vmatrix}\,. \tag{9.39}$$

The r.h.s. of (9.38) reads

$$-\langle \vec{x}_1, \vec{x}_2, \ldots \vec{x}_N | \Psi^{\mathrm{F}}(n_1, \ldots, n_{j+1}, n_j \ldots n_S)\rangle \;=$$

$$-\frac{1}{\sqrt{N!}} \begin{vmatrix} \ldots & \phi_{j+1}(\vec{x}_1) & \phi_j(\vec{x}_1) & \ldots \\ \ldots & \phi_{j+1}(\vec{x}_2) & \phi_j(\vec{x}_2) & \ldots \\ \vdots & \vdots & \vdots & \vdots \\ \ldots & \phi_{j+1}(\vec{x}_N) & \phi_j(\vec{x}_N) & \ldots \end{vmatrix} . \tag{9.40}$$

Because of the property of the determinant to change sign when two columns are interchanged, the expressions (9.39) and (9.40) are identical.
Obviously, it is necessary to specify for a fermion wave function the order in which the single–particle states $|\lambda_j\rangle$ are occupied. For this purpose one must adhere to a strict convention: the labelling of single–particle states by indices $j = 1, 2, \ldots$ must be chosen once and for all at the beginning of a calculation and these states must be occupied always in the order of increasing indices. A proper example is the two particle fermion wave function

$$\begin{aligned} &\langle \vec{x}_1, \vec{x}_2 | \Psi^{\mathrm{F}}(n_1 = 0, n_2 = 1, n_3 = 0, n_4 = 0, n_5 = 1, n_6 = n_7 = \ldots n_S = 0)\rangle \\ &= \frac{1}{\sqrt{2}} \begin{vmatrix} \phi_2(\vec{x}_1) & \phi_5(\vec{x}_1) \\ \phi_2(\vec{x}_2) & \phi_5(\vec{x}_2) \end{vmatrix} = \frac{1}{\sqrt{2}} \left[\phi_2(\vec{x}_1)\phi_5(\vec{x}_2) - \phi_5(\vec{x}_1)\phi_2(\vec{x}_2)\right] . \end{aligned} \tag{9.41}$$

Before we consider the definition of fermion creation and annihilation operators we need to take notice of the fact that one also needs to define at which position in the wave function, i.e., at which column of the Slater determinant (9.19), particles are being created or annihilated. *One adopts the convention that particles are created and annihilated by these operators always at the first position of the wave function,* i.e., at the first column of the Slater determinant (9.19). This requires one, however, in order to be consistent with this convention, that occupancies are always ordered according to a monotonous increase of the single–particle state index, to move the particle to the first position (to be annihilated there) or to move it from the first position to its canonical position (after it had been created at the first position). This change of position is connected with a possible sign change $(-1)^{q_j}$ where q_j is defined for a given $\mathcal{N} = (n_1, n_2, \ldots n_S)$ as follows:

$$q_j \;=\; \sum_{k=1}^{j-1} n_k \;, \tag{9.42}$$

i.e., q_j is the number of states $|k\rangle$ with $k < j$ which are occupied in a fermion wave function. We will be adopting below the notational convention that q'_j corresponds to occupancies $\mathcal{N}' = (n'_1, n'_2, \ldots n'_S)$.
We are now ready to define the annihilation operator c_j for fermions in the single–particle state $|j\rangle$ as follows

$$\begin{aligned} c_j \, |\Psi^{\mathrm{F}}(n_1, n_2, \ldots n_j \ldots n_S)\rangle \;= \\ n_j \, (-1)^{q_j} \, |\Psi^{\mathrm{F}}(\underbrace{n_1, n_2, n_3 \ldots n_{j-1}}_{q_j \text{ states occupied}}, n_j \to 0, n_{j+1} \ldots n_S)\rangle \end{aligned} \tag{9.43}$$

In case that the single particle state $|j\rangle$ is not occupied, i.e., $n_j = 0$, the rhs. vanishes.

The operator $c_j^\dagger$ adjoint to c_j exhibits the following property

$$\begin{aligned} c_j^\dagger \,|\Psi^{\rm F}(n_1, n_2, \ldots n_j \ldots n_S)\rangle \;=\; \\ (-1)^{q_j}(1 \,-\, n_j)\,|\Psi^{\rm F}(\underbrace{n_1, n_2, n_3 \ldots n_{j-1}}_{q_j \text{ states occupied}}, n_j \to 1, n_{j+1} \ldots n_S)\rangle \end{aligned} \tag{9.44}$$

The operators thus defined obey the commutation properties

$$[c_j, c_k]_+ \;=\; 0\,; \quad [c_j^\dagger, c_k^\dagger]_+ \;=\; 0 \tag{9.45}$$

$$[c_j, c_k^\dagger]_+ \;= \delta_{j\,k} \tag{9.46}$$

where we have used the definition of the so-called *anti-commutators* $[A, B]_+ \,=\, AB \,+\, BA$.
We first show that (9.44) follows from the definition (9.43). Let $\mathbb{C}_j$ be the matrix corresponding to the operator c_j in the basis of fermion states (9.17, 9.19). The elements of $\mathbb{C}_j$ are then (note that n_j only assumes values 0 or 1)

$$(\mathbb{C}_j)_{\mathcal{N}'\mathcal{N}} \;=\; \langle\Psi^{\rm F}(\mathcal{N}')|\; c_j\,|\Psi^{\rm F}(\mathcal{N})\rangle \;=\; n_j\,(-1)^{q_j}\,\delta_{n'_j\,n_j-1}\prod_{\substack{k=1\\k\neq j}}^{S}\delta_{n'_k n_k}\,. \tag{9.47}$$

Let $\mathbb{C}_j^\dagger$ be the matrix adjoint to $\mathbb{C}_j$. One obtains for its matrix elements

$$\begin{aligned} \left(\mathbb{C}_j^\dagger\right)_{\mathcal{N}\mathcal{N}'} \;&=\; \langle\Psi^{\rm F}(\mathcal{N}')|\; c_j^\dagger\,|\Psi^{\rm F}(\mathcal{N})\rangle \;=\; \overline{(\mathbb{C}_j)_{\mathcal{N}'\mathcal{N}}} \\ &=\; n'_j\,(-1)^{q'_j}\,\delta_{n_j\,n'_j-1}\prod_{\substack{k=1\\k\neq j}}^{S}\delta_{n_k n'_k} \\ &=\; (-1)^{q_j}(1-n_j)\,\delta_{n'_j\,n_j+1}\prod_{\substack{k=1\\k\neq j}}^{S}\delta_{n'_k n_k}\,. \end{aligned} \tag{9.48}$$

From this follows (9.44). We have used here the definition $q'_j \;=\; \sum_{k<j} n'_k$ along with $q'_j \;=\; q_j$, otherwise, the last factor on the rhs. of (9.48) vanishes.
We will prove now the anti-commutation property (9.46). The anti-commutation properties (9.45) follow in an analogous way and will not be derived here. We consider (9.46) first for the case $j = k$. It holds

$$\begin{aligned} c_j^\dagger c_j|\Psi^{\rm F}(\mathcal{N})\rangle \;&=\; c_j^\dagger(-1)^{q_j}n_j\,|\Psi^{\rm F}(n_1 \ldots n_j \to 0 \ldots n_S)\rangle \;= \\ (-1)^{q_j}(1-n_j+1)(-1)^{q_j}n_j\,|\Psi^{\rm F}(\mathcal{N})\rangle \;&=\; n_j\Psi^{\rm F}(\mathcal{N})\rangle \end{aligned} \tag{9.49}$$

The derivation involves realization of the fact that a non-zero result is obtained only in case $n_j = 1$. Similarly one obtains

$$\begin{aligned} c_j c_j^\dagger|\Psi^{\rm F}(\mathcal{N})\rangle \;&=\; c_j(-1)^{q_j}(1-n_j)\,|\Psi^{\rm F}(n_1 \ldots n_j \to 1 \ldots n_S)\rangle \;= \\ (-1)^{q_j}(n_j+1))(-1)^{q_j}(1-n_j)\,|\Psi^{\rm F}(\mathcal{N})\rangle \;&=\; (1-n_j)\,|\Psi^{\rm F}(\mathcal{N})\rangle \end{aligned} \tag{9.50}$$

(9.49) and (9.50) yield

$$(c_j c_j^\dagger \, + \, c_j^\dagger c_j \,) \, |\Psi^{\rm F}(\mathcal{N})\rangle \; = \; |\Psi^{\rm F}(\mathcal{N})\rangle \; . \tag{9.51}$$

For $j < k$ one obtains

$$\begin{aligned} c_j c_k^\dagger |\Psi^{\rm F}(\mathcal{N})\rangle \; &= \; (-1)^{q_j+q_k} n_j (1-n_k) |\Psi^{\rm F}(n_1 \ldots n_j = 0 \ldots n_k = 1 \ldots n_S)\rangle \\ c_k^\dagger c_j |\Psi^{\rm F}(\mathcal{N})\rangle \; &= \; (-1)^{q_j+q_k-1} n_j (1-n_k) |\Psi^{\rm F}(n_1 \ldots n_j = 0 \ldots n_k = 1 \ldots n_S)\rangle \end{aligned}$$

and, hence,

$$(c_j c_k^\dagger \, + \, c_k^\dagger c_j \,) \, |\Psi^{\rm F}(\mathcal{N})\rangle \; = \; 0 \; . \tag{9.52}$$

One can obtain similarly the same relationship in the case $j > k$. Since (9.51, 9.52) hold for any $|\Psi^{\rm F}(\mathcal{N})\rangle$ one can conclude (9.46).
Equation (9.49) shows that the operator

$$\hat{N}_j \; = \; c_j^\dagger c_j \tag{9.53}$$

is diagonal in the occupation number representation, i.e., in the basis $|\Psi^{\rm F}(\mathcal{N})\rangle$, with eigenvalues equal to the occupation numbers n_j

$$\hat{N}_j \, |\Psi^{\rm F}(\mathcal{N})\rangle \; = \; n_j \, |\Psi^{\rm F}(\mathcal{N})\rangle \tag{9.54}$$

$\hat{N}_j$, hence, is referred to as the *occupation number operator*. Correspondingly,

$$\hat{N} \; = \; \sum_{j=1}^{S} \hat{N}_j \tag{9.55}$$

is called the *particle number operator*.
It is an interesting exercise to demonstrate that $\hat{N}_j$ only has eigenvalues 0 or 1, for which purpose one needs to envoke only the algebraic (anti-commutation) properties (9.46). The stated property follows from the idempotence of $\hat{N}_j$ which can be derived as follows

$$\hat{N}_j^2 \; = \; c_j^\dagger c_j c_j^\dagger c_j \; = \; c_j^\dagger \, (1 - c_j^\dagger c_j) c_j \; = \; c_j^\dagger c_j \; - \; c_j^\dagger c_j^\dagger c_j c_j \; = \; c_j^\dagger c_j \; = \; \hat{N}_j \; . \tag{9.56}$$

Here we have made use of $c_j c_j = 0$ and $c_j^\dagger c_j^\dagger = 0$ which follows from $n_j \in \{0, 1\}$.
We finally note that the creation operators $c_j^\dagger$ allow one to construct any fermion state $|\Psi^{\rm F}(\mathcal{N})\rangle$ from the vacuum state $|0\rangle$ defined as above (see (9.29))

$$|\Psi^{\rm F}(\ldots n_{\lambda_1} \ldots n_{\lambda_2} \ldots n_{\lambda_N} \ldots)\rangle \; = \; c_{\lambda_1}^\dagger c_{\lambda_2}^\dagger \cdots c_{\lambda_N}^\dagger \, |0\rangle \; . \tag{9.57}$$

On the l.h.s. of this equation we meant to indicate only those N occupation numbers n_{λ_j} which are non-vanishing. On the r.h.s. the creation operators must operate in the proper order, i.e., an operator c_j must be on the left of an operator c_k for $j < k$.
In the following we will apply fermion operators $c_j^\dagger$ and c_j only to electrons which carry spin $\frac{1}{2}$. We will often separate the states $|j\rangle$, coordinates $\vec{x}$ and index j into a space part and a spin part, i.e.,

$$j \to j, \sigma \quad (\sigma = \pm \tfrac{1}{2}) \; ; \qquad \langle \vec{x} | j \rangle \; \to \; \phi_j(\vec{r}) \, |\tfrac{1}{2}\sigma\rangle \; . \tag{9.58}$$

9.3 One– and Two–Particle Operators

Definition

Operators acting on N particle systems of the type

$$\hat{F} \;=\; \sum_{j=1}^{N} \hat{f}(\vec{x}_j) \tag{9.59}$$

are called *single–particle operators*. The operators $\hat{f}(\vec{x})$ which constitute $\hat{F}$ are called the *associated generic single–particle operators*[1]. Operators of the type

$$\hat{G} \;=\; \frac{1}{2} \sum_{\substack{j,k=1 \\ j\neq k}}^{N} \hat{g}(\vec{x}_j, \vec{x}_k) \tag{9.60}$$

correspondingly are called *two–particle operators*. The operators $\hat{g}(\vec{x}, \vec{x}')$ which constitute $\hat{G}$ are called the *associated generic two–particle operators*. These operators had been introduced already in Eq. (9.6) above. The essential aspect of definition (9.59, 9.60) is that the sum over particles in (9.59) and over pairs of particles in (9.60) involves generic operators – acting on $\vec{x}_j$ and on $\vec{x}_j, \vec{x}_k$, respectively – which are all identical. An operator

$$\hat{R} \;=\; \sum_{j=1}^{N} \hat{r}_j(\vec{x}_j) \tag{9.61}$$

is not a one particle operator as long as the N operators $\hat{r}_j(\vec{x})$, $j = 1, 2, \ldots, N$ are not all identical. We seek to determine now how one–particle operators $\hat{F}$ and two–particle operators $\hat{G}$ act on many–boson and many–fermion states $|\Psi^{\mathrm{B}}(\mathcal{N})\rangle$ and $|\Psi^{\mathrm{F}}(\mathcal{N})\rangle$, respectively. The action of the operators is obviously described through the *many–particle state* matrix elements

$$\langle \Psi^{\mathrm{B,F}}(\mathcal{N}')|\hat{F}|\Psi^{\mathrm{B,F}}(\mathcal{N})\rangle \;; \qquad \langle \Psi^{\mathrm{B,F}}(\mathcal{N}')|\hat{G}|\Psi^{\mathrm{B,F}}(\mathcal{N})\rangle \;. \tag{9.62}$$

These matrix elements can be obtained by chosing the operators and many-particle states in the position representation, i.e., (9.59, 9.59) and (9.15, 9.17) and integrating over all particle space-spin coordinates $\vec{x}_1, \vec{x}_2, \ldots \vec{x}_N$. This procedure is most tedious and becomes essentially impossible in the general case that many-particle basis functions are linear combinations of wave functions of the type (9.15, 9.17).

Since the many–particle states are built-up from a basis $|r\rangle$, $r = 1, 2, \ldots S$ of single–particle spin-orbital states (see page 241) and the operators are specified through the associated generic operators $\hat{f}$ and $\hat{g}$, one expects that the matrix elements can actually be expressed in terms of matrix elements involving solely the generic operators and the single–particle states, namely,

$$\langle r|\hat{f}|s\rangle \;=\; \int d^3r\, \psi_r^*(\vec{r}, \sigma_r)\, \hat{f}(\vec{x})\, \psi_s(\vec{r}, \sigma_s) \tag{9.63}$$

[1]This expression has been specified for the purpose of these notes to distinguish $\hat{F}$ and $\hat{f}$ and is not common terminology. We adopt a similar expression to distinguish two–particle operators $\hat{G}$ and $\hat{g}$.

$$\langle r,s|\hat{g}|t,u\rangle \;=\; \int d^3r \int d^3r'\, \psi_r^*(\vec{r},\sigma_r)\psi_s^*(\vec{r}\,',\sigma_s)\, \hat{g}(\vec{x},\vec{x}\,')\, \psi_t(\vec{r},\sigma_t)\psi_u(\vec{r}\,',\sigma_u)\;. \tag{9.64}$$

We like to note an important symmetry of the matrix element $\langle r,s|\hat{g}|t,u\rangle$ which follows directly from the definition (9.64)

$$\langle r,s|\hat{g}|t,u\rangle \;=\; \langle s,r|\hat{g}|u,t\rangle\;. \tag{9.65}$$

This symmetry will be exploited repeatedly below. Notice, that the symmetry implies that one can switch *simultaneously* the pairs of indices r,s and t,u, *one cannot switch the indices of one pair individually.*

The aim of the present section is to express the matrix elements of single– and two–particle operators in a basis of many–particle states (9.62) in terms of the matrix elements of single–particle states (9.63, 9.64). We will show that the boson and fermion creation and annihilation operators serve this purpose.

Examples of One– and Two–Particle Operators

An example for a single–particle operator is the *kinetic energy operator*

$$\hat{T} \;=\; \sum_{j=1}^{N} \left(-\frac{\hbar^2}{2m}\right) \frac{\partial^2}{\partial \vec{r}_j^{\,2}}\;. \tag{9.66}$$

Here we adopt the notation for the Laplacian

$$\frac{\partial^2}{\partial \vec{r}^{\,2}} \;=\; \frac{\partial^2}{\partial x_1^2} \;+\; \frac{\partial^2}{\partial x_2^2} \;+\; \frac{\partial^2}{\partial x_3^2}\;. \tag{9.67}$$

In this case the generic single–particle operator, according to (9.59), is $\hat{t}(\vec{x}) \;=\; (-\hbar^2/2m)(\partial^2/\partial\vec{r}^{\,2})$. The matrix elements of this operator in a single–particle basis are

$$\langle r\sigma|\hat{t}|s\sigma'\rangle \;=\; \delta_{\sigma\sigma'} \int d^3r\, \phi_r^*(\vec{r}) \left(-\frac{\hbar^2}{2m}\right) \frac{\partial^2}{\partial \vec{r}^{\,2}}\, \phi_s(\vec{r})\;. \tag{9.68}$$

Another single–particle operator is the *one–particle density operator*

$$\hat{\rho}(\vec{\vec{r}}\,') \;=\; \sum_{j=1}^{N} \delta(\vec{\vec{r}} \;-\; \vec{r}_j)\;. \tag{9.69}$$

In this case the generic operator is $\hat{f}(\vec{x}) \;=\; \delta(\vec{\vec{r}} \;-\; \vec{r})$. The matrix elements of the generic operator in the one–particle basis are

$$\langle r\sigma|\hat{f}|s\sigma'\rangle \;=\; \delta_{\sigma\sigma'} \int d^3r\, \phi_r^*(\vec{r})\, \delta(\vec{\vec{r}} \;-\; \vec{r})\, \phi_s(\vec{r}) \;=\; \delta_{\sigma\sigma'}\; \phi_r^*(\vec{\vec{r}})\, \phi_s(\vec{\vec{r}})\, \delta_{\sigma\sigma'} \tag{9.70}$$

Both operators, i.e., (9.66) and (9.69), are spin-independent as is evident from the factor $\delta_{\sigma\sigma'}$.

An example for a two–particle operator is the *Coulomb repulsion operator*

$$\hat{V} \;=\; \frac{1}{2} \sum_{\substack{j,k=1 \\ j\neq k}}^{N} \frac{q^2}{|\vec{r}_j \;-\; \vec{r}_k|} \tag{9.71}$$

which is also spin-independent. In this case the generic operator is $\hat{v}(\vec{x},\vec{x}') = q^2/|\vec{r}_1 - \vec{r}_2|$. The matrix elements of the generic operator in terms of single–particle states are

$$\begin{aligned}\langle r,\sigma;\, s,\sigma'|\,\hat{t}\,|t,\sigma'';\, u,\sigma'''\rangle \;=\\ \int d^3r_1 \int d^3r_2\, \phi_r^*(\vec{r}_1)\phi_s^*(\vec{r}_2)\,\frac{q^2}{|\vec{r}_1-\vec{r}_2|}\,\phi_t(\vec{r}_1)\phi_u(\vec{r}_2)\,\delta_{\sigma\sigma''}\delta_{\sigma'\sigma'''}\end{aligned} \tag{9.72}$$

As a final example of one– and two–particle operators we consider operators of total spin. The operator

$$\vec{S} \;=\; \sum_{j=1}^{N} \vec{S}_j \tag{9.73}$$

is a one-particle operator. Here, $\vec{S}_j$ is the spin operator for particle j with components (in the spherical representation) $S_{j,+}$, $S_{j,-}$, $S_{j,3}$. The generic operator is given through the Pauli matrices, i.e., $\vec{\hat{s}} = \frac{1}{2}\vec{\sigma}$. The *non-vanishing* matrix elements of the spherical components of $\vec{s}$ are ($\alpha = +\frac{1}{2}, \beta = -\frac{1}{2}$)

$$\begin{aligned}\langle r,\alpha|\hat{s}_+|s,\beta\rangle &= \delta_{rs}\\ \langle r,\beta|\hat{s}_-|s,\alpha\rangle &= \delta_{rs}\\ \langle r,\alpha|\hat{s}_3|s,\alpha\rangle &= +\tfrac{1}{2}\,\delta_{rs}\\ \langle r,\beta|\hat{s}_3|s,\beta\rangle &= -\tfrac{1}{2}\,\delta_{rs}\end{aligned} \tag{9.74}$$

The operator

$$S^2 \;=\; \left(\sum_{j=1}^{N}\vec{S}_j\right)^2 \;=\; \sum_{j,k=1}^{N}\vec{S}_j\cdot\vec{S}_k \tag{9.75}$$

is a two–particle operator, however, not in the strict sense of our definition above since the restriction $j \neq k$ does not apply in the summation. However, one can obviously extract the term $\sum_j S_j^2$ to be left with a two–particle operator in the strict sense of our definition. The generic operator is

$$\hat{s}_{12} \;=\; \vec{S}_1\cdot\vec{S}_2 \;=\; \frac{1}{2}S_{1,+}S_{2,-} \;+\; \frac{1}{2}S_{1,-}S_{2,+} \;+\; S_{1,3}S_{2,3} \tag{9.76}$$

the single–particle state matrix elements of which are

$$\begin{aligned}\langle r,\sigma;\, s,\sigma'|\,\hat{s}_{12}\,|t,\sigma'';\, u,\sigma'''\rangle \;=\; \delta_{rt}\,\delta_{su}\Big(\tfrac{1}{2}\delta_{\frac{1}{2}\sigma}\delta_{-\frac{1}{2}\sigma''}\delta_{-\frac{1}{2}\sigma'}\delta_{\frac{1}{2}\sigma'''} \;+\\ +\;\tfrac{1}{2}\delta_{-\frac{1}{2}\sigma}\delta_{\frac{1}{2}\sigma''}\delta_{\frac{1}{2}\sigma'}\delta_{-\frac{1}{2}\sigma'''} \;+\; \tfrac{1}{4}\,\delta_{\sigma\sigma''}\delta_{\sigma'\sigma'''}\left(\,\delta_{\sigma\sigma'} \;-\; \delta_{-\sigma\sigma'}\,\right)\Big).\end{aligned} \tag{9.77}$$

The operators (9.74) and (9.77) are orbital-independent as evidenced by the factors δ_{rs} and $\delta_{rt}\delta_{su}$ in the respective formulas.

Definition in Terms of Creation and Annihilation Operators

In order to express many–particle state matrix elements (9.62) through single–particle state matrix elements (9.63, 9.64) one replaces the operators $\hat{F}$ in (9.59) and $\hat{G}$ in (9.60) as follows

$$\sum_{j=1}^{N} \hat{f}(\vec{x}_j) \;\rightarrow\; \begin{cases} \sum_{r,s=1}^{S} \langle r|\hat{f}|s\rangle \, a_r^\dagger a_s & \text{bosons} \\ \sum_{r,s=1}^{S} \langle r|\hat{f}|s\rangle \, c_r^\dagger c_s & \text{fermions} \end{cases} \tag{9.78}$$

$$\frac{1}{2} \sum_{\substack{j,k=1 \\ j\neq k}}^{N} \hat{g}(\vec{x}_j, \vec{x}_k) \;\rightarrow\; \begin{cases} \frac{1}{2} \sum_{r,s,t,u=1}^{S} \langle r,s|\hat{g}|t,u\rangle \, a_r^\dagger a_s^\dagger a_t a_u & \text{bosons} \\ \frac{1}{2} \sum_{r,s,t,u=1}^{S} \langle r,s|\hat{g}|t,u\rangle \, c_r^\dagger c_s^\dagger c_u c_t & \text{fermions} \end{cases} \tag{9.79}$$

It is crucial to notice in the expression for the fermion two–particle operator that the order of the annihilation operators in (9.79) is opposite to that in the matrix element $\langle r,s|\hat{g}|t,u\rangle$, namely $c_u c_t$, and **not** $c_t c_u$.

Expressions (9.78) and (9.79) have the following implication: The operators (9.78, 9.79) in the basis of many–particle states $|\Psi^{\mathrm{B,F}}(\mathcal{N})\rangle$ have the same values as the respective matrix elements (9.62). In order to determine these matrix elements one needs to evaluate first the matrix elements of the generic operators $\langle r|\hat{f}|s\rangle$ $\langle r,s|\hat{g}|t,u\rangle$ as defined in (9.63) and (9.64), respectively, and then in a second step the matrix elements

$$\langle \Psi^{\mathrm{B}}(\mathcal{N}')|a_r^\dagger a_s|\Psi^{\mathrm{B}}(\mathcal{N})\rangle \;, \quad \langle \Psi^{\mathrm{B}}(\mathcal{N}')|a_r^\dagger a_s^\dagger a_t a_u|\Psi^{\mathrm{B}}(\mathcal{N})\rangle \quad \text{bosons} \tag{9.80}$$

$$\langle \Psi^{\mathrm{F}}(\mathcal{N}')|c_r^\dagger c_s|\Psi^{\mathrm{F}}(\mathcal{N})\rangle \;, \quad \langle \Psi^{\mathrm{F}}(\mathcal{N}')|c_r^\dagger c_s^\dagger c_t c_u|\Psi^{\mathrm{F}}(\mathcal{N})\rangle \quad \text{fermions} \tag{9.81}$$

For the latter matrix elements simple rules can be derived from the algebraic properties of the boson and fermion operators (9.23, 9.24) and (9.45, 9.46), respectively. These rules will be provided below only for the case of fermions. To show that the resulting values of the matrix elements (9.62) are correct one needs to compare the result derived with conventional derivations of the matrix elements[2].

The Matrix Elements $\langle \Psi^{\mathrm{F}}(\mathcal{N}')|c_r^\dagger c_s|\Psi^{\mathrm{F}}(\mathcal{N})\rangle$

We consider first the case $r = s$. The matrix elements are then actually those of the number operator $c_r^\dagger c_r$ which is diagonal in the occupation number representation, i.e.,

$$\langle \Psi^{\mathrm{F}}(\mathcal{N}')|c_r^\dagger c_r|\Psi^{\mathrm{F}}(\mathcal{N})\rangle \;=\; n_r \prod_{r=1}^{S} \delta_{n'_r\, n_r} \;. \tag{9.82}$$

In case of $r \neq s$ one obtains

$$\begin{aligned} &\langle \Psi^{\mathrm{F}}(\mathcal{N}')|c_r^\dagger c_s|\Psi^{\mathrm{F}}(\mathcal{N})\rangle \;= \\ &\langle c_r\, \Psi^{\mathrm{F}}(\mathcal{N}')|c_s\, \Psi^{\mathrm{F}}(\mathcal{N})\rangle \;=\; n'_r(-1)^{q'_r}\, n_s(-1)^{q_s}\, \delta_{n'_r-1\, n_r} \delta_{n'_s\, n_s-1} \prod_{\substack{t=1 \\ t\neq r,s}} \delta_{n'_t\, n_t} \end{aligned} \tag{9.83}$$

[2]We refer the reader to Condon and Shortley's 'Theory of Atomic Spectra', pp.171 and pp.176 where the matrix elements of fermion states can be found.

From (9.82, 9.84) we can conclude that for *diagonal elements* holds

$$\langle \Psi^{\mathrm{F}}(\mathcal{N})|\hat{F}|\Psi^{\mathrm{F}}(\mathcal{N})\rangle \;=\; \sum_{r=1}^{S} n_r \,\langle r|\hat{f}|r\rangle \tag{9.84}$$

The *off-diagonal elements* with $n'_r = n_r$ except for $r = s,t$, in which case $n'_s = n_s + 1$ and $n'_t = n_t - 1$ holds, are

$$\langle \Psi^{\mathrm{F}}(\mathcal{N}')|\hat{F}|\Psi^{\mathrm{F}}(\mathcal{N})\rangle \;=\; (1-n_s)(-1)^{q'_s}\, n_t (-1)^{q_t}\,\langle s|\hat{f}|t\rangle \tag{9.85}$$

All other matrix elements vanish, i.e., those for which $\mathcal{N}'$ and $\mathcal{N}$ differ in more than two occupation numbers. Comparision with the results in Condon&Shortley (pp. 171) shows that the operator defined in (9.78) does yield the same results as the operator defined in (9.59) when the matrix elements are determined between Slater determinant wave functions.

The Matrix Elements $\langle \Psi^{\mathrm{F}}(\mathcal{N}')|c_r^\dagger c_s^\dagger c_t c_u|\Psi^{\mathrm{F}}(\mathcal{N})\rangle$

Before we determine these matrix elements we will investigate which possible combination of indices r,s,t,u need to be considered. Starting from the definition (9.79) and using $\left(c_r^\dagger\right)^2 = 0$, $c_r^2 = 0$ we can write

$$\hat{G} \;=\; \frac{1}{2} \sum_{\substack{r,s,t,u \\ r\neq s,\, t\neq u}}^{S} \langle rs|\hat{g}|tu\rangle c_r^\dagger c_s^\dagger c_u c_t \tag{9.86}$$

For the combination of indices r,s,t,u in this sum three possibilities exist

Case 0 two of the four indices are different;

Case 1 three of the four indices are different;

Case 2 all four indices are different.

Accordingly, we split the sum in (9.86) into three contributions, namely, $\hat{G} = \hat{G}_0 + \hat{G}_1 + \hat{G}_2$, where each contribution corresponds to one of the three cases mentioned.
The first contribution $\hat{G}_0$ is

$$\hat{G}_0 \;=\; \frac{1}{2} \sum_{\substack{r,s=1 \\ \text{a.i.d.}}}^{S} \langle r\, s|\hat{g}|r\, s\rangle c_r^\dagger c_s^\dagger c_s c_r \;+\; \frac{1}{2} \sum_{\substack{r,s=1 \\ \text{a.i.d.}}}^{S} \langle r\, s|\hat{g}|s\, r\rangle c_r^\dagger c_s^\dagger c_r c_s \;. \tag{9.87}$$

Here and in the following we denote by 'a.i.d.' (*all indices different*) that only tuples (r,s), (r,s,t), (r,s,t,u) are included in the summation for which all indices are different, i.e., for which holds $r \neq s$, $r \neq t$ etc. The anti-commutation property (9.45) yields together with the definition of the occupation number operator (9.53)

$$\hat{G}_0 \;=\; \frac{1}{2} \sum_{\substack{r,s=1 \\ \text{a.i.d.}}}^{S} \Big(\langle r\, s|\hat{g}|r\, s\rangle \;-\; \langle r\, s|\hat{g}|s\, r\rangle \Big)\, \hat{N}_r\, \hat{N}_s \;. \tag{9.88}$$

Obviously, this contribution to the two–particle operator is diagonal in the occupation number representation. As we will find, this part accounts for all diagonal contributions to $\hat{G}$.
The second contribution $\hat{G}_1$ is

$$\begin{aligned}\hat{G}_1 &= \tfrac{1}{2}\sum_{\substack{r,s,t=1\\ \text{a.i.d.}}}^{S}\langle r\,s|\hat{g}|r\,t\rangle c_r^\dagger c_s^\dagger c_t c_r \\ &+ \tfrac{1}{2}\sum_{\substack{r,s,t=1\\ \text{a.i.d.}}}^{S}\langle r\,s|\hat{g}|t\,s\rangle c_r^\dagger c_s^\dagger c_s c_t \\ &+ \tfrac{1}{2}\sum_{\substack{r,s,t=1\\ \text{a.i.d.}}}^{S}\langle r\,s|\hat{g}|s\,t\rangle c_r^\dagger c_s^\dagger c_t c_s \\ &+ \tfrac{1}{2}\sum_{\substack{r,s,t=1\\ \text{a.i.d.}}}^{S}\langle r\,s|\hat{g}|t\,r\rangle c_r^\dagger c_s^\dagger c_r c_t\ . \end{aligned} \tag{9.89}$$

Commutation of $c_t c_u$ according to (9.45), employing the symmetry property (9.65), relabelling summation indices and using $c_r^\dagger c_s^\dagger c_t c_r = c_s^\dagger c_r^\dagger c_r c_t = c_s^\dagger c_t \hat{N}_r$ yields

$$\hat{G}_1 = \sum_{\substack{r,s,t=1\\ \text{a.i.d.}}}^{S}\Big(\langle r\,s|\hat{g}|r\,t\rangle - \langle r\,s|\hat{g}|t\,r\rangle\Big) c_s^\dagger c_t \hat{N}_r\ . \tag{9.90}$$

Since $\hat{N}_r$ is diagonal, this contribution obviously behaves similarly to the off-diagonal part (9.84) of the one–particle operator $\hat{F}$.
A similar series of transformations allows one to express $\hat{G}_2$ which can be written

$$\begin{aligned}\hat{G}_2 &= \tfrac{1}{2}\sum_{\substack{r,s,t,u\\ r>s,t>u}}^{S}\langle r\,s|\hat{g}|t\,u\rangle c_r^\dagger c_s^\dagger c_u c_t \\ &+ \tfrac{1}{2}\sum_{\substack{r,s,t,u\\ r>s,t<u}}^{S}\langle r\,s|\hat{g}|t\,u\rangle c_r^\dagger c_s^\dagger c_u c_t \\ &+ \tfrac{1}{2}\sum_{\substack{r,s,t,u\\ r<s,t>u}}^{S}\langle r\,s|\hat{g}|t\,u\rangle c_r^\dagger c_s^\dagger c_u c_t \\ &+ \tfrac{1}{2}\sum_{\substack{r,s,t,u\\ r<s,t<u}}^{S}\langle r\,s|\hat{g}|t\,u\rangle c_r^\dagger c_s^\dagger c_u c_t \end{aligned} \tag{9.91}$$

as follows

$$\hat{G}_2 = \sum_{\substack{r,s,t,u\\ r>s,t>u}}^{S}\Big(\langle r\,s|\hat{g}|t\,u\rangle - \langle r\,s|\hat{g}|u\,t\rangle\Big) c_r^\dagger c_s^\dagger c_u c_t\ . \tag{9.92}$$

This contribution to $\hat{G}$ has non-vanishing matrix elements only when $\mathcal{N}'$ differs from $\mathcal{N}$ in *four* occupation numbers n_r, n_s, n_t, n_u. A similar contribution does not arise for $\hat{F}$.
We can state now the matrix elements $\langle\Psi^{\mathrm{F}}(\mathcal{N}')|\hat{G}|\Psi^{\mathrm{F}}(\mathcal{N})\rangle$. For this purpose we consider three cases which actually correspond to the three cases considered below Eq. (9.86).

Case 0 This case covers diagonal matrix elements, i.e., $\mathcal{N}' = \mathcal{N}$ or $n_r' = n_r$ for all r. Only $\hat{G}_0$, given in (9.88), contributes in this case. One obtains

$$\langle\Psi^{\mathrm{F}}(\mathcal{N})|\hat{G}|\Psi^{\mathrm{F}}(\mathcal{N})\rangle = \frac{1}{2}\sum_{r,s}^{S} n_r n_s\Big(\langle r,\,s|\hat{g}|r,\,s\rangle - \langle r,\,s|\hat{g}|s,\,r\rangle\Big)\ . \tag{9.93}$$

Case 1 This case covers off-diagonal elements with $n'_r = n_r$ except for $r = s, t$ for which holds $n'_s = n_s + 1$ and $n'_t = n_t - 1$. Only $\hat{G}_1$, given in (9.90), contributes in this case. One obtains

$$\begin{aligned} &\langle \Psi^{\mathrm{F}}(\mathcal{N}')|\hat{G}|\Psi^{\mathrm{F}}(\mathcal{N})\rangle \;= \\ &(-1)^{q'_s+q_t}(1-n_s)n_t \sum\nolimits_{\substack{r \\ r\neq s,t}} n_r \left(\langle r,s|\hat{g}|r,t\rangle \;-\; \langle r,s|\hat{g}|t,r\rangle \right) \end{aligned} \tag{9.94}$$

Case 2 This case covers off-diagonal elements with $n'_r = n_r$ except for $r = s, t, u, v$ for which holds $n'_s = n_s + 1$, $n'_t = n_t + 1$, $n'_u = n_u - 1$, $n'_v = n_v - 1$. Only $\hat{G}_2$, given in (9.92), contributes in this case. One obtains

$$\begin{aligned} &\langle \Psi^{\mathrm{F}}(\mathcal{N}')|\hat{G}|\Psi^{\mathrm{F}}(\mathcal{N})\rangle \;= \\ &\pm (-1)^{q'_s+q'_r+q_u+q_v}(1-n_s)(1-n_r)n_u n_v \left(\langle r,s|\hat{g}|u,v\rangle \;-\; \langle r,s|\hat{g}|v,u\rangle \right) \end{aligned} \tag{9.95}$$

In the latter formula the '+'-sign applies for $s < r$, $u < v$ or $s > r$, $u > v$, the '−'-sign applies for $s < r$, $u > v$ or $s > r$, $u < v$.

All matrix elements which are not covered by the three cases above vanish. In particular, for *non-vanishing* matrix elements $\langle \Psi^{\mathrm{F}}(\mathcal{N}')|\hat{G}|\Psi^{\mathrm{F}}(\mathcal{N})\rangle$ *at most four occupation numbers* n'_r and n_r *can differ*. Furthermore, *particle number is conserved*, i.e., it holds $\sum_{r=1}^{S} n'_r = \sum_{r=1}^{S} n_r$.

Exercise 9.3.1: Derive (9.90) from (9.89).
Exercise 9.3.2: Derive (9.92) from (9.91).
Exercise 9.3.3: Derive (9.95) from (9.92).
Exercise 9.3.4: Show that the matrix elements of $\hat{F}$ and of $\hat{G}$ conserve particle number.
Exercise 9.3.5: Derive the non-vanishing matrix elements (9.62) for bosons.

Spin Operators

In this section we will briefly consider the spin operators $\vec{\hat{S}}$ given in (9.73) and $\hat{S}^2$ given in (9.75) which are, respectively, one–particle and two–particle operators. The matrix elements of the corresponding generic operators are (9.74) and (9.76). We like to express these operators through fermion creation and annihilation operators. One obtains for the three components of $\vec{\hat{S}}$

$$\begin{aligned} \hat{S}_3 &= \frac{1}{2}\sum_{r=1}^{S} \left(\hat{N}_{r\alpha} \;-\; \hat{N}_{r\beta} \right) \\ \hat{S}_+ &= \sum_{r=1}^{S} c^{\dagger}_{r\alpha} c_{r\beta} \\ \hat{S}_- &= \sum_{r=1}^{S} c^{\dagger}_{r\beta} c_{r\alpha} \end{aligned} \tag{9.96}$$

For $\hat{S}^2$ one obtains

$$\hat{S}^2 \;=\; \frac{1}{2}\hat{N} \;+\; \frac{1}{4}\sum_{r,s=1}^{S}(\hat{N}_{r\alpha} \;-\; \hat{N}_{r\beta})(\hat{N}_{s\alpha} \;-\; \hat{N}_{s\beta}) \;-\; \sum_{r,s=1}^{S} c^{\dagger}_{r\alpha}c^{\dagger}_{s\beta}c_{r\beta}c_{s\alpha}\;. \tag{9.97}$$

The summation in (9.96, 9.97) is over the orbital states, the spin part of the single–particle states is represented by $\alpha = \frac{1}{2}$ and $\beta = -\frac{1}{2}$.

Exercise 9.3.6: Derive (9.96) and (9.97).

9.4 Independent-Particle Models

In the remaining part of this chapter we will apply the technique of operators of 2nd quantization to the description of many–fermion systems as they arise, e.g., in crystals, molecules, atoms and nuclei. In all these systems the *Hamiltonian* is a sum of one– and two–particle operators

$$H \;=\; \sum_{r=1}^{S}\langle r|\hat{t}|s\rangle c^{\dagger}_{r}c_{s} \;+\; \frac{1}{2}\sum_{r,s,t,u}^{S}\langle r,s|\hat{v}|u,v\rangle c^{\dagger}_{r}c^{\dagger}_{s}c_{u}c_{t}\;. \tag{9.98}$$

In many cases the Hamiltonian is spin-independent and the equivalent notation

$$H \;=\; \sum_{\substack{r,s=1\\ \sigma}}^{S}\langle r|\hat{t}|s\rangle c^{\dagger}_{r\sigma}c_{s\sigma} \;+\; \frac{1}{2}\sum_{\substack{r,s,t,u=1\\ \sigma,\sigma'}}^{S}\langle r,s|\hat{v}|t,u\rangle c^{\dagger}_{r\sigma}c^{\dagger}_{s\sigma'}c_{u\sigma'}c_{t\sigma} \tag{9.99}$$

will be used. In the latter case the indices r, s, t, u refer only to the orbital part of the single–particle basis.

If it were not for the two–particle contribution to the Hamiltonian, which represents the interactions between particles, the description of many–particle systems, e.g., evaluation of their stationary states and excitation energies, would be a simple exercise in linear algebra. Unfortunately, the two–particle operator makes the description of many–particle systems a very hard problem. The fortunate side of this is, however, that the two–particle operator representing interactions between particles induces a variety of interesting phenomena.

Actually, the study of problems governed by Hamiltonians of the type (9.98, 9.99) has preoccupied an important part of all intellectual efforts in Theoretical Physics during the past fifty years. In fact, it is one of the main intellectual achievements of Physics among the Sciences to have addressed systematically the study of systems composed of many strongly interacting components. The outcome of these studies is that even when interactions between components (particles) are simple, the concerted behaviour of interacting systems can deviate qualitatively from that of systems made up of independent components. Examples is the laser action, superconductivity, but also ordinary phenomena like freezing and evaporation. Often, the systems investigated involve a macroscopic number of components such that statistical mechanical approaches are envoked. The great generality of the concepts developed in the context of many–particle systems can be judged from the fact

that today these concepts are providing insight into the functioning of biological brains, another prototypical systems of many strongly interacting components, namely of neurons[3].
Before continuing one may finally point out that the material world as we know it and as it establishes the varieties of natural substances, ranging from nuclei to the molecules of living systems , ultimately depend in a crucial way on the fermion character of its constituent building blocks, electrons and nucleons. If it were'nt for the fermion nature, all material systems would condensate into states which would depend little on particle number. The *Aufbau principle*, according to which nuclei and atoms change their qualitative properties when they grow larger, would not exist. The electronic properties of atoms with different numbers of electrons would differ little, Chemistry essentially would know only one element and Life would not exist.

Exercise 9.4.1: Imagine that in a closed, water-filled jar all electrons of water turn their fermion nature into a boson nature. What might happen?

Independent-Particle Hamiltonian

A many–fermion system governed by an independent-particle Hamiltonian, i.e., a Hamiltonian without a two–particle operator contribution accounting for interactions among the particles, can be described in a rather straightforward way. *We will restrict our description in the following to systems composed of an even number (2N) of particles and to spin-independent interactions.* The Hamiltonian of such system is

$$H_o \;=\; \sum_{\substack{r,s=1\\ \sigma}}^{S} \langle r|\hat{t}|s\rangle c^{\dagger}_{r\sigma} c_{s\sigma} \tag{9.100}$$

We will denote the matrix elements of $\hat{t}$ as $\langle r|\hat{t}|s\rangle \;=\; t_{rs}$.

Transformation to a New Set of Creation and Annihilation Operators

Our aim is to represent the Hamiltonian (9.100) in a form in which the factors $c^{\dagger}_{r\sigma} c_{s\sigma}$ become occupation number operators. In such a representation H_o is diagonal and the stationary states can be stated readily. To alter the representation of H_o we apply a unitary transformation $\mathbf{U}$ to the single–particle states $|r\rangle$ to obtain a new basis of states $|m)$, i.e., $\{|m), m \;=\; 1, 2, \ldots S\}$, where

$$|m) \;=\; \sum_{r=1}^{S} U_{rm} |r\rangle \; . \tag{9.101}$$

The unitary property of U_{rn}

$$\mathbf{U}^{\dagger}\mathbf{U} \;=\; \mathbf{U}\,\mathbf{U}^{\dagger} \;=\; \mathbb{1} \tag{9.102}$$

or

$$\sum_{r=1}^{S} U *_{rm}\, U_{rn} \;=\; \delta_{mn} \; , \quad \sum_{m=1}^{S} U_{rm} U *_{sm} \;=\; \delta_{rs} \tag{9.103}$$

[3] An account of some of these efforts can be found in *Modelling Brain Function – The World of Attractor Neural Networks* by D.J. Amit (Cambridge University Press, New York, 1989)

allows one to express conversely $|r\rangle$ in terms of $|m)$

$$|r\rangle \;=\; \sum_{m=1}^{S} U^*_{rm}|m) \;. \tag{9.104}$$

One can, hence, expand

$$t_{rs} \;=\; \sum_{mn}^{S} \tilde{t}_{mn} U_{rm} U^*_{sn} \tag{9.105}$$

where

$$\tilde{t}_{mn} \;=\; (m|\hat{t}|n) \;=\; \sum_{r,s=1}^{S} t_{rs} U^*_{rm} U_{sn} \;=\; \left(U^\dagger \hat{t} U\right)_{mn} \tag{9.106}$$

which together with (9.100) yields

$$H_o \;=\; \tilde{t}_{mn} d^\dagger_{m\sigma} d_{n\sigma} \tag{9.107}$$

where

$$d^\dagger_{n\sigma} \;=\; \sum_{r=1}^{S} U_{rn} c^\dagger_{r\sigma} \;, \quad d_{n\sigma} \;=\; \sum_{r=1}^{S} U^*_{rn} c_{r\sigma} \tag{9.108}$$

These operators describe the creation and annihilation operators of fermions in states $|n)$ which are linear combinations (9.101) of states $|r\rangle$. The unitary property (9.103) allows one to express [c.f. (9.104)]

$$c^\dagger_{r\sigma} \;=\; \sum_{m=1}^{S} U^*_{rm} d^\dagger_{m\sigma} \;, \quad c_{s\sigma} \;=\; \sum_{n=1}^{S} U_{sn} d_{n\sigma} \;. \tag{9.109}$$

The operators (9.108) obey the same anti-commutation properties (9.45, 9.46) as $c^\dagger_{r\sigma}$ and $c_{r\sigma}$, namely,

$$[d_{m\sigma}, d_{n\sigma'}]_+ \;=\; 0 \;; \quad [d^\dagger_{m\sigma}, d^\dagger_{n\sigma'}]_+ \;=\; 0 \tag{9.110}$$

$$[d_{m\sigma}, d^\dagger_{n\sigma'}]_+ \;=\; \delta_{mn}\, \delta_{\sigma\sigma'} \tag{9.111}$$

and, accordingly, they are Fermion creation and annihilation operators. In a basis of states

$$\prod_{j=1}^{N} d^\dagger_{n_j\sigma_j} \; |0\rangle \tag{9.112}$$

, i.e., of N-Fermion states in which single particle states $|n_j)$ as defined in (9.101) are occupied, the operators $d^\dagger_{n\sigma}$ and $d_{n\sigma}$, behave exactly like the operators $c^\dagger_{r\sigma}$ and $c_{r\sigma}$ behave for states $|\Psi^F(\mathcal{N})\rangle$; for example, the expressions drived above for the matrix elements of one- and two-particle operators hold in an analogous way for $d^\dagger_{n\sigma}$ and $d_{n\sigma}$.

We demonstrate here the anti-commutation property (9.111), the anti-commutation properties (9.110) can be demonstrated similarly. The l.h.s. of (9.111) can be written, using (9.108) and (9.46),

$$[d_{m\sigma}, d^\dagger_{n\sigma'}]_+ \;=\; \sum_{r,s} U^*_{rm} U_{sn} \, [c_{r\sigma}, c^\dagger_{s\sigma'}]_+ \;=\; \delta_{\sigma\sigma'} \sum_{r} U^*_{rm} U_{rn} \;. \tag{9.113}$$

The unitarity property $\sum_r U^*_{rm} U_{rn} \;=\; \delta_{mn}$ yields immediately identity (**??**).

Diagonal Representation

The transformation (9.101) gives us the freedom to choose the matrix (9.106) diagonal, i.e.,

$$\tilde{t}_{mn} \;=\; \epsilon_n \, \delta_{mn} \; . \tag{9.114}$$

In this case H_o has the desired form

$$H_o \;=\; \sum_{n=1}^{S} \epsilon_n d^{\dagger}_{n\sigma} d_{n\sigma} \tag{9.115}$$

and involves solely occupation number operators $\hat{\tilde{N}}_{n\sigma} \;=\; d^{\dagger}_{n\sigma} d_{n\sigma}$. It is, hence, a simple matter to state many–particle states in the new representation.
Before proceeding we like to address the issue how the new representation, i.e., transformation matrices U_{rn} and energy values ϵ_n, is obtained. The condition (9.114) together with (9.106) is equivalent to

$$\sum_{s=1}^{S} t_{rs} U_{sn} \;=\; \epsilon_n U_{rn} \quad \forall r, r = 1, 2, \ldots S \tag{9.116}$$

which follows from (9.106) and the unitary property of U_{rn}. This equation poses the eigenvalue problem for the hermitian matrix (t_{rs}). The corresponding eigenvalues $\epsilon_n, n \;=\; 1, 2, \ldots S$ are real, the associated (properly normalized) eigenvectors V_n are the columns of (U_{rn}), i.e., $V_n^T \;=\; (U_{1n}, U_{2n}, \ldots U_{Sn})$.
Eigenstates of (9.115) can be stated immediately since any many–particle wave function in the occupation number representation is suitable. We apply for this purpose (9.57) to the case of $2N$ particles. The $2N$ fermion state

$$\prod_{j=1}^{2N} d^{\dagger}_{n_j \, \sigma_j} |0\rangle \quad \text{where } (n_j, \sigma_j) \;\neq\; (n_k, \sigma_k) \text{ for } j \neq k \tag{9.117}$$

are eigenstates of (9.115) with eigenvalues

$$E(n_1, \sigma_1;\, n_2, \sigma_2;\, \ldots n_{2N}, \sigma_{2N};\,) \;=\; \sum_{j=1}^{2N} \epsilon_{n_j} \; . \tag{9.118}$$

Ground State

In case of an ordering $\epsilon_m \;<\; \epsilon_n$ for $m < n$ the state of lowest energy, the so-called ground state, is

$$|\Phi_o\rangle \;=\; \prod_{j=1}^{N} d^{\dagger}_{j\,\alpha} d^{\dagger}_{j\,\beta} |0\rangle \; . \tag{9.119}$$

In this state the N lowest orbital eigenstates of t_{rs} are occupied each by an electron with spin $|\alpha\rangle \;=\; |\frac{1}{2}, \frac{1}{2}\rangle$ and $|\beta\rangle \;=\; |\frac{1}{2}, -\frac{1}{2}\rangle$. A *non-degenerate* orbital state which is occupied by two fermions, i.e., a fermion in a spin state $|\frac{1}{2},\, \frac{1}{2}\rangle$ as well as a fermion in a spin state $|\frac{1}{2},\, -\frac{1}{2}\rangle$ is called a *closed*

shell. The ground state has only closed shells. We will demonstrate now that this property endows the ground state with total spin zero.
In order to investigate the total spin of $|\Phi_o\rangle$ we apply to this state the total spin operator $\hat{S}^2$ as given by (9.97). The spin operator in the present representation of single-particle states is

$$\hat{S}^2 \;=\; \frac{1}{2}\hat{N} + \frac{1}{4}\sum_{m,n=1}^{S}(\hat{\bar{N}}_{m\alpha} - \hat{\bar{N}}_{m\beta})(\hat{\bar{N}}_{n\alpha} - \hat{\bar{N}}_{n\beta}) - \sum_{m,n=1}^{S} d^\dagger_{m\alpha}d^\dagger_{n\beta}d_{m\beta}d_{n\alpha} \tag{9.120}$$

where the occupation number operators refer to the single-particle states $|m)$. The action of the operator (9.120) is particularly simple for closed shells. One can conclude immediately that the second term in (9.97), namely, $\frac{1}{4}\sum_{m,n=1}^{S}(\hat{\bar{N}}_{m\alpha} - \hat{\bar{N}}_{m\beta})(\hat{\bar{N}}_{n\alpha} - \hat{\bar{N}}_{n\beta})$ does not give any contribution if either $|m)$ or $|n)$ are closed shells. Similarly, the third term $-\sum_{m,n=1}^{S} c^\dagger_{m\alpha}c^\dagger_{n\beta}c_{m\beta}c_{n\alpha}$ does not give a contribution if $m \neq n$ and either $|m)$ or $|n)$ are closed shells. In case of $m = n$ it gives a contribution '1' for each closed shell which cancels the contribution of the first term. Alltogether, one can conclude that $|\Phi_o\rangle$ is an eigenstate of $\hat{S}^2$ with eigenvalue zero, i.e., is a *singlet state*.

Excited States

We want to construct now excited states for the system governed by the independent-particle Hamiltonian (9.100). Obviously, the states with energy closest to the ground state are those in which a particle is promoted from the highest occupied state $|N)$ to the lowest unoccupied state $|N+1)$. There are four such states, namely,

$$\begin{aligned}
|\alpha,\alpha\rangle\rangle &= d^\dagger_{N+1\alpha}d_{N\alpha}|\Phi_o\rangle \\
|\alpha,\beta\rangle\rangle &= d^\dagger_{N+1\alpha}d_{N\beta}|\Phi_o\rangle \\
|\beta,\alpha\rangle\rangle &= d^\dagger_{N+1\beta}d_{N\alpha}|\Phi_o\rangle \\
|\beta,\beta\rangle\rangle &= d^\dagger_{N+1\beta}d_{N\beta}|\Phi_o\rangle
\end{aligned} \tag{9.121}$$

All four states have the same excitation energy, i.e., energy above the ground state, of $\Delta E = \epsilon_{N+1} - \epsilon_N$. However, the states differ in their spin character. The two states $|\alpha,\beta\rangle\rangle$ and $|\beta,\alpha\rangle\rangle$ are eigenstates of $\hat{S}^2$ given in (9.120), both with eigenvalues 2

$$\hat{S}^2|\alpha,\beta\rangle\rangle \;=\; 2\,|\alpha,\beta\rangle\rangle\,; \quad \hat{S}^2||\beta,\alpha\rangle\rangle \;=\; 2\,|\beta,\alpha\rangle\rangle\,. \tag{9.122}$$

This property can be derived as follows: The two states are obviously diagonal with respect to the following contribution to $\hat{S}^2$

$$\hat{\Sigma}_1 \;=\; \frac{1}{2}\hat{N} + \frac{1}{4}\sum_{m,n=1}^{N-1}(\hat{N}_{m\alpha} - \hat{N}_{m\beta})(\hat{N}_{n\alpha} - \hat{N}_{n\beta}) - \sum_{m,n=1}^{N-1} d^\dagger_{m\alpha}d^\dagger_{n\beta}d_{m\beta}d_{n\alpha} \tag{9.123}$$

which acts only on the $N-1$ closed shells (except for $\hat{N}$) of the two states and has an eigenvalue $N + 0 - (N-1) = 1$ in both cases. The remaining contributions to $\hat{S}^2$ which act on partially occupied orbitals $|N\rangle$ and $|N+1\rangle$ are

$$\hat{\Sigma}_2 \;=\; \frac{1}{4}\sum_{m,n=N}^{N+1}(\hat{N}_{m\alpha} - \hat{N}_{m\beta})(\hat{N}_{n\alpha} - \hat{N}_{n\beta}) \tag{9.124}$$

and

$$\hat{\Sigma}_3 \;=\; -\sum_{m,n=N}^{N+1} d^\dagger_{m\alpha} d^\dagger_{n\beta} d_{m\beta} d_{n\alpha} \; . \tag{9.125}$$

Contributions to $\hat{S}^2$ acting on states $|N+2\rangle, |N+3\rangle, \ldots$ do not need to be considered since they give vanishing contributions. The two states $|\alpha,\beta\rangle\rangle$, $|\beta,\alpha\rangle\rangle$ are, of course, also eigenstates of $\hat{\Sigma}_2$, both with eigenvalues 1. The action of $\hat{\Sigma}_3$ on the two states produces , for example, for $|\alpha,\beta\rangle\rangle$

$$\hat{\Sigma}_3|\alpha,\beta\rangle\rangle \;=\; -\sum_{m,n=N}^{N+1} d^\dagger_{m\alpha} d^\dagger_{n\beta} d_{m\beta} d_{n\alpha} d^\dagger_{N+1\alpha} d_{N\beta}|\Phi_o\rangle \;=\; 0 \tag{9.126}$$

which follows from the occurence of squares of fermion operators which, of course, vanish. The same result holds for $|\beta,\alpha\rangle\rangle$. One can, hence, conclude that (9.122) is correct.
An eigenvalue '2' of the spin operator $\hat{S}^2$ identifies the states $|\alpha,\beta\rangle\rangle$ and $|\beta,\alpha\rangle\rangle$ as triplet states. The remaining states $|\alpha,\alpha\rangle\rangle$ and $|\beta,\beta\rangle\rangle$ are not eigenstates of $\hat{S}^2$, however, the linear combinations $\sqrt{1/2}(|\alpha,\alpha\rangle\rangle \pm |\beta,\beta\rangle\rangle$ qualify as eigenstates, one with eigenvalue zero (singlet) and one with eigenvalue '2' (triplet).

Exercise 9.4.2: Construct four operators $\hat{O}_{\ell m}$

$$\hat{O}_{\ell m} \;=\; \sum_{\sigma,\sigma'} \gamma^{\ell m}_{\sigma,\sigma'} d^\dagger_{N+1,\sigma} d_{N\sigma'} \tag{9.127}$$

such that $\hat{O}_{\ell m}|\Phi_o\rangle$ are triplet and singlet states, i.e., apropriate eigenstates of the operators $\hat{S}^2$ and $\hat{S}_3$.

Example: $2N$ **Independent Electrons on a Ring**

In order to illustrate the procedure to obtain eigenstates of one-particle Hamiltonians (9.100) outlined above we consider a system of $2N$ electrons which move in a set of atomic orbitals $|r\rangle$, $r = 1, 2, \ldots 2N$ which are located on a ring. The system is assumed to posses a $2N$–fold symmetry axis and interactions connect only orbital states $|r\rangle \rightarrow |r \pm 1\rangle$. We identify the states $|2N+1\rangle = |1\rangle$ to avoid cumbersome summation limits, etc. The system is described by the Hamiltonian

$$\hat{H}_o \;=\; -t \sum_{\substack{n=1 \\ \sigma}}^{2N} \left(c^\dagger_{n+1\,\sigma} c_{n\,\sigma} \;+\; c^\dagger_{n\,\sigma} c_{n+1\,\sigma} \right) \; . \tag{9.128}$$

The cyclic property of the system is expressed through

$$c_{2N+1\,\sigma} \;=\; c_{1\,\sigma} \; ; \quad c^\dagger_{2N+1\,\sigma} \;=\; c^\dagger_{1\,\sigma} \tag{9.129}$$

The symmetry of the Hamiltonian suggests to choose a new representation defined through the operators ($r = 1, 2, \ldots 2N$)

$$d_{r\sigma} \;=\; \frac{1}{\sqrt{2N}} \sum_{n=1}^{2N} e^{-irn\pi/N} \, c_{n\sigma} \tag{9.130}$$

$$d^\dagger_{r\sigma} \;=\; \frac{1}{\sqrt{2N}} \sum_{n=1}^{2N} e^{irn\pi/N}\, c^\dagger_{n\sigma} \tag{9.131}$$

(9.130, 9.131) constitute a unitary transformation $\mathbf{U}$ defined through $U_{rn} \;=\; \frac{1}{\sqrt{2N}} e^{irn\pi/N}$. The unitarity property follows from

$$\sum_{r=1}^{2N} e^{irs\pi/N} \;=\; 0 \quad \text{for } 0 < s < 2N \;., \tag{9.132}$$

an identity which can be derived employing the well-known expression for a finite geometric series

$$\sum_{s=1}^{m} a^s \;=\; a\,\frac{1-a^m}{1-a} \;. \tag{9.133}$$

Application to the l.h.s. of (9.132) gives for $r \neq 0,\, 2N,\, 4N, \ldots$

$$e^{is\pi/N}\,\frac{1 \,-\, e^{2N\, is\pi/N}}{1 \,-\, e^{is\pi/N}} \tag{9.134}$$

which, indeed, vanishes for integer s. One can then express[cf. (9.109)]

$$c_{n\sigma} \;=\; \frac{1}{\sqrt{2N}} \sum_{r=1}^{2N} e^{irn\pi/N}\, d_{r\sigma} \tag{9.135}$$

$$c^\dagger_{n\sigma} \;=\; \frac{1}{\sqrt{2N}} \sum_{r=1}^{2N} e^{-irn\pi/N}\, d^\dagger_{r\sigma} \tag{9.136}$$

and, thereby, one obtains the new representation of $\hat{H}_o$

$$\hat{H}_o \;=\; -t \sum_{\substack{r,s=1\\ \sigma}}^{2N} \tilde{t}_{rs}\, d^\dagger_{r\sigma} d_{s\sigma} \tag{9.137}$$

where

$$\tilde{t}_{rs} \;=\; e^{-ir\pi/N}\,\frac{1}{2N} \sum_{n=1}^{2N} e^{i(s-r)n\pi/N} \;+\; e^{is\pi/N}\,\frac{1}{2N} \sum_{n=1}^{2N} e^{i(r-s)n\pi/N} \quad . \tag{9.138}$$

Using (9.132) one can conclude

$$\tilde{t}_{rs} \;=\; \left(e^{ir\pi/N} \;+\; e^{-ir\pi/N} \right) \delta_{rs} \tag{9.139}$$

from which follows

$$\hat{H}_o \;=\; \sum_{\substack{r=1\\ \sigma}}^{2N} \left(-2t\, \cos\frac{r\pi}{N} \right)\; d^\dagger_{r\sigma} d_{r\sigma} \;. \tag{9.140}$$

The Hamiltonian is now diagonal and the construction outlined above can be applied.

Exercise 9.4.3: Construct the ground state and the lowest energy excited state(s) for a system of $2N$ electrons moving in a linear chain of single-electron orbitals $|n\rangle$, $n = 1, 2, \ldots 2N$ with interactions between neighbouring $(n, n \pm 1)$ orbitals and Hamiltonian

$$\hat{H}_o \;=\; -\,t \sum_{\substack{n=1 \\ \sigma}}^{2N-1} \left(c^\dagger_{n+1\,\sigma} c_{n\,\sigma} \;+\; c^\dagger_{n\,\sigma} c_{n+1\,\sigma} \right) . \tag{9.141}$$

State the excitation energy ΔE.

9.5 Self-Consistent Field Theory

Observations of many-fermion systems show that properties of such systems often can be explained to a surprisingly good degree on the basis of the assumption that the fermions move like independent particles. The most famous example is the periodic table of the elements. The regularities of the elements can be rationalized in terms of electrons moving in hydrogen atom-type orbitals, i.e., in states which are constructed according to the 'Aufbauprinzip' in a manner which neglects the mutual repulsion between electrons. A similar principle accounts approximately for properties of nuclei as described by the nuclear shell model. The properties of molecules can be understood largely in terms of models which place electrons into so-called molecular orbitals in which the electrons move as if they were not interacting with other electrons. Many electronic properties of solids can be understood in a similar way, i.e., assuming that electrons move independently from each other in so-called bands, e.g., valence bands or conduction bands. Of course, properties of many-fermion properties deviant from such simple description is then of as much interest as independent fermion descriptions are useful.

Since independent-particle behaviour is so universal one may wonder why the presence of particle-particle interactions, i.e., the everpresent Coulomb repulsion between electrons, does not spoil it. The reason is mainly connected with the fact that most many-fermion systems are found in their ground state or removed from it through low energy excitations, and that independent-particle behaviour surfaces mainly because it applies well to the ground state. There are two reasons why this is so: first, if one places fermions into an independent-particle ground state of the type (9.119) then perturbations due to pair interactions according to the Pauli exclusion principle can only involve independent particle states not occupied in the ground state, i.e., those with energy *above* that of the (energetically) highest occupied single-particle state. Hence, the fermion nature restricts the possibility for perturbations on the system, in particular, if one can choose the single-particle states such that the residual perturbations are small. We will present in this and the following section a method which constructs such optimal single-fermion states. These states do not altogether neglect the pair interactions, rather they assume that each particle experiences the average interaction due to the fermions frozen into the ground state.

A second reason why independent-particle models can be successful is that the pair interaction for electrons is the slowly decaying Coulomb interaction. The slow decay of the interaction makes the motion of any one electron rather independent of the exact position of most other electrons, i.e., one expects that mean-field descriptions should be rather sufficient.

In this section we will consider systems of $2N$ fermions described by the spin-independent Hamiltonian

$$H \;=\; \sum_{\substack{r,s=1\\ \sigma}}^{S} \langle r|\hat{t}|s\rangle c^{\dagger}_{r\sigma}c_{s\sigma} \;+\; \frac{1}{2}\sum_{\substack{r,s,t,u=1\\ \sigma,\sigma'}}^{S} \langle r,s|\hat{v}|t,u\rangle c^{\dagger}_{r\sigma}c^{\dagger}_{s\sigma'}c_{u\sigma'}c_{t\sigma} \tag{9.142}$$

and derive a single-particle operator which approximates this Hamiltonian.
In our formulas below we will adopt strictly the convention that indices r, s, t, u refer to the initial basis of single-particle states $|r\rangle, r = 1, 2, \ldots, S$ on which (9.142) is based. We will adopt a second set of single-particle states $\{|\tilde{m}\rangle, m = 1, 2, \ldots S\}$ the elements of which will be labelled by indices m, n, p, q. The states $|\tilde{m}\rangle$ are connected with the states $|r\rangle$ through $(m = 1, 2, \ldots S)$

$$|\tilde{m}\rangle \;=\; \sum_{r=1}^{S} U_{rm}\,|r\rangle\;; \quad |r\rangle \;=\; \sum_{m=1}^{S} U^{*}_{rm}\,|\tilde{m}\rangle\;. \tag{9.143}$$

These S states are associated with the creation and annihilation operators

$$\begin{aligned} d^{\dagger}_{m\sigma} \;&=\; \sum_{r=1}^{S} U_{rm}c^{\dagger}_{r\sigma} \\ d_{m\sigma} \;&=\; \sum_{r=1}^{S} U^{*}_{rm}c_{r\sigma} \end{aligned} \tag{9.144}$$

defined as in (9.108). The states $|\tilde{m}\rangle$ serve to define a reference state of the system

$$|\Phi_o(\mathbf{U})\rangle \;=\; \prod_{j=1}^{N} d^{\dagger}_{j\,\alpha}d^{\dagger}_{j\,\beta}|0\rangle\;. \tag{9.145}$$

which will play a pivotal role: it is this the state in which the fermions are assumed to be moving and relative to which the mean interactions acting on individual fermions is determined.
$\mathbf{U} = (U_{rm})$ as defined in (9.143) with $r, m = 1, 2, \ldots S$ is a unitary matrix. If one restricts $m = 1, 2, \ldots N$, as in in (9.145), the corresponding $\mathbf{U} = (U_{rm})$ forms an $S \times N$ matrix and, naturally, is not unitary.

Mean-Field Potential

The mean field for the reference state $|\Phi_o(\mathbf{U})\rangle$ is defined in a straightforward way by averaging the two-particle contribution to (9.142) over this reference state to turn the contribution into an effective one-particle operator. This is done as follows:

$$\begin{aligned} &\hat{V}_{mf}\left(|\Phi_o(\mathbf{U})\rangle\right) \;= \\ &\frac{1}{2}\sum_{\substack{r,s,t,u=1\\ \sigma,\sigma'}}^{S} \langle r,s|\hat{v}|t,u\rangle\Big(\langle\langle c^{\dagger}_{r\sigma}c^{\dagger}_{s\sigma'}\rangle\rangle c_{u\sigma'}c_{t\sigma} \;+\; c^{\dagger}_{r\sigma}c^{\dagger}_{s\sigma'}\langle\langle c_{u\sigma'}c_{t\sigma}\rangle\rangle \\ &\qquad + c^{\dagger}_{r\sigma}\langle\langle c^{\dagger}_{s\sigma'}c_{u\sigma'}\rangle\rangle c_{t\sigma} \;+\; c^{\dagger}_{s\sigma'}\langle\langle c^{\dagger}_{r\sigma}c_{t\sigma}\rangle\rangle c_{u\sigma'} \\ &\qquad - c^{\dagger}_{r\sigma}\langle\langle c^{\dagger}_{s\sigma'}c_{t\sigma}\rangle\rangle c_{u\sigma'} \;-\; c^{\dagger}_{s\sigma'}\langle\langle c^{\dagger}_{r\sigma}c_{u\sigma'}\rangle\rangle c_{t\sigma} \quad \Big) \end{aligned} \tag{9.146}$$

Here $\langle\langle \quad \cdots \quad \rangle\rangle$ denotes the average

$$\langle\langle \mathcal{O} \rangle\rangle \;=\; \langle \Phi_o(\mathbf{U}) |\, \mathcal{O}\, | \Phi_o(\mathbf{U}) \rangle \;. \tag{9.147}$$

Since the reference state $|\Phi_o(\mathbf{U})\rangle$ is defined in terms of the single-particle states $|\tilde{m}\rangle$ it is preferable to switch the representation of (9.146) accordingly. Since the averages affect only two creation-annihilation operators actually a mixed representation is most suitable

$$\begin{aligned}
\hat{V}_{mf}\left(|\Phi_o(\mathbf{U})\rangle\right) \;&=\; \frac{1}{2} \sum_{\substack{m,n,t,u=1 \\ \sigma,\sigma'}}^{S} \langle \tilde{m}, \tilde{n} | \hat{v} | t, u \rangle \langle\langle d^\dagger_{m\sigma} d^\dagger_{n\sigma'} \rangle\rangle c_{u\sigma'} c_{t\sigma} \\
&+ \frac{1}{2} \sum_{\substack{r,s,m,n=1 \\ \sigma,\sigma'}}^{S} \langle r, s | \hat{v} | \tilde{m}, \tilde{n} \rangle c^\dagger_{r\sigma} c^\dagger_{s\sigma'} \langle\langle d_{m\sigma'} d_{n\sigma} \rangle\rangle \\
&+ \frac{1}{2} \sum_{\substack{r,m,n,t=1 \\ \sigma,\sigma'}}^{S} \langle r, \tilde{m} | \hat{v} | t, \tilde{n} \rangle c^\dagger_{r\sigma} \langle\langle d^\dagger_{m\sigma'} d_{n\sigma'} \rangle\rangle c_{t\sigma} \\
&+ \frac{1}{2} \sum_{\substack{m,s,n,u=1 \\ \sigma,\sigma'}}^{S} \langle \tilde{m}, s | \hat{v} | \tilde{n}, u \rangle c^\dagger_{s\sigma'} \langle\langle d^\dagger_{m\sigma} d_{n\sigma} \rangle\rangle c_{u\sigma'} \\
&- \frac{1}{2} \sum_{\substack{r,m,n,u=1 \\ \sigma,\sigma'}}^{S} \langle r, \tilde{m} | \hat{v} | \tilde{n}, u \rangle c^\dagger_{r\sigma} \langle\langle d^\dagger_{m\sigma'} d_{n\sigma} \rangle\rangle c_{u\sigma'} \\
&- \frac{1}{2} \sum_{\substack{m,s,t,n=1 \\ \sigma,\sigma'}}^{S} \langle \tilde{m}, s | \hat{v} | t, \tilde{n} \rangle c^\dagger_{s\sigma'} \langle\langle d^\dagger_{m\sigma} d_{n\sigma'} \rangle\rangle c_{t\sigma}
\end{aligned} \tag{9.148}$$

where, for example,

$$\langle r, \tilde{m} |\, \hat{v}\, | t, \tilde{n} \rangle \;=\; \sum_{s,u}^{S} \langle r, s |\, \hat{v}\, t, u \rangle\, U^*_{sm}\, U_{un} \;. \tag{9.149}$$

The remaining matrix elements appearing in (9.148) are obtained in a similar way by transforming only two of the four single-particle states into the new representation.
For the averages $\langle\langle \quad \cdots \quad \rangle\rangle$ in (9.148) one obtains, by means of the rules (9.84),

$$\begin{aligned}
\langle\langle d^\dagger_{m\sigma} d^\dagger_{n\sigma'} \rangle\rangle \;&=\; \langle \Phi_o(\mathbf{U}) | d^\dagger_{m\sigma} d^\dagger_{n\sigma'} | \Phi_o(\mathbf{U}) \rangle \;=\; 0 \\
\langle\langle d_{m\sigma} d_{n\sigma'} \rangle\rangle \;&=\; \langle \Phi_o(\mathbf{U}) | d_{m\sigma} d_{n\sigma'} | \Phi_o(\mathbf{U}) \rangle \;=\; 0 \\
\langle\langle d^\dagger_{m\sigma} d_{n\sigma'} \rangle\rangle \;&=\; \langle \Phi_o(\mathbf{U}) | d^\dagger_{m\sigma} d_{n\sigma'} | \Phi_o(\mathbf{U}) \rangle \\
&=\; \begin{cases} \delta_{mn}\delta_{\sigma\sigma'} & m \le N \\ 0 & m > N \end{cases}
\end{aligned} \tag{9.150}$$

and, hence,

$$\hat{V}_{mf}\left(|\Phi_o(\mathbf{U})\rangle\right) \;=\; \frac{1}{2} \sum_{m \le N} \sum_{\substack{r,t=1 \\ \sigma,\sigma'}}^{S} \langle r, \tilde{m} | \hat{v} | t, \tilde{n} \rangle c^\dagger_{r\sigma} c_{t\sigma}\; \delta_{\sigma'\sigma'}$$

$$+\frac{1}{2}\sum_{m\leq N}\sum_{\substack{s,u=1\\ \sigma,\sigma'}}^{S}\langle\tilde{m},s|\hat{v}|\tilde{n},u\rangle c^{\dagger}_{s\sigma'}c_{u\sigma'}\;\delta_{\sigma\sigma}$$

$$-\frac{1}{2}\sum_{m\leq N}\sum_{\substack{r,u=1\\ \sigma}}^{S}\langle r,\tilde{m}|\hat{v}|\tilde{n},u\rangle c^{\dagger}_{r\sigma}c_{u\sigma}$$

$$-\frac{1}{2}\sum_{m\leq N}\sum_{\substack{s,t=1\\ \sigma}}^{S}\langle\tilde{m},s|\hat{v}|t,\tilde{n}\rangle c^{\dagger}_{s\sigma}c_{t\sigma}\;. \tag{9.151}$$

Carrying out the sum $\sum_\sigma \delta_{\sigma\sigma} = 2$ leads to a factor 2 for the first two terms. Exploiting the symmetry $\langle r,s|\,\hat{v}\,|t,u\rangle = \langle s,r|\,\hat{v}\,|u,t\rangle$ and renaming dummy summation indices one can demonstrate that term one and term two as well as term three and four are identical and one obtains

$$\begin{aligned}\hat{V}_{mf}\left(|\Phi_o(\mathbf{U})\rangle\right) &= 2\sum_{\substack{r,m,t=1\\ \sigma,\sigma'}}^{S}\langle r,\tilde{m}|\hat{v}|t,\tilde{n}\rangle c^{\dagger}_{r\sigma}c_{t\sigma}\;\delta_{\sigma'\sigma'}\\ &\quad-\sum_{\substack{r,m,u=1\\ \sigma,\sigma'}}^{S}\langle r,\tilde{m}|\hat{v}|\tilde{n},u\rangle c^{\dagger}_{r\sigma}c_{u\sigma'}\end{aligned} \tag{9.152}$$

By means of expression (9.149) for the matrix elements in the mixed representation one can state the right hand side explicitly in terms of U_{rm}. The one-particle operator (9.152) can then be written in the form

$$\hat{V}_{mf}\left(|\Phi_o(\mathbf{U})\rangle\right) = \sum_{\substack{r,s\\ \sigma}}\langle r|\,\hat{v}_{mf}\left(|\Phi_o(\mathbf{U})\rangle\right)\,|s\rangle\,c^{\dagger}_{r\sigma}c_{s\sigma} \tag{9.153}$$

where

$$\langle r|\,\hat{v}_{mf}\left(|\Phi_o(\mathbf{U})\rangle\right)\,|s\rangle = \sum_{t,u}\left(\,2\,\langle r,t|\,\hat{v}\,|s,u\rangle - \langle r,t|\,\hat{v}\,|u,s\rangle\,\right)\left(\sum_{m\leq N}U^{*}_{tm}U_{um}\right) \tag{9.154}$$

The mean field approximation replaces then the Hamiltonian (9.142) by $\hat{H}_{mf}(\mathbf{U})$ defined through

$$\hat{H}_{mf}(\mathbf{U}) = \sum_{\substack{r,s=1\\ \sigma}}^{S}\left(\,\langle r|\hat{t}|s\rangle + \langle r|\,\hat{v}_{mf}\left(|\Phi_o(\mathbf{U})\rangle\right)\,|s\rangle\,\right)c^{\dagger}_{r\sigma}c_{s\sigma} \tag{9.155}$$

which is a function of the $S\times N$–matrix

$$\mathbf{U} = (U_{rm})_{r=1,2,\ldots S;\,m=1,2,\ldots N} \tag{9.156}$$

which defines the reference state $|\Phi_o(\mathbf{U})\rangle$. The mean field approximation, as introduced here, leaves open the question how the reference state should be chosen. At this point this state is completely arbitrary. We will address now a proper choice of the reference state.

9.6 Self-Consistent Field Algorithm

The Self-Consistent Field (SCF) approximation, often also referred to as the Hartree-Fock approximation, is based on the mean field approach and provides an algorithm to construct a reference state $|\Phi_o(\mathbf{U})\rangle$ for a $2N$ fermion system described by a spin-independent Hamiltonian (9.142). The procedure determines the reference state as the ground state of the independent-particle Hamiltonian (9.155) following the method outlined in Section 9.4. This procedure, however, can achieve this goal only iteratively, assuming in an intial step (1) a properly chosen reference state, determining in a step (2) the corresponding mean field Hamiltonian (9.155), and obtaining in a step (3) the ground state of this one-particle Hamiltonian according to the construction in Section 9.4 and defining this as the new reference state; steps (2) and (3) are being repeated M times until the procedure converges, i.e., the reference state resulting from step $2M+1$ (within numerical error) is equal to the reference state resulting from step $2M-1$. The state thus determined is referred to as the self-consistent independent-particle ground state, the independent-particle nature stemming from the fact that the functional form of the ground state, i.e., (9.119, 9.145), is exact only for an independent-particle Hamiltonian. We will argue below that the self-consistent field ground state, under conditions which often are realized, is the lowest energy independent-particle ground state one can construct.
Let us state now the construction of the self-consistent field ground state in more detail.

SCF-Algorithm, Step 1: Choosing an Initial Reference State

One defines the Hamiltonian

$$\hat{H}^{(1)}_{mf} \;=\; \sum_{\substack{r,s=1\\ \sigma}}^{S} \langle r|\hat{t}|s\rangle c^{\dagger}_{r\sigma} c_{s\sigma} \tag{9.157}$$

and determines the associated diagonal representation defined through

$$\sum_{s=1}^{S} \langle r|\hat{t}|s\rangle U^{(1)}_{sm} \;=\; \epsilon^{(1)}_m \, U^{(1)}_{rm} \;, \quad r = 1,2,\ldots S \;, \quad m = 1,2,\ldots S \;. \tag{9.158}$$

where the labels m are ordered to obey the condition

$$\epsilon^{(1)}_m \;< \epsilon^{(1)}_n \quad \text{for } m \;<\; n \;. \tag{9.159}$$

One defines the reference state $|\Phi_o(\mathbf{U}^{(1)})\rangle$ using the definition (9.145), i.e. $\mathbf{U}^{(1)}$ is the $S\times N$–matrix of the first N column vectors of $U^{(1)}_{rm},\ r = 1,2,\ldots S \;, \quad m = 1,2,\ldots S$.

SCF-Algorithm, Step $2M$, $M = 1,2,\ldots$: Determine Mean Field Hamiltonian

In this step the mean field Hamiltonian

$$\begin{aligned} \hat{H}^{(M)}_{mf} \;&=\; \sum_{\substack{r,s=1\\ \sigma}}^{S} \Big(\langle r|\hat{t}|s\rangle \;+\; \langle r|\, \hat{v}_{mf} \left(|\Phi_o(\mathbf{U}^{(M)})\rangle \right) |s\rangle \Big) \, c^{\dagger}_{r\sigma} c_{s\sigma} \\ &=\; \sum_{\substack{r,s=1\\ \sigma}}^{S} \langle r|\hat{h}^{(M)}_{mf}|s\rangle \, c^{\dagger}_{r\sigma} c_{s\sigma} \end{aligned} \tag{9.160}$$

is evaluated, i.e., the $S \times S$–matrix

$$\langle r|h_{mf}^{(M)}|s\rangle \;=\; \langle r|\hat{t}|s\rangle \;+\; \sum_{t,u=1}^{S} \left(2\langle r,t|\,\hat{v}\,|s,u\rangle \;-\; \langle r,t|\,\hat{v}\,|u,s\rangle \right) \left(\sum_{m=1}^{N} U_{tm}^{*(M)} U_{um}^{(M)} \right) \tag{9.161}$$

is calculated.

SCF-Algorithm, Step $2M+1$, $M = 1, 2, \ldots$: Diagonalize Mean Field Hamiltonian

In this step the eigenvalue problem

$$\sum_{s=1}^{S} \langle r|h_{mf}^{(M)}|s\rangle U_{sm}^{(M+1)} \;=\; \epsilon_m^{(M+1)}\, U_{rm}^{(M+1)}\,, \quad r = 1, 2, \ldots S\,,\; m = 1, 2, \ldots S \tag{9.162}$$

is solved adopting an ordering of the labels m which obeys the condition

$$\epsilon_m^{(M+1)} < \epsilon_n^{(M+1)} \quad \text{for } m \,<\, n\,. \tag{9.163}$$

The result yields the reference state $|\Phi_o(\mathbf{U}^{(M+1)})\rangle$ using the definition (9.145), i.e. $\mathbf{U}^{(M+1)}$ is the $S \times N$–matrix of the first N column vectors of $U_{rm}^{(M+1)}$, $r = 1, 2, \ldots S$, $m = 1, 2, \ldots S$.

SCF-Algorithm: Continuation and Convergence Condition

When step $2M+1$ is completed, step $2M+2$ is carried out, etc. The procedure is continued until it is detected that the one-particle density differences $|\sum_{m=1}^{N} U_{rm}^{*(M+1)} U_{um}^{(M+1)} \;-\; \sum_{m=1}^{N} U_{rm}^{*(M)} U_{um}^{(M)}|$ for $r, s \,=\, 1, 2 \ldots S$ do not exceed a preset threshold, indicating that the state converged.

Exercise 9.6.1: Let $\hat{P}_{rs}$ be the one-particle operator with the generic operator $\hat{p}_{rs} \,=\, |r\rangle\langle s|$, i.e.,

$$\langle t, \sigma|\hat{p}_{rs}|u, \sigma'\rangle \;=\; \delta_{tr}\delta_{su}\delta_{\sigma\sigma'}\,. \tag{9.164}$$

(a) Determine the expectation value of $\hat{P}_{rs}$ for the state as defined in (9.145).
(b) How can the expectation values of the operator $\hat{P}_{rr}$ be interpreted.
(c) Show that any spin-independent one-particle operator

$$\hat{F} \;=\; \sum_{\substack{r,s=1 \\ \sigma}}^{S} \langle r|\hat{f}|s\rangle\, c_{r\sigma}^{\dagger} c_{s\sigma} \tag{9.165}$$

can be written

$$\hat{F} \;=\; \sum_{r,s=1}^{S} \langle r|\hat{f}|s\rangle\, \hat{P}_{rs} \tag{9.166}$$

(d) Define a similar two-particle operator $\hat{P}_{rstu}$.
(e) Consider the creation operators

$$g_{q\sigma}^{\dagger} \;=\; \sum_{m=1}^{N} V_{mq}\, d_{m\sigma}^{\dagger} \tag{9.167}$$

which are connected with $d^\dagger_{m\sigma}$ through a unitary $N \times N$–matrix (V_{qm}). Note that this is **not** an $S \times S$–matrix! Show that the reference state defined through

$$\prod_{q=1}^{N} g^\dagger_{q\alpha} g^\dagger_{q\beta} |0\rangle \tag{9.168}$$

has the same expectation values for $\hat{P}_{rs}$ and $\hat{P}_{rstu}$ as the reference state (9.145).
(f) Can one distinguish the states (9.145) and (9.168) through a physical observation?
Exercise 9.6.2: Determine the SCF ground state and its energy expectation value for a system of two particles described through the Hamiltonian ($S = 2$)

$$\begin{aligned} H \;&=\; \sum_\sigma \Big(\epsilon\,(\, c^\dagger_{1\sigma} c_{1\sigma} \,+\, c^\dagger_{2\sigma} c_{2\sigma}\,) \;-\; t\,(\, c^\dagger_{1\sigma} c_{2\sigma} \,+\, c^\dagger_{2\sigma} c_{1\sigma})\Big) \\ &+ 2v\,(\, c^\dagger_{1\alpha} c_{1\alpha} c^\dagger_{1\beta} c_{1\beta} \,+\, c^\dagger_{2\alpha} c_{2\alpha} c^\dagger_{2\beta} c_{2\beta}\,) \,+ v \sum_{\sigma,\sigma'} (\, c^\dagger_{1\sigma} c_{1\sigma} c^\dagger_{2\sigma'} c_{2\sigma'} \\ &+ c^\dagger_{2\sigma} c_{2\sigma} c^\dagger_{1\sigma'} c_{1\sigma'}\,)\,. \end{aligned} \tag{9.169}$$

9.7 Properties of the SCF Ground State

We want to investigate now the properties of the SCF ground state. We begin by summarizing the result of the SCF algorithm.
We denote the representation which results from the SCF algorithm after its convergence as follows:

$$|\tilde{m}\rangle \;=\; \sum_{r=1}^{S} U^{(SCF)}_{rm}\, |r\rangle\,. \tag{9.170}$$

Here $|r\rangle$, $r = 1, 2, \ldots S$ denotes the initial single-particle states and $\left(U^{(SCF)}_{rm}\right)$ is the unitary $S \times S$–transformation matrix obtained in (9.162,9.163). For this representation holds

$$\langle\tilde{m}|\,\hat{t}\,|\tilde{n}\rangle \;+\; \sum_{m'}^{N} \big(\, 2\langle\tilde{m}\tilde{m}'|\,\hat{v}\,|\tilde{n}\tilde{m}'\rangle \;-\; \langle\tilde{m}\tilde{m}'|\,\hat{v}\,|\tilde{m}'\tilde{n}\rangle\,\big) \;=\; \epsilon^{(SCF)}_m\, \delta_{mn} \tag{9.171}$$

where the convention (9.163), i.e.,

$$\epsilon^{(SCF)}_m \,< \epsilon^{(SCF)}_n \quad \text{for } m \,<\, n \tag{9.172}$$

is obeyed. Equation (9.171) implies that for the mean field Hamiltonian holds

$$\hat{H}_{mf}(\mathbf{U}^{(SCF)}) \;=\; \sum_{\substack{m=1 \\ \sigma}}^{S} \epsilon^{(SCF)}_m\, d^\dagger_{m\sigma} d_{m\sigma}\,, \tag{9.173}$$

i.e., the mean-field Hamiltonian is diagonal in the SCF representation. Finally, the SCF ground state is

$$\|SCF\rangle \;=\; \prod_{m=1}^{N} d^{\dagger}_{m\alpha} d^{\dagger}_{m\beta} |0\rangle \; . \tag{9.174}$$

The first property we consider is the total spin character of $\|SCF\rangle$. Since this state is composed only of closed shells one can conclude following Section 9.4 that $\|SCF\rangle$ is a total singlet state.
We like to determine next the energy expectation value of $\|SCF\rangle$. Within the mean field approximation as described by (9.173) holds

$$\langle SCF\|\hat{H}_{mf}(\mathbf{U}^{(SCF)})\|SCF\rangle \;=\; 2\sum_{r=1}^{N} \epsilon_m^{(SCF)} \; . \tag{9.175}$$

However, one can also determine the expectation value of $\|SCF\rangle$ for the Hamiltonian (9.142), i.e., $\langle SCF\|\hat{H}\|SCF\rangle$. Employing expressions (9.84) and (9.93) for the expectation values (diagonal elements) of the one-particle and two-particle part of the Hamiltonian (9.142), exploiting the spin-independence of (9.142), one obtains

$$\begin{aligned} &\langle SCF\|\hat{H}\|SCF\rangle \;= \\ &2\sum_{m=1}^{N} \langle \tilde{m}|\hat{t}|\tilde{m}\rangle \;+\; \sum_{m,m'=1}^{N} \left(2\langle \tilde{m}\tilde{m}'|\,\hat{v}\,|\tilde{m}\tilde{m}'\rangle \;-\; \langle \tilde{m}\tilde{m}'|\,\hat{v}\,|\tilde{m}'\tilde{m}\rangle \right) \; . \end{aligned} \tag{9.176}$$

Comparision with (9.171) yields

$$\begin{aligned} &\langle SCF\|\hat{H}\|SCF\rangle \;= \\ &2\sum_{m=1}^{N} \epsilon_m^{(SCF)} \;-\; \sum_{m,m'=1}^{N} \left(2\langle \tilde{m}\tilde{m}'|\,\hat{v}\,|\tilde{m}\tilde{m}'\rangle \;-\; \langle \tilde{m}\tilde{m}'|\,\hat{v}\,|\tilde{m}'\tilde{m}\rangle \right) \; . \end{aligned} \tag{9.177}$$

The second term originates from the fact that in the mean field approximation one assumes that each of the fermions is subject to an average pair interaction involving a $2N$-particle reference state. Since the system under consideration has only $2N$ particles altogether, the mean field potential $\hat{V}_{mf}\left(|\Phi_o(\mathbf{U})\rangle\right)$ defined in (9.146) counts a particle twice, once as a member of the ground state, and once as a probe particle being subject to the mean pair interaction. This over-counting leads to the correction term in (9.177).

Exercise 9.7.1: Reformulate (9.171) in terms of matrix elements $\langle r|\hat{t}|s\rangle$, $\langle r,s|\hat{v}|t,u\rangle$ and $U_{rm}^{(SCF)}$. Show that the resulting eigenvalue problem of the type (9.162) is non-linear in $U_{rm}^{(SCF)}$.
Exercise 9.7.2: Derive (9.177).

9.8 Mean Field Theory for Macroscopic Systems

The SCF algorithm has been developed in Sections 9.4–9.7 for finite systems. In the present Section we want to adapt the algorithm to systems containing a macroscopically large number of fermions. We consider as an example the Hubbard model of an infinite linear lattice and apply the model to describe magnetic instabilities in metals.

The Hubbard Model

The Hubbard model serves today the important role as the simplest manifestation of strongly correlated electron systems which arise in many instances in molecular and solid state physics, for example, in the two-dimensional copper oxide lattices of high temperature superconductors. The one-dimensional Hubbard model is described by the Hamiltonian

$$\hat{H} = -t \sum_{r\sigma} \left(c^\dagger_{r+1,\sigma} c_{r\sigma} + c^\dagger_{r\sigma} c_{r+1,\sigma} \right) + \frac{v}{2} \sum_{\substack{r \\ \sigma\sigma'}} c^\dagger_{r\sigma} c^\dagger_{r\sigma'} c_{r\sigma'} c_{r\sigma} \ . \tag{9.178}$$

Here the index r, $r \in \mathbb{Z}$, describes the sites of a linear lattice. We assume that there are altogether $S = 2N_0$ lattice sites, labeled $r = -N_0+1, \ldots, N_0$ which are populated by $2N$ electrons (we choose an even number of electrons for convenience). Both N_0 and N are macroscopically large numbers. The operator $c^\dagger_{r\sigma}$ ($c_{r\sigma}$) creates (destroys) an electron with spin σ ($\sigma = \alpha, \beta$ for 'up' and 'down' spins, respectively) in state $|r\rangle$ at lattice site r. t describes the coupling of state $|r\rangle$ to states $|r \pm 1\rangle$ at the two neighboring lattice sites; t is assumed to be a positive, real number. According to (9.178) v contributes only in case that two electrons occupy the same lattice site, i.e., v represents the 'on-site' Coulomb repulsion.
The first term on the r.h.s. of (9.178), is identical to that of the independent–particle Hamiltonian H_0 (9.128), i.e.,

$$\hat{H}_o = -t \sum_{r\sigma} \left(c^\dagger_{r+1,\sigma} c_{r\sigma} + c^\dagger_{r\sigma} c_{r+1,\sigma} \right) \ . \tag{9.179}$$

The *potential energy* term, i.e., the second term on ther.h.s. of (9.178), in the usual two–particle operator form, reads

$$\hat{V} \ = \ \frac{1}{2} \sum_{\substack{r,s,t,u \\ \sigma\sigma'}} \langle r, s|\hat{v}|t, u\rangle \, c^\dagger_{r\sigma} c^\dagger_{s\sigma'} c_{u\sigma'} c_{t\sigma} \ . \tag{9.180}$$

where

$$\langle r, s|\hat{v}|u, t\rangle = v \ \delta_{ru} \delta_{st} \delta_{rs} \ . \tag{9.181}$$

The Hubbard model, as stated through Hamiltonian (9.178), is characterized through the parameters t, v of the Hubbard Hamiltonian (9.178) and through the so-called *filling factor* n

$$n = \frac{2N}{S} = \frac{N}{N_0} \ . \tag{9.182}$$

n can assume values $0 \leq n \leq 2$. The case $n = 1$ is termed the *half filled band* case, referring to the fact that in this case the *band* of single particle energies (9.192) derived below is filled half.
In spite of the simplicity of its Hamiltonian the Hubbard model, except in case of one-dimensional systems, cannot be solved exactly. Due to the lack of an exact solution in dimensions higher than one, approximation schemes like the self-consistent field approximation play an important role.

Since we are dealing presently with a macroscopic system, boundary effects are assumed to be negligible, a freedom which we employ to adopt so-called *periodic boundary conditions* like those for the example "$2N$ independent electrons on a ring" on page 262. Accordingly, we identify $|r + N_0\rangle = |r\rangle$ and later let N_0 go to infinity.

SCF Ground State for the Hubbard Model

We seek to determine the ground state of the Hubbard model stated by (9.178) within the framework of the self-consistent field theory. For this purpose we apply the SCF theory presented in Section 9.6. Accordingly, we assume that the SCF ground state is given by (9.174), i.e., by

$$||SCF\rangle = \prod_{m=0}^{N-1} d^\dagger_{m\alpha} d^\dagger_{m\beta} |0\rangle \ , \tag{9.183}$$

where the operators $d^\dagger_{m\sigma}$ are related to the original creation operators $c^\dagger_{r\sigma}$ through the unitary transformation $\mathbf{U}$ defined in (9.144). The self-consistent field Hamiltonian $\hat{H}_{mf}$ which corresponds to the Hubbard Hamiltonian (9.178) can be determined by applying (9.155, 9.154). Using (9.181, 9.179), one obtains for the mean field Hamiltonian

$$\hat{H}_{mf}(\mathbf{U}) = \hat{H}_0 + v \sum_{r,\sigma} \left(\sum_{m=0}^{N-1} U^*_{rm} U_{rm} \right) c^\dagger_{r\sigma} c_{r\sigma} \ . \tag{9.184}$$

We are now ready to apply *step 1* of the SCF algorithm and diagonalize term $\hat{H}_0$. This task has been solved already on page 262. The unitary transformation which diagonalizes $\hat{H}_0$, in fact, is

$$U_{rm} = \frac{1}{\sqrt{2N_0}} \exp\left(irm\pi/N_0\right) \ . \tag{9.185}$$

According to (9.140), the operator $\hat{H}_0$ can be re-written in the diagonal form

$$\hat{H}_0 = \sum_{\substack{m=-N_0+1 \\ \sigma}}^{N_0} \epsilon_m \, d^\dagger_{m\sigma} d_{m\sigma} \tag{9.186}$$

where

$$\epsilon_m = -2t \cos \frac{m\pi}{N_0} \ . \tag{9.187}$$

The energy levels are labeled such that $\epsilon_0 < \epsilon_{\pm 1} < \ldots < \epsilon_{\pm(N_0-1)} < \epsilon_{N_0}$ holds as can be readily verified.

We can now embark on *step 2* of the SCF algorithm. Inserting (9.185) into (9.184) and using (9.186) one obtains for the mean field Hamiltonian

$$\hat{H}^{(2)}_{mf} = \sum_{\substack{m=-N_0+1 \\ \sigma}}^{N_0} \epsilon_m \, d^\dagger_{m\sigma} d_{m\sigma} + v \frac{N}{2N_0} \sum_{r\sigma} c^\dagger_{r\sigma} c_{r\sigma} \ . \tag{9.188}$$

Expressing $c^\dagger_{r\sigma}$, $c_{r\sigma}$ through $d^\dagger_{m\sigma}$, $d_{n\sigma}$ according to (9.109) and using (9.185) one can prove

$$\sum_{r\sigma} c^\dagger_{r\sigma} c_{r\sigma} = \sum_{\substack{m=-N_0+1 \\ \sigma}}^{N_0} d^\dagger_{m\sigma} d_{m\sigma} \; . \tag{9.189}$$

and, hence,

$$\hat{H}_{mf} = \hat{H}^{(2)}_{mf} = \sum_{\substack{m=-N_0+1 \\ \sigma}}^{N_0} \left(\epsilon_m + \frac{vn}{2}\right) d^\dagger_{m\sigma} d_{m\sigma} \tag{9.190}$$

We have used in the latter expression the definition (9.182) of the filling factor n [c.f. (9.182)]. The Hamiltonian $\hat{H}_{mf}$, as given in (9.190), is already diagonal, i.e., the self–consistency condition is exactly met and the SCF algorithm converged after two steps. In other words, the ground state of the Hubbard model in the self-consistent field approximation can be determined exactly. The energy E_{mf} of this state can be calculated readily by using (9.177) and

$$\langle \tilde{m}\tilde{m}'|\hat{v}|\tilde{m}\tilde{m}'\rangle = \langle \tilde{m}\tilde{m}'|\hat{v}|\tilde{m}'\tilde{m}\rangle = \frac{v}{2N_0}$$

which holds in the present case. E_{mf} is then

$$E_{mf} = 2\sum_{m=0}^{N-1}\left(\epsilon_m + \frac{vn}{2}\right) - N\frac{vn}{2} \;=\; 2\sum_{m=0}^{N-1}\left(\epsilon_m + \frac{vn}{4}\right) \tag{9.191}$$

where ϵ_m is given in (9.187). Apparently, the electrons in the present description behave like independent particles with energies $\bar{\epsilon}_m = \epsilon_m + \frac{vn}{4}$.
So far we have not exploited the fact that N_0 and N are macroscopically large quantities. When N_0 becomes macroscopically large the single particle discrete energy levels (9.187) form a quasi-continuous energy band. Indeed, for $k_m = m\pi/N_0$ holds $k_m \in \left\{-\pi + \frac{1}{N_0}, -\pi + \frac{2}{N_0} \ldots, \pi\right\}$ and, in the limit $N_0 \to \infty$, holds $k \equiv k_m \in]-\pi, \pi]$. The energy ϵ_m can be expressed as a function of a continuous variable $k \in]-\pi, \pi]$, namely, through the so-called *dispersion law*

$$\epsilon(k) = -2t\cos k \; , \qquad -\pi \le k \le \pi \; . \tag{9.192}$$

This dispersion law is presented in Fig. 9.1. One can see that $-2t \le \epsilon(k) \le 2t$ holds. The bottom of the energy band corresponds to $\epsilon(0) = -2t$ and the width of the energy band is $4t$. The ground state of the system, i.e., the state with the lowest possible energy, can be obtained, according to the Pauli principle, by filling up this energy band with electrons from the bottom of the band (which corresponds to $k = 0$) to a maximum energy, called the *Fermi energy*, denoted by ϵ_F (see Fig. 9.1).

Alternative Description of the Mean Field Approximation

The state (9.183) is not the only candidate for the mean field ground state of the Hubbard model. To demonstrate this we consider an alternative formulation of the mean field approximation for Hamiltonian (9.178). Indeed, there exist several ways of formulating mean field approximations, even for one and the same Hamiltonian. Usually, the results of different formulations agree qualitatively, but may differ quantitatively.

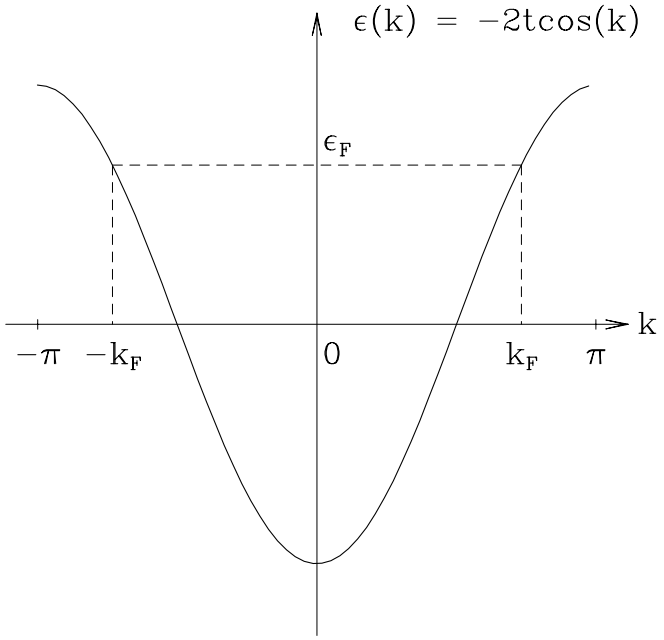


Figure 9.1: The dispersion law $\epsilon(k)$ for independent electrons on an infinite one dimensional lattice. $\epsilon_F = \epsilon(\pm k_F)$ denotes the Fermi energy.

First, let us express the interaction term in (9.178) in terms of the occupation number operators $\hat{N}_{r\sigma} = c^\dagger_{r\sigma} c_{r\sigma}$. The anti-commutation properties (9.45, 9.46) of the fermionic creation and annihilation operators yield

$$c^\dagger_{r\sigma} c^\dagger_{r\sigma'} c_{r\sigma'} c_{r\sigma} = -c^\dagger_{r\sigma} c^\dagger_{r\sigma'} c_{r\sigma} c_{r\sigma'} =$$

$$c^\dagger_{r\sigma} c_{r\sigma} c^\dagger_{r\sigma'} c_{r\sigma'} - c^\dagger_{r\sigma} c_{r\sigma} \delta_{\sigma\sigma'} = \hat{N}_{r\sigma} \hat{N}_{r\sigma'} - \hat{N}^2_{r\sigma} \delta_{\sigma\sigma'} \ , \tag{9.193}$$

where, in the last term on the right hand side, we have employed (9.56). Inserting (9.193) into (9.178) and performing the summation over the spin indices results in

$$\hat{H} = \hat{H}_0 + v \sum_r \hat{N}_{r\alpha} \hat{N}_{r\beta} \ . \tag{9.194}$$

For a given reference state, say the one given in (9.183), the operator $\hat{N}_{r\sigma}$ can be written

$$\hat{N}_{r\sigma} = \langle N_{r\sigma} \rangle + \delta \left(N_{r\sigma} \right) \ , \tag{9.195}$$

where $\langle N_{r\sigma} \rangle$ is the mean occupation number of the one particle state $|r\sigma\rangle$, i.e.,

$$\langle N_{r\sigma} \rangle \ = \ \langle SCF || N_{r\sigma} || SCF \rangle \ , \tag{9.196}$$

and where $\delta \left(N_{r\sigma} \right) = \hat{N}_{r\sigma} - \langle N_{r\sigma} \rangle$ describes the *fluctuations* of $\hat{N}_{r\sigma}$, i.e., the deviation of $\hat{N}_{r\sigma}$ from its mean value $\langle N_{r\sigma} \rangle$ for the corresponding reference state. The mean field approximation neglects the

fluctuations to second order and the mean field Hamiltonian is obtained from (9.194) by dropping all the terms which contain the fluctuations to second order. In this approximation one can express

$$\begin{aligned}\hat{N}_{r\alpha}\hat{N}_{r\beta} &= \left(\langle N_{r\alpha}\rangle + \delta\left(N_{r\alpha}\right)\right)\left(\langle N_{r\beta}\rangle + \delta\left(N_{r\beta}\right)\right) \\ &\approx \langle N_{r\alpha}\rangle\langle N_{r\beta}\rangle + \langle N_{r\alpha}\rangle\delta\left(N_{r\beta}\right) + \langle N_{r\beta}\rangle\delta\left(N_{r\alpha}\right) \\ &= \langle N_{r\alpha}\rangle\hat{N}_{r\beta} + \langle N_{r\beta}\rangle\hat{N}_{r\alpha} - \langle N_{r\alpha}\rangle\langle N_{r\beta}\rangle \end{aligned} \tag{9.197}$$

and the corresponding mean field Hamiltonian becomes

$$\hat{H}_{mf} = \hat{H}_0 + v\sum_r\left(\langle N_{r\alpha}\rangle\hat{N}_{r\beta} + \langle N_{r\beta}\rangle\hat{N}_{r\alpha}\right) - v\sum_r\langle N_{r\alpha}\rangle\langle N_{r\beta}\rangle\ . \tag{9.198}$$

Let us assume that the ground state of the Hubbard model, in the mean field approximation, is given by (9.183). Because of

$$\begin{aligned}\langle N_{r\sigma}\rangle &= \langle c^\dagger_{r\sigma}c_{r\sigma}\rangle \;=\; \sum_{m,m'}U_{rm}U^*_{rm'}\langle d^\dagger_{m\sigma}d_{m'\sigma}\rangle \\ &= \sum_{m=0}^{N-1}U_{rm}U^*_{rm} \;=\; \frac{N}{2N_0} \;=\; \frac{n}{2}\,,\end{aligned}$$

holds

$$\begin{aligned}\hat{H}_{mf} &= \hat{H}_0 + \frac{mv}{2}\sum_{r\sigma}c^\dagger_{r\sigma}c_{r\sigma} - 2N_0\frac{vn^2}{2} \\ &= \sum_{m\sigma}\left(\epsilon_m + \frac{nv}{2}\right)d^\dagger_{m\sigma}d_{m\sigma} - N\frac{vn}{2}\ ,\end{aligned} \tag{9.199}$$

which is identical to expression (9.191) for the ground state energy. Note that the procedure employed in (9.195), i.e., separating the occupation number operator into a mean value plus fluctuation and neglecting the fluctuations in 2nd order, yields essentially the same mean field Hamiltonian (9.198) as the one obtained by using the mean field potential (9.146).

Spin-Polarized Mean Field Ground State

We want to consider now a ground state for $\hat{H}_{mf}$ defined through

$$\langle N_{r\sigma}\rangle = n_\sigma = \text{const} \tag{9.200}$$

allowing, however, that n_α and n_β assume different values. For such state the mean occupation number $n_{r\sigma}$ is uniform at all lattice sites, but the mean number of particles with spin α can be different from the mean number of particles with spin β. Therefore, the ground state of the system can have a non-zero local spin and, consequently, non-vanishing magnetization. In case $n_\alpha = n_\beta$ the following construction will lead to a ground state which coincides with the non-magnetic ground state (9.183) .
To determine a magnetic ground state we note

$$\sum_r\hat{N}_{r\sigma} = \sum_{mm'}\left(\sum_r U_{rm}U^*_{rm'}\right)d^\dagger_{m\sigma}d_{m'\sigma} = \sum_{mm'}\delta_{mm'}\hat{N}_{m\sigma} = \sum_m\hat{N}_{m\sigma}\ ,$$

from which we obtain

$$n_\sigma = \frac{1}{2N_0}\sum_m \langle N_{m\sigma}\rangle \ . \tag{9.201}$$

The mean field Hamiltonian is then

$$\hat{H}_{mf} = \sum_m \left[(\epsilon_m + vn_\beta)\,\hat{N}_{m\alpha} + (\epsilon_m + vn_\alpha)\,\hat{N}_{m\beta}\right] - 2N_0 v n_\alpha n_\beta \ . \tag{9.202}$$

This Hamiltonian describes a system of non-interacting electrons with a spin-dependent dispersion law

$$\epsilon_{m\sigma} = \epsilon_m + vn_{-\sigma} \ . \tag{9.203}$$

The actual values of n_σ ($\sigma = \alpha, \beta$) are determined by minimizing the energy density of the ground state

$$\mathcal{E}_{mf} \equiv \frac{\langle H_{mf}\rangle}{2N_0} = \frac{1}{2N_0}\sum_{m,\sigma}\left[\epsilon_{m\sigma}\langle N_{m\sigma}\rangle\right] - vn_\alpha n_\beta \tag{9.204}$$

with respect to n_α and n_β, subject to the constraint

$$n = n_\alpha + n_\beta \ . \tag{9.205}$$

Such minimization is realized through the method of *Lagrangian multipliers*. According to this method one minimizes

$$\tilde{\mathcal{E}}_{mf} = \mathcal{E}_{mf} + \mu\,(n - n_\alpha - n_\beta) \ , \tag{9.206}$$

where μ is the Lagrangian multiplier considered an independent variable, i.e., $\tilde{\mathcal{E}}_{mf}$ is minimized with respect to n_α, n_β and μ. The additional term in (9.206) ascertains that condition (9.205) is met.

At this point we exploit the fact that our system is macroscopically large, i.e., $N_0 \sim 10^{23}$. In this limit the discrete energy spectrum (9.203) is provided by the continuous function

$$\epsilon_\sigma(k) = \epsilon(k) + vn_{-\sigma} \ , \tag{9.207}$$

where $\epsilon(k)$ is given by (9.192). This last equation tells us that the electrons with spin α (β) are accommodated by an energy sub–band $\epsilon_\alpha(k)$ ($\epsilon_\beta(k)$) which is obtained from the dispersion law of the non-interactive electrons $\epsilon(k)$ by an overall shift of vn_β (vn_α). The ground state (corresponding to the lowest possible energy) of the many electron system is obtained by filling these two energy sub–bands with electrons up to the same maximum energy value, the so-called Fermi energy $\epsilon_F = \mu$ (see below), as is shown schematically in Figure 9.2a. Since $\epsilon_\beta(k) - \epsilon_\alpha(k) = v(n_\alpha - n_\beta)$, one can see that an uneven occupation by electrons of the two energy bands yields a relative shift of the energy bands with respect to each other, e.g., if $n_\alpha > n_\beta$ then $\epsilon_\beta(k) > \epsilon_\alpha(k)$, such that the larger n_σ, the smaller is $\epsilon_\sigma(k)$ for a given k. Therefore, one can expect the system to lower its energy by assuming $n_\alpha \neq n_\beta$ and, hence, a spin–polarized ground state.

The values of n_α and n_β are determined by minimizing the energy density (9.206). For this purpose the sum in (9.206) needs to be evaluated. In the limit $N_0 \to \infty$ this sum can be expressed as an integral. In fact, for any function $f\left(\frac{m\pi}{N_0}\right)$ holds (compare with the definition of a definite integral as the limit of the corresponding Riemann sum)

$$\lim_{N_0\to\infty}\frac{1}{2N_0}\sum_{m=-N_0+1}^{N_0} f\left(\frac{m\pi}{N_0}\right) = \int_{-\pi}^{\pi}\frac{dk}{2\pi}f(k) \ . \tag{9.208}$$

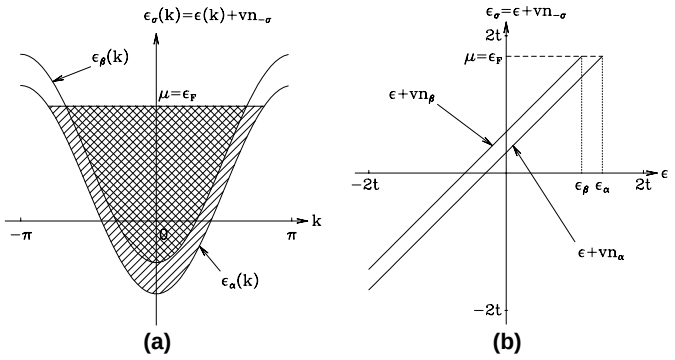


Figure 9.2: (a) Relative position of the two energy sub–bands $\epsilon_\alpha(k)$ and $\epsilon_\beta(k)$ for $\Delta = n_\alpha - n_\beta > 0$. The shaded areas represent the filled portions of these bands by electrons. (b) Graphical definition of the limiting energies ϵ_α and ϵ_β.

If the function f depends explicitly only on $\epsilon(k)$ one can state

$$\int_{-\pi}^{\pi} \frac{dk}{2\pi} f\left(\epsilon(k)\right) = \int_{-\infty}^{\infty} \rho(\epsilon) d\epsilon f(\epsilon) \tag{9.209}$$

where

$$\rho(\epsilon) = \int_{-\pi}^{\pi} \frac{dk}{2\pi} \delta\left(\epsilon - \epsilon(k)\right) \ . \tag{9.210}$$

Here $\delta(x)$ is the *Dirac–delta* function and $\rho(\epsilon)$, which is usually called the *density of states*, gives the the number of available one particle states per site, unit energy and a given spin orientation of the particle.
For a given dispersion law $\epsilon(k)$ the density of states $\rho(\epsilon)$ can be calculated by employing equation (9.210). In our case $\epsilon(k)$ is given by (9.192) and, therefore, holds

$$\rho(\epsilon) = \int_{-\pi}^{\pi} \frac{dk}{2\pi} \delta(\epsilon + 2t\cos k) \ . \tag{9.211}$$

In order to calculate the integral we recall the following property of the Dirac–delta function

$$\delta\left[f(x)\right] = \sum_i \frac{\delta(x - x_i)}{|f'(x_i)|} \ , \tag{9.212}$$

where x_i are the simple roots of the function $f(x)$. Since the equation $f(k) \equiv 2t\cos k + \epsilon = 0$ has two simple roots, namely $k_{1,2} = \pm(\pi - \arccos(\epsilon/2t))$, and noting

$$|f'(k)| = |2t\sin k| = 2t\sqrt{1-\cos^2 k} = \sqrt{(2t)^2 - \epsilon^2} \ ,$$

one obtains

$$\delta(\epsilon - \epsilon(k)) = \frac{\delta(k + \pi - \arccos(\epsilon/2t)) + \delta(k - \pi + \arccos(\epsilon/2t))}{\sqrt{(2t)^2 - \epsilon^2}} \ .$$

Inserting this last result into (9.211) and using

$$\int_{-\pi}^{\pi} dk \delta(k \pm (\pi - \arccos(\epsilon/2t))) = \theta(2t - |\epsilon|) \ ,$$

where $\theta(x)$ is the *step function*, i.e., $\theta(x) = 1$ if $x > 0$ and $\theta(x) = 0$ if $x < 0$, one arrives at the following expression of the density of states

$$\rho(\epsilon) = \frac{\theta(2t - |\epsilon|)}{\pi\sqrt{(2t)^2 - \epsilon^2}} \; . \tag{9.213}$$

The presence of the θ function in the above formula guaranties that the density of states vanishes for $|\epsilon| > 2t$, i.e., outside the energy band $\epsilon(k)$; if we assume that ϵ is restricted to the interval $]-2t, 2t[$ then the θ function in (9.213) can simply be replaced by 1.
We can employ the results obtained to express the ground state energy density (9.204) through an integral expression. One obtains

$$\tilde{\mathcal{E}}_{mf} = \int_{-2t}^{\epsilon_\alpha} \epsilon\rho(\epsilon)d\epsilon + \int_{-2t}^{\epsilon_\beta} \epsilon\rho(\epsilon)d\epsilon + vn_\alpha n_\beta + \mu(n - n_\alpha - n_\beta) \; , \tag{9.214}$$

where ϵ_σ denotes the value of $\epsilon(k)$ which corresponds to the top of the filled portion of the energy sub–band $\epsilon_\sigma(k)$ [c.f. Fig. 9.2], i.e.,

$$\epsilon_F = \epsilon_\alpha + vn_\beta = \epsilon_\beta + vn_\alpha \; . \tag{9.215}$$

n_α and n_β in (9.214) are given by the sum (9.201). Replacing again the sum by an integral over energy one can write

$$n_\sigma \; = \; \int_{-2t}^{\epsilon_\sigma} \rho(\epsilon)d\epsilon \tag{9.216}$$

Using (9.213) and carrying out the resulting integral yields

$$n_\sigma = \frac{1}{\pi} \arcsin(\epsilon_\sigma/2t) \; . \tag{9.217}$$

Since $\tilde{\mathcal{E}}_{mf}$ [c.f. (9.206)] depends only on continuous quantities, namely, on $\epsilon_\alpha, \epsilon_\beta$ and μ, the necessary conditions for a ground state of minimum energy are

$$\frac{\partial\tilde{\mathcal{E}}_{mf}}{\partial\epsilon_\alpha} = 0 \; , \qquad \frac{\partial\tilde{\mathcal{E}}_{mf}}{\partial\epsilon_\beta} = 0 \; , \qquad \frac{\partial\tilde{\mathcal{E}}_{mf}}{\partial\mu} = 0 \; . \tag{9.218}$$

The derivatives can be readily determined and the conditions (9.218) read

$$\frac{\partial\tilde{\mathcal{E}}_{mf}}{\partial\epsilon_\alpha} = \epsilon_\alpha + vn_\beta - \mu = 0 \; , \tag{9.219}$$

$$\frac{\partial\tilde{\mathcal{E}}_{mf}}{\partial\epsilon_\beta} = \epsilon_\beta + vn_\alpha - \mu = 0 \; , \tag{9.220}$$

and

$$\frac{\partial\tilde{\mathcal{E}}_{mf}}{\partial\mu} = n - n_\alpha - n_\beta = 0 \; . \tag{9.221}$$

Equations (9.219–9.221) and (9.216) allow one, in principle, to determine the unknown quantities ϵ_α, ϵ_β, μ, n_α and n_β as a function of the parameters of the Hubbard model, namely, of t, v and n. From (9.219, 9.220) μ can be readily obtained

$$\mu = \frac{1}{2}\,(\epsilon_\alpha + \epsilon_\beta) + \frac{vn}{2} \; , \tag{9.222}$$

Comparision of (9.222) and (9.215) reveals that μ is indeed equal to the Fermi energy ϵ_F (see also Figure 9.2b).
One still needs to determine ϵ_α and ϵ_β. For this purpose we consider the *magnetization* of the mean field ground state

$$\Delta \equiv n_\alpha - n_\beta \; . \tag{9.223}$$

$\Delta = 0$ corresponds to the ground state (9.183) without magnetization. One can express Δ as a function of ϵ_α and ϵ_β in two different ways. First, from (9.219–9.220) and (9.223 one obtains

$$\Delta = \frac{\epsilon_\alpha - \epsilon_\beta}{v} \; , \tag{9.224}$$

and second, from (9.217) and (9.223) follows

$$\Delta = \int_{\epsilon_\beta}^{\epsilon_\alpha} \rho(\epsilon) d\epsilon = \int_0^{v\Delta} \rho\left(\epsilon_\beta + \epsilon\right) d\epsilon \; . \tag{9.225}$$

Defining the function

$$f(\Delta) = \int_0^{v\Delta} \rho\left(\epsilon_\beta + \epsilon\right) d\epsilon \; , \tag{9.226}$$

one can obtain Δ by solving (cf. (9.226)

$$\Delta = f(\Delta) \; . \tag{9.227}$$

Since $f(0) = 0$ one can infer that $\Delta = 0$ is always a solution (the so called *paramagnetic* solution) of condition (9.227).
Equation (9.227) can be solved graphically. For this purpose one plots $f(\Delta)$ versus Δ as is shown schematically in Figure 9.3. We assume in the following discussion that the implicit Δ dependence of ϵ_β in $f(\Delta)$ can be neglected.
From the definition (9.226) follows that $f(\Delta)$ is a monotonically increasing function of Δ. Indeed, the slope of $f(\Delta)$ is positive for all $\Delta \geq 0$, as reflected by the expression

$$\frac{d[f(\Delta)]}{d\Delta} = v\rho(\epsilon_\beta + \Delta) = \frac{v}{\pi\sqrt{(2t)^2 - (v\Delta + \epsilon_\beta)^2}} > 0 \; . \tag{9.228}$$

Since the total number of states per site is finite, $f(\Delta)$ converges to a maximum value as $\Delta \to \Delta_{max} = n$. Accordingly, condition (9.227) will have a second, non trivial, so called *ferromagnetic* solution, if, and only if, the slope of $f(\Delta)$ at the origin is larger than 1 [c.f. Figure 9.3]. If the slope of $f(\Delta)$ at the origin is less than 1, condition (9.227) has only the trivial solution $\Delta = 0$.
The slope of $f(\Delta)$ at the origin depends on the Hubbard model parameter v. For v above a critical value v_c, the Hubbard model has a ground state with $\Delta \neq 0$, i.e., a magnetic ground state. One can determine v_c by equating the slope of $f(\Delta)$ at the origin to 1, i.e., through the condition [c.f. (9.228)]

$$v_c \rho\left(\epsilon_\beta\right) = 1 \; , \tag{9.229}$$

This cindition is known as the *Stoner criterion* and gives the critical value of v which determines the onset of the ferromagnetic long range order. The Stoner criterion (9.229) has a wider range of validity than one can infer from the present analysis of the Hubbard model.

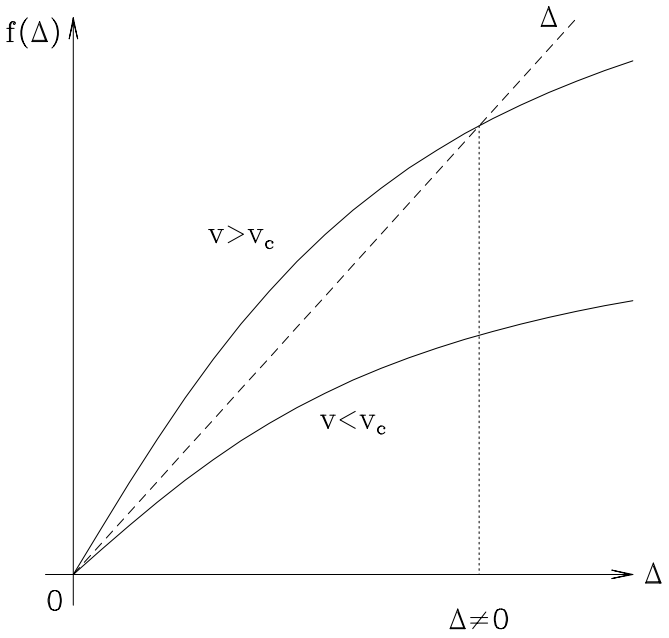


Figure 9.3: Graphical solution of the mean field equation $\Delta = f(\Delta)$.

By parameterizing ϵ_α and ϵ_β

$$\epsilon_\alpha = 2t \sin x \ , \qquad \epsilon_\beta = 2t \sin y \ , \tag{9.230}$$

one obtains from (9.219, 9.220, 9.205) the following set of equations determining x, y and, hence, ϵ_α, ϵ_β

$$x + y = \pi(n-1) \tag{9.231}$$

$$x - y = \pi \frac{2t}{v}(\sin x - \sin y) = \pi\Delta \ . \tag{9.232}$$

For a given $\frac{2t}{v}$ the values of x and y can be obtained numerically and other relevant quantities mentioned above can be calculated.
We like to determine finally the dependence of $\frac{v_c}{2t}$ on the filling factor n. The corresponding analytical expression provides the phase diagram of the ground state: when v approaches v_c from above, the magnetization Δ of the ground state vanishes and remains zero for values $v < v_c$. Equation (9.232) states that $x - y$ is proportional to Δ. Since Δ vanishes near v_c we set $x = y + \delta$ where δ is a small quantity. From (9.231) follows

$$y = \frac{\pi}{2}(n-1) - \frac{\delta}{2} \approx \frac{\pi}{2}(n-1) \ . \tag{9.233}$$

Employing the approximation $\sin(y+\delta) - \sin y \approx \delta \cos y$ for small δ, (9.232) together with (9.233) yield the desired expression

$$\frac{v_c}{2t} = \pi \cos\left[\frac{\pi}{2}(n-1)\right] \ . \tag{9.234}$$

The resulting phase diagram of the ground state, i.e., the plot of $\frac{v_c}{2t}$ vs. n, is presented in Figure 9.4. As already mentioned, the ground state is characterized by the quantities (v, t, n). Accordingly, any

mean field ground state is represented by a point in the phase diagram. The magnetic nature of the ground state of the system depends on whether the representative point lies inside the ferromagnetic or inside the paramagnetic domain of the phase diagram.

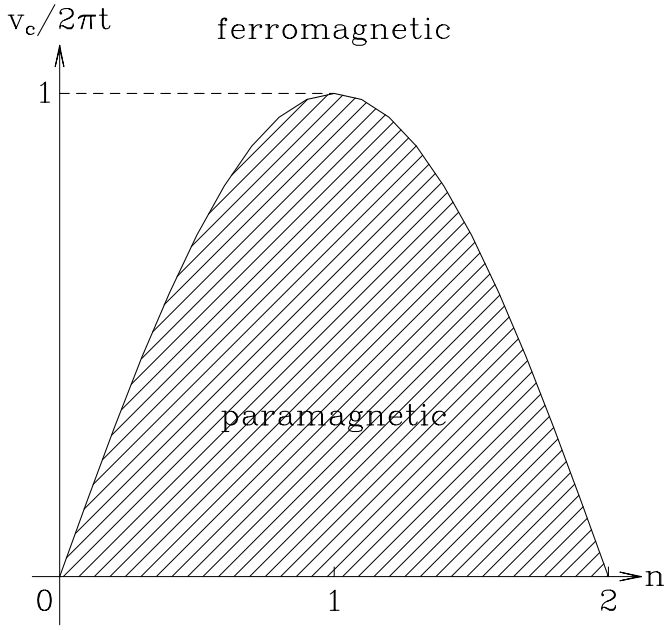


Figure 9.4: Ground state phase diagram of the mean field Hamiltonian (9.202).

Concluding Remarks

At the end of our analysis of the mean field ground state of the one-dimensional Hubbard model, it is natural to ask oneself if the results obtained reflect, at least qualitatively, the properties of the real ground state of the Hubbard model. Well, the answer is no. The real ground state of the one-dimensional Hubbard model is quite different in nature from the mean field ground state discussed here. The reason is that the effect of quantum fluctuations, which are neglected in the mean field theory, in one spatial dimension is very strong.
Nevertheless, the analysis presented above is not quite useless. First, once we specify a class of possible states to which the real ground state might belong to, the method described above gives a systematic way of singling out the state with the lowest possible energy which might be a good candidate for the real ground state of the system. Second, since in one spatial dimension there are many exact results available, the application of the mean field theory for these systems provides an excellent testing opportunity of these theories by comparing their predicted results with the exact ones. Third, the mean field theory of the one-dimensional Hubbard model can be extended in a straightforward way to higher spatial dimensions. In general, the effect of quantum fluctuations is getting less important as the dimensionality of the system is increased. For example, in three spatial dimensions the theory presented above works fine in the case of the transitional–metal oxides such as NiO and CoO. However, strong electron correlation effects , even in these materials, can lead to serious modifications of the mean field ground state.

In case of two spatial dimensions things are more complicated. On the one hand, the lack of exact solutions, and on the other hand, the strong effect of quantum fluctuations of the strongly correlated electron system described by the Hubbard Hamiltonian makes all the presently existing solutions questionable. In fact, in the case of the copper oxide high temperature superconductors, which involve strongly correlated quasi two-dimensional electron systems, a reliable microscopic theory is still lacking. The available mean field theories cannot account for all the striking, unusual physical properties of these materials.

Chapter 10

Relativistic Quantum Mechanics

In this Chapter we will address the issue that the laws of physics must be formulated in a form which is Lorentz–invariant, i.e., the description should not allow one to differentiate between frames of reference which are moving relative to each other with a constant uniform velocity $\vec{v}$. The transformations beween such frames according to the Theory of Special Relativity are described by Lorentz transformations. In case that $\vec{v}$ is oriented along the x_1–axis, i.e., $\vec{v} = v_1\hat{x}_1$, these transformations are

$$x_1\prime = \frac{x_1 - v_1 t}{\sqrt{1 - \left(\frac{v_1}{c}\right)^2}}\ ,\ t' = \frac{t - \frac{v_1}{c^2}x_1}{\sqrt{1 - \left(\frac{v_1}{c}\right)^2}}\ ,\ x_2' = x_2\ ;\ x_3' = x_3 \tag{10.1}$$

which connect space time coordinates (x_1, x_2, x_3, t) in one frame with space time coordinates (x_1', x_2', x_3', t') in another frame. Here c denotes the velocity of light. We will introduce below Lorentz-invariant differential equations which take the place of the Schrödinger equation of a particle of mass m and charge q in an electromagnetic field [c.f. (refeq:ham2, 8.45)] described by an electrical potential $V(\vec{r}, t)$ and a vector potential $\vec{A}(\vec{r}, t)$

$$i\hbar\frac{\partial}{\partial t}\psi(\vec{r}, t) = \left[\frac{1}{2m}\left(\frac{\hbar}{i}\nabla - \frac{q}{c}\vec{A}(\vec{r}, t)\right)^2 + qV(\vec{r}, t)\right]\psi(\vec{r}, t) \tag{10.2}$$

The replacement of (10.2) by Lorentz–invariant equations will have two surprising and extremely important consequences: some of the equations need to be formulated in a representation for which the wave functions $\psi(\vec{r}, t)$ are vectors of dimension larger one, the components representing the spin attribute of particles and also representing together with a particle its anti-particle. We will find that actually several Lorentz–invariant equations which replace (10.2) will result, any of these equations being specific for certain classes of particles, e.g., spin–0 particles, spin–$\frac{1}{2}$ particles, etc. As mentioned, some of the equations describe a particle together with its anti-particle. It is not possible to uncouple the equations to describe only a single type particle without affecting negatively the Lorentz invariance of the equations. Furthermore, the equations need to be interpreted as actually describing many–particle-systems: the equivalence of mass and energy in relativistic formulations of physics allows that energy converts into particles such that any particle described will have 'companions' which assume at least a virtual existence.
Obviously, it will be necessary to begin this Chapter with an investigation of the group of Lorentz transformations and their representation in the space of position $\vec{r}$ and time t. The representation

in Sect. 10.1 will be extended in Sect. 10.4 to cover fields, i.e., wave functions $\psi(\vec{r},t)$ and vectors with functions $\psi(\vec{r},t)$ as components. This will provide us with a general set of Lorentz–invariant equations which for various particles take the place of the Schrödinger equation. Before introducing these general Lorentz–invariant field equations we will provide in Sects. 10.5, 10.7 a heuristic derivation of the two most widely used and best known Lorentz–invariant field equations, namely the Klein–Gordon (Sect. 10.5) and the Dirac (Sect. 10.7) equation.

10.1 Natural Representation of the Lorentz Group

In this Section we consider the natural representation of the Lorentz group $\mathcal{L}$, i.e. the group of Lorentz transformations (10.1). Rather than starting from (10.1), however, we will provide a more basic definition of the transformations. We will find that this definition will lead us back to the transformation law (10.1), but in a setting of representation theory methods as applied in Secti. 5 to the groups SO(3) and SU(2) of rotation transformations of space coordinates and of spin.
The elements $L \in \mathcal{L}$ act on 4–dimensional vectors of position– and time–coordinates. We will denote these vectors as follows

$$x^\mu \stackrel{\text{def}}{=} (x^0, x^1, x^2, x^3) \tag{10.3}$$

where $x^0 = ct$ describes the time coordinate and $(x^1, x^2, x^3) = \vec{r}$ describes the space coordinates. Note that the components of x^μ all have the same dimension, namely that of length. We will, henceforth, assume new units for time such that the velocity of light c becomes $c = 1$. This choice implies $dim(\text{time}) = dim(\text{length})$.

Minkowski Space

Historically, the Lorentz transformations were formulated in a space in which the time component of x^μ was chosen as a purely imaginary number and the space components real. This space is called the Minkowski space. The reason for this choice is that the transformations (10.1) leave the quantity

$$s^2 = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 \tag{10.4}$$

invariant, i.e., for the transformed space-time–cordinates $x'^\mu = (x'^0, x'^1, x'^2, x'^3)$ holds

$$(x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = (x'^0)^2 - (x'^1)^2 - (x'^2)^2 - (x'^3)^2 \,. \tag{10.5}$$

One can interprete the quantity $\sqrt{-s^2}$ as a distance in a 4–dimensional Euclidean space if one chooses the time component purely imaginary. In such a space Lorentz transformations correspond to 4-dimensional rotations. Rather than following this avenue we will introduce Lorentz transformations within a setting which does not require real and imaginary coordinates.

The Group of Lorentz Transformations $\mathcal{L}$ = O(3,1)

The Lorentz transformations L describe the relationship between space-time coordinates x^μ of two reference frames which move relative to each other with uniform fixed velocity $\vec{v}$ and which might be reoriented relative to each other by a rotation around a common origin. Denoting by x^μ the

coordinates in one reference frame and by x'^{μ} the coordinates in the other reference frame, the Lorentz transformations constitute a linear transformation which we denote by

$$x'^{\mu} \;=\; \sum_{\nu=0}^{3} L^{\mu}{}_{\nu} x_{\nu} \; . \tag{10.6}$$

Here $L^{\mu}{}_{\nu}$ are the elements of a 4×4–matrix representing the Lorentz transformation. The upper index closer to 'L' denotes the first index of the matrix and the lower index ν further away from 'L' denotes the second index. [A more conventional notation would be $L_{\mu\nu}$, however, the latter notation will be used for different quantities further below.] The following possibilities exist for the positioning of the indices $\mu, \nu \;=\; 0, 1, 2, 3$:

$$\text{4-vector: } x^{\mu},\; x_{\mu} \; ; \quad 4 \times 4 \text{ tensor: } A^{\mu}{}_{\nu},\; A_{\mu}{}^{\nu},\; A^{\mu\nu},\; A_{\mu\nu} \; . \tag{10.7}$$

The reason for the notation is two-fold. First, the notation in (10.6) allows us to introduce the so-called *summation conventon*: any time the *same* index appears in an upper **and** a lower position, summation over that index is assumed without explicitly noting it, i.e.,

$$\underbrace{y_{\mu} x^{\mu}}_{\text{new}} \;=\; \underbrace{\sum_{\mu=0}^{3} y_{\mu} x^{\mu}}_{\text{old}} \; ; \; \underbrace{A^{\mu}{}_{\nu} x^{\nu}}_{\text{new}} \;=\; \underbrace{\sum_{\nu=0}^{3} A^{\mu}{}_{\nu} x^{\nu}}_{\text{old}} \; ; \; \underbrace{A^{\mu}{}_{\nu} B^{\nu}{}_{\rho}}_{\text{new}} \;=\; \underbrace{\sum_{\nu=0}^{3} A^{\mu}{}_{\nu} B^{\nu}{}_{\rho}}_{\text{old}} \; . \tag{10.8}$$

The summation convention allows us to write (10.6) $x'^{\mu} \;=\; L^{\mu}{}_{\nu} x^{\nu}$. The second reason is that upper and lower positions allow us to accomodate the expression (10.5) into scalar products. This will be explained further below.
The Lorentz transformations are non-singular 4×4–matrices with real coefficients, i.e., $L \in \mathrm{GL}(4, \mathbb{R})$, the latter set constituting a group. The Lorentz transformations form the subgroup of all matrices which leave the expression (10.5) invariant. This condition can be written

$$x^{\mu} g_{\mu\nu} x^{\nu} \;=\; x'^{\mu} g_{\mu\nu} x'^{\nu} \tag{10.9}$$

where

$$(\, g_{\mu\nu} \,) \;=\; \begin{pmatrix} 1 & 0 & 0 & \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \;=\; \mathbf{g} \; . \tag{10.10}$$

Combining condition (10.9) and (10.6) yields

$$L^{\mu}{}_{\rho} g_{\mu\nu} L^{\nu}{}_{\sigma} \, x^{\rho} x^{\sigma} \;=\; g_{\rho\sigma} \, x^{\rho} x^{\sigma} \; . \tag{10.11}$$

Since this holds for any x^{μ} it must be true

$$L^{\mu}{}_{\rho} g_{\mu\nu} L^{\nu}{}_{\sigma} \;=\; g_{\rho\sigma} \; . \tag{10.12}$$

This condition specifies the key property of Lorentz transformations. We will exploit this property below to determine the general form of the Lorentz transformations. The subset of $\mathrm{GL}(4, \mathbb{R})$, the

elements of which satisfy this condition, is called O(3,1). This set is identical with the set of all Lorentz transformations $\mathcal{L}$. We want to show now $\mathcal{L} = \mathrm{O}(3,1) \subset \mathrm{GL}(4, \mathbb{R})$ is a group.
To simplify the following proof of the key group properties we like to adopt the conventional matrix notation for $L^{\mu}{}_{\nu}$

$$\mathbf{L} \;=\; (\, L^{\mu}{}_{\nu}\,) \;=\; \begin{pmatrix} L^0{}_0 & L^0{}_1 & L^0{}_2 & L^0{}_3 \\ L^1{}_0 & L^1{}_1 & L^1{}_2 & L^1{}_3 \\ L^2{}_0 & L^2{}_1 & L^2{}_2 & L^3{}_3 \\ L^3{}_0 & L^3{}_1 & L^3{}_2 & L^3{}_3 \end{pmatrix} . \tag{10.13}$$

Using the definition (10.10) of $\mathbf{g}$ one can rewrite the invariance property (10.12)

$$\mathbf{L}^T \mathbf{g} \mathbf{L} \;=\; \mathbf{g}\,. \tag{10.14}$$

From this one can obtain using

$$\mathbf{g}^2 \;=\; \mathbb{1} \tag{10.15}$$

$(\mathbf{g}\mathbf{L}^T\mathbf{g})\mathbf{L} \;=\; \mathbb{1}$ and, hence, the inverse of $\mathbf{L}$

$$\mathbf{L}^{-1} \;=\; \mathbf{g}\,\mathbf{L}^T\,\mathbf{g} \;=\; \begin{pmatrix} L^0{}_0 & -L^1{}_0 & -L^2{}_0 & -L^3{}_0 \\ -L^0{}_1 & L^1{}_1 & L^2{}_1 & L^3{}_1 \\ -L^0{}_2 & L^1{}_2 & L^2{}_2 & L^3{}_2 \\ -L^0{}_3 & L^1{}_3 & L^2{}_3 & L^3{}_3 \end{pmatrix} . \tag{10.16}$$

The corresponding expression for $(\mathbf{L}^T)^{-1}$ is obviously

$$(\mathbf{L}^T)^{-1} \;=\; (\mathbf{L}^{-1})^T \;=\; \mathbf{g}\,\mathbf{L}\,\mathbf{g}\,. \tag{10.17}$$

To demonstrate the group property of O(3,1), i.e., of

$$\mathrm{O}(3,1) \;=\; \{\, \mathbf{L}, \mathbf{L} \in \mathrm{GL}(4,\mathbb{R}),\, \mathbf{L}^T\mathbf{g}\mathbf{L} \;=\; \mathbf{g}\,\}\,, \tag{10.18}$$

we note first that the identity matrix $\mathbb{1}$ is an element of O(3,1) since it satisfies (10.14). We consider then $\mathbf{L}_1, \mathbf{L}_2 \in \mathrm{O}(3,1)$. For $\mathbf{L}_3 \;=\; \mathbf{L}_1\mathbf{L}_2$ holds

$$\mathbf{L}_3^T\,\mathbf{g}\,\mathbf{L}_3 \;=\; \mathbf{L}_2^T\mathbf{L}_1^T\,\mathbf{g}\,\mathbf{L}_1\mathbf{L}_2 \;=\; \mathbf{L}_2^T\,(\mathbf{L}_1^T\,\mathbf{g}\mathbf{L}_1)\,\mathbf{L}_2 \;=\; \mathbf{L}_2^T\,\mathbf{g}\,\mathbf{L}_2 \;=\; \mathbf{g}\,, \tag{10.19}$$

i.e., $\mathbf{L}_3 \in \mathrm{O}(3,1)$. One can also show that if $\mathbf{L} \in \mathrm{O}(3,1)$, the associated inverse obeys (10.14), i.e., $\mathbf{L}^{-1} \in \mathrm{O}(3,1)$. In fact, employing expressions (10.16, 10.17) one obtains

$$(\mathbf{L}^{-1})^T\,\mathbf{g}\,\mathbf{L}^{-1} \;=\; \mathbf{g}\mathbf{L}\mathbf{g}\mathbf{g}\mathbf{g}\mathbf{L}^T\mathbf{g} \;=\; \mathbf{g}\mathbf{L}\mathbf{g}\mathbf{L}^T\mathbf{g}\,. \tag{10.20}$$

Multiplying (10.14) from the right by $\mathbf{g}\mathbf{L}^T$ and using (10.15) one can derive $\mathbf{L}^T\mathbf{g}\mathbf{L}\mathbf{g}\mathbf{L}^T \;=\; \mathbf{L}^T$ and multiplying this from the left by by $\mathbf{g}(\mathbf{L}^T)^{-1}$ yields

$$\mathbf{L}\,\mathbf{g}\,\mathbf{L}^T \;=\; \mathbf{g} \tag{10.21}$$

Using this result to simplify the r.h.s. of (10.20) results in the desired property

$$(\mathbf{L}^{-1})^T\,\mathbf{g}\,\mathbf{L}^{-1} \;=\; \mathbf{g}\,, \tag{10.22}$$

i.e., property (10.14) holds for the inverse of $\mathbf{L}$. This stipulates that O(3,1) is, in fact, a group.

Classification of Lorentz Transformations

We like to classify now the elements of $\mathcal{L} = \mathrm{O}(3,1)$. For this purpose we consider first the value of $\det L$. A statement on this value can be made on account of property (10.14). Using $\det AB = \det A \det B$ and $\det A^T = \det A$ yields $(\det L)^2 = 1$ or

$$\det L = \pm 1 \,. \tag{10.23}$$

One can classify Lorentz transformations according to the value of the determinant into two distinct classes.
A second class property follows from (10.14) which we employ in the formulation (10.12). Considering in (10.12) the case $\rho = 0$, $\sigma = 0$ yields

$$\left(L^0{}_0\right)^2 - \left(L^1{}_0\right)^2 - \left(L^2{}_0\right)^2 - \left(L^3{}_0\right)^2 = 1 \,. \tag{10.24}$$

or since $(L^1{}_0)^2 + (L^2{}_0)^2 + (L^3{}_0)^2 \geq 0$ it holds $(L^0{}_0)^2 \geq 1$. From this we can conclude

$$L^0{}_0 \geq 1 \quad \text{or} \qquad L^0{}_0 \leq -1 \,, \tag{10.25}$$

i.e., there exist two other distinct classes. Properties (10.23) and (10.25) can be stated as follows: The set of all Lorentz transformations $\mathcal{L}$ is given as the union

$$\mathcal{L} = \mathcal{L}^{\uparrow}_{+} \cup \mathcal{L}^{\downarrow}_{+} \cup \mathcal{L}^{\uparrow}_{-} \cup \mathcal{L}^{\downarrow}_{-} \tag{10.26}$$

where $\mathcal{L}^{\uparrow}_{+}$, $\mathcal{L}^{\downarrow}_{+}$, $\mathcal{L}^{\uparrow}_{-}$, $\mathcal{L}^{\downarrow}_{-}$ are disjunct sets defined as follows

$$\mathcal{L}^{\uparrow}_{+} = \{\, L, L \in \mathrm{O}(3,1),\, \det L = 1,\, L^0{}_0 \geq 1\} \,; \tag{10.27}$$

$$\mathcal{L}^{\downarrow}_{+} = \{\, L, L \in \mathrm{O}(3,1),\, \det L = 1,\, L^0{}_0 \leq -1\} \,; \tag{10.28}$$

$$\mathcal{L}^{\uparrow}_{-} = \{\, L, L \in \mathrm{O}(3,1),\, \det L = -1,\, L^0{}_0 \geq 1\} \,; \tag{10.29}$$

$$\mathcal{L}^{\downarrow}_{-} = \{\, L, L \in \mathrm{O}(3,1),\, \det L = -1,\, L^0{}_0 \leq -1\} \,. \tag{10.30}$$

It holds $\mathbf{g} \in \mathcal{L}$ and $-\mathbb{1} \in \mathcal{L}$ as one can readily verify testing for property (10.14). One can also verify that one can write

$$\mathcal{L}^{\uparrow}_{-} = \mathbf{g}\mathcal{L}^{\uparrow}_{+} = \mathcal{L}^{\uparrow}_{+}\mathbf{g} \,; \tag{10.31}$$

$$\mathcal{L}^{\downarrow}_{+} = -\mathcal{L}^{\uparrow}_{+} \,; \tag{10.32}$$

$$\mathcal{L}^{\downarrow}_{-} = -\,\mathbf{g}\mathcal{L}^{\uparrow}_{+} = -\,\mathcal{L}^{\uparrow}_{+}\mathbf{g} \tag{10.33}$$

where we used the definition $a\mathcal{M} = \{M_1, \exists M_2, M_2 \in \mathcal{M}, M_1 = a\, M_2\}$. The above shows that the set of proper Lorentz transformations $\mathcal{L}^{\uparrow}_{+}$ allows one to generate all Lorentz transformations, except for the trivial factors $\mathbf{g}$ and $-\mathbb{1}$. It is, hence, entirely suitable to investigate first only Lorentz transformations in $\mathcal{L}^{\uparrow}_{+}$.
We start our investigation by demonstrating that $\mathcal{L}^{\uparrow}_{+}$ forms a group. Obviously, $\mathcal{L}^{\uparrow}_{+}$ contains $\mathbb{1}$. We can also demonstrate that for $A, B \in \mathcal{L}^{\uparrow}_{+}$ holds $C = AB \in \mathcal{L}^{\uparrow}_{+}$. For this purpose we consider the value of $C^0{}_0 = A^0{}_\mu B^\mu{}_0 = \sum_{j=1}^{3} A^0{}_j B^j{}_0 + A^0{}_0 B^0{}_0$. Schwartz's inequality yields

$$\left(\sum_{j=1}^{3} A^0{}_j B^j{}_0 \right)^2 \leq \sum_{j=1}^{3} \left(A^0{}_j\right)^2 \sum_{j=1}^{3} \left(B^j{}_0\right)^2 \,. \tag{10.34}$$

From (10.12) follows $(B^0{}_0)^2 - \sum_{j=1}^3 (B^j{}_0)^2 = 1$ or $\sum_{j=1}^3 (B^j{}_0)^2 = (B^0{}_0)^2 - 1$. Similarly, one can conclude from (10.21) $\sum_{j=1}^3 (A^0{}_j)^2 = (A^0{}_0)^2 - 1$. (10.34) provides then the estimate

$$\left(\sum_{j=1}^{3} A^0{}_j B^j{}_0 \right)^2 \leq \left[(A^0{}_0)^2 - 1 \right] \left[(B^0{}_0)^2 - 1 \right] < (A^0{}_0)^2 \, (B^0{}_0)^2 \, . \tag{10.35}$$

One can conclude, therefore, $|\sum_{j=1}^3 A^0{}_j B^j{}_0| < A^0{}_0 B^0{}_0$. Since $A^0{}_0 \geq 1$ and $B^0{}_0 \geq 1$, obviously $A^0{}_0 B^0{}_0 \geq 1$. Using the above expression for $C^0{}_0$ one can state $C^0{}_0 > 0$. In fact, since the group property of O(3,1) ascertains $\mathbf{C}^T \mathbf{g} \mathbf{C} = \mathbf{g}$ it must hold $C^0{}_0 \geq 1$.
The next group property of $\mathcal{L}_+^\uparrow$ to be demonstrated is the existence of the inverse. For the inverse of any $\mathbf{L} \in \mathcal{L}_+^\uparrow$ holds (10.16). This relationship shows $(\mathbf{L}^{-1})^0{}_0 = L^0{}_0$, from which one can conclude $\mathbf{L}^{-1} \in \mathcal{L}_+^\uparrow$. We also note that the identity operator $\mathbb{1}$ has elements

$$\mathbb{1}^\mu{}_\nu = \delta^\mu{}_\nu \tag{10.36}$$

where we defined[1]

$$\delta^\mu{}_\nu = \begin{cases} 1 & \text{for } \mu = \nu \\ 0 & \text{for } \mu \neq \nu \end{cases} \tag{10.37}$$

It holds, $\mathbb{1}^0{}_0 = \geq 1$ and, hence, $\mathbb{1} \in \mathcal{L}_+^\uparrow$. Since the associative property holds for matrix multiplication we have verified that $\mathcal{L}_+^\uparrow$ is indeed a subgroup of SO(3,1).
$\mathcal{L}_+^\uparrow$ is called the subgroup of *proper, orthochronous Lorentz transformations*. In the following we will consider solely this subgroup of SO(3,1).

Infinitesimal Lorentz transformations

The transformations in $\mathcal{L}_+^\uparrow$ have the property that they are continously connected to the identity $\mathbb{1}$, i.e., these transformations can be parametrized such that a continuous variation of the parameters connects any element of $\mathcal{L}_+^\uparrow$ with $\mathbb{1}$. This property will be exploited now in that we consider first transformations in a small neighborhood of $\mathbb{1}$ which we parametrize by infinitesimal parameters. We will then employ the Lie group properties to generate all transformations in $\mathcal{L}_+^\uparrow$.
Accordingly, we consider transformations

$$L^\mu{}_\nu = \delta^\mu{}_\nu + \epsilon^\mu{}_\nu \, ; \quad \epsilon^\mu{}_\nu \quad \text{small} \, . \tag{10.38}$$

For these transformations, obviously, holds $L^0{}_0 > 0$ and the value of the determinant is close to unity, i.e., if we enforce (10.14) actually $L^0{}_0 \geq 1$ and $\det \mathbf{L} = 1$ must hold. Property (10.14) implies

$$\left(\mathbb{1} + \epsilon^T \right) \mathbf{g} \left(\mathbb{1} + \epsilon \right) = \mathbf{g} \tag{10.39}$$

where we have employed the matrix form ϵ defined as in (10.13). To order $O(\epsilon^2)$ holds

$$\epsilon^T \mathbf{g} + \mathbf{g} \epsilon = 0 \, . \tag{10.40}$$

[1]It should be noted that according to our present definition holds $\delta_{\mu\nu} = g_{\mu\rho} \delta^\rho{}_\nu$ and, accordingly, $\delta_{00} = 1$ and $\delta_{11} = \delta_{22} = \delta_{33} = -1$.

Using (10.15) one can conclude

$$\epsilon^T \;=\; -\,\mathbf{g}\,\epsilon\,\mathbf{g} \tag{10.41}$$

which reads explicitly

$$\begin{pmatrix} \epsilon^0{}_0 & \epsilon^1{}_0 & \epsilon^2{}_0 & \epsilon^3{}_0 \\ \epsilon^0{}_1 & \epsilon^1{}_1 & \epsilon^2{}_1 & \epsilon^3{}_1 \\ \epsilon^0{}_2 & \epsilon^1{}_2 & \epsilon^2{}_2 & \epsilon^3{}_2 \\ \epsilon^0{}_3 & \epsilon^1{}_3 & \epsilon^2{}_3 & \epsilon^3{}_3 \end{pmatrix} = \begin{pmatrix} -\epsilon^0{}_0 & \epsilon^0{}_1 & \epsilon^0{}_2 & \epsilon^0{}_3 \\ \epsilon^1{}_0 & -\epsilon^1{}_1 & -\epsilon^1{}_2 & -\epsilon^1{}_3 \\ \epsilon^2{}_0 & -\epsilon^2{}_1 & -\epsilon^2{}_2 & -\epsilon^2{}_3 \\ \epsilon^3{}_0 & -\epsilon^3{}_1 & -\epsilon^3{}_2 & -\epsilon^3{}_3 \end{pmatrix} . \tag{10.42}$$

This relationship implies

$$\begin{aligned} \epsilon^\mu{}_\mu &= 0 \\ \epsilon^0{}_j &= \epsilon^j{}_0 \,, \qquad j \,=\, 1,2,3 \\ \epsilon^j{}_k &= -\,\epsilon^k{}_j \,, \quad j,k \,=\, 1,2,3 \end{aligned} \tag{10.43}$$

Inspection shows that the matrix ϵ has 6 independent elements and can be written

$$\epsilon(\vartheta_1,\vartheta_2,\vartheta_3,w_1,w_2,w_3) \;=\; \begin{pmatrix} 0 & -w_1 & -w_2 & -w_3 \\ -w_1 & 0 & -\vartheta_3 & \vartheta_2 \\ -w_2 & \vartheta_3 & 0 & -\vartheta_1 \\ -w_3 & -\vartheta_2 & \vartheta_1 & 0 \end{pmatrix} . \tag{10.44}$$

This result allows us now to define six generators for the Lorentz transformations($k \,=\, 1,2,3$)

$$\mathbf{J}_k \;=\; \epsilon(\vartheta_k = 1, \text{other five parameters zero}) \tag{10.45}$$

$$\mathbf{K}_k \;=\; \epsilon(w_k = 1, \text{other five parameters zero}) \,. \tag{10.46}$$

The generators are explicitly

$$\mathbf{J}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} ; \; \mathbf{J}_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} ; \; \mathbf{J}_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{10.47}$$

$$\mathbf{K}_1 = \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} ; \; \mathbf{K}_2 = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} ; \; \mathbf{K}_3 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \tag{10.48}$$

These commutators obey the following commutation relationships

$$\begin{aligned} [\,\mathbf{J}_k,\,\mathbf{J}_\ell\,] &= \epsilon_{k\ell m}\,\mathbf{J}_m \\ [\,\mathbf{K}_k,\,\mathbf{K}_\ell\,] &= -\,\epsilon_{k\ell m}\,\mathbf{J}_m \\ [\,\mathbf{J}_k,\,\mathbf{K}_\ell\,] &= \epsilon_{k\ell m}\,\mathbf{K}_m \,. \end{aligned} \tag{10.49}$$

The operators also obey

$$\vec{\mathbf{J}}\cdot\vec{\mathbf{K}} \;=\; \mathbf{J}_1\mathbf{J}_1 \,+\, \mathbf{J}_2\mathbf{J}_2 \,+\, \mathbf{J}_3\mathbf{J}_3 \;=\; 0 \tag{10.50}$$

as can be readily verified.

Exercise 7.1:
Demonstrate the commutation relationships (10.49, 10.50).

The commutation relationships (10.49) define the Lie algebra associated with the Lie group $\mathcal{L}_+^\uparrow$. The commutation relationships imply that the algebra of the generators $\mathbf{J}_k$, $\mathbf{K}_k$, $k = 1, 2, 3$ is closed. Following the treatment of the rotation group SO(3) one can express the elements of $\mathcal{L}_+^\uparrow$ through the exponential operators

$$\mathbf{L}(\vec{\vartheta},\, \vec{w}) \;=\; \exp\left(\vec{\vartheta} \cdot \vec{\mathbf{J}} \;+\; \vec{w} \cdot \vec{\mathbf{K}} \right) \quad ; \quad \vec{\vartheta}, \vec{w} \;\in\; \mathbb{R}^3 \tag{10.51}$$

where we have defined $\vec{\vartheta} \cdot \vec{\mathbf{J}} \;=\; \sum_{k=1}^{3} \vartheta_k \mathbf{J}_k$ and $\vec{w} \cdot \vec{\mathbf{K}} \;=\; \sum_{k=1}^{3} w_k \mathbf{K}_k$. One can readily show, following the algebra in Chapter 5, and using the relationship

$$\mathbf{J}_k \;=\; \begin{pmatrix} 0 & 0 \\ 0 & L_k \end{pmatrix} \tag{10.52}$$

where the 3×3–matrices L_k are the generators of SO(3) defined in Chapter 5, that the transformations (10.51) for $\vec{w} \;=\; 0$ correspond to rotations of the spatial coordinates, i.e.,

$$\mathbf{L}(\vec{\vartheta},\, \vec{w} \;=\; 0) \;=\; \begin{pmatrix} 0 & 0 \\ 0 & R(\vec{\vartheta}) \end{pmatrix} . \tag{10.53}$$

Here $R(\vec{\vartheta})$ are the 3×3–rotation matrices constructed in Chapter 5. For the parameters ϑ_k of the Lorentz transformations holds obviously

$$\vartheta_k \;\in\; [0, 2\pi[\quad , \; k = 1, 2, 3 \tag{10.54}$$

which, however, constitutes an overcomplete parametrization of the rotations (see Chapter 5).
We consider now the Lorentz transformations for $\vec{\vartheta} \;=\; 0$ which are referred to as 'boosts'. A boost in the x_1–direction is $L \;=\; \exp(w_1 \mathbf{K}_1)$. To determine the explicit form of this transformation we evaluate the exponential operator by Taylor expansion. In analogy to equation (5.35) it issufficient to consider in the present case the 2×2–matrix

$$L' \;=\; \exp\left(w_1 \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \right) \;=\; \sum_{n=0}^{\infty} \frac{w_1^n}{n!} \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}^n \tag{10.55}$$

since

$$\exp\left(w_1 \mathbf{K}_1\right) \;=\; \exp \begin{pmatrix} & & 0 & 0 \\ & L' & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{10.56}$$

Using the idempotence property

$$\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}^2 \;=\; \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \;=\; \mathbb{1} \tag{10.57}$$

one can carry out the Taylor expansion above:

$$\begin{aligned} L' &= \sum_{n=0}^{\infty} \frac{w_1^{2n}}{(2n)!} \; \mathbb{1} \; + \; \sum_{n=0}^{\infty} \frac{w_1^{2n+1}}{(2n+1)!} \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \\ &= \cosh w_1 \, \mathbb{1} \; + \; \sinh w_1 \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} \cosh w_1 & -\sinh w_1 \\ -\sinh w_1 & \cosh w_1 \end{pmatrix} . \end{aligned} \tag{10.58}$$

The conventional form (10.1) of the Lorentz transformations is obtained through the parameter change

$$v_1 \; = \; \frac{\sinh w_1}{\cosh w_1} \; = \; \tanh w_1 \tag{10.59}$$

Using $\cosh^2 w_1 - \sinh^2 w_1 = 1$ one can identify $\sinh w_1 = \sqrt{\cosh^2 w_1 - 1}$ and $\cosh w_1 = \sqrt{\sinh^2 w_1 + 1}$. Correspondingly, one obtains from (10.59)

$$v_1 \; = \; \frac{\sqrt{\cosh^2 w_1 - 1}}{\cosh w_1} \; = \; \frac{\sinh w_1}{\sqrt{\sinh^2 w_1 + 1}} \, . \tag{10.60}$$

These two equations yield

$$\cosh w_1 \; = \; 1/\sqrt{1 - v_1^2} \, ; \quad \sinh w_1 \; = \; v_1/\sqrt{1 - v_1^2} \, , \tag{10.61}$$

and (10.56, 10.59) can be written

$$\exp\left(w_1 \mathbf{K}_1\right) \; = \; \begin{pmatrix} \frac{1}{\sqrt{1-v_1^2}} & \frac{-v_1}{\sqrt{1-v_1^2}} & 0 & 0 \\ \frac{-v_1}{\sqrt{1-v_1^2}} & \frac{1}{\sqrt{1-v_1^2}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{10.62}$$

According to (10.3, 10.6, 10.51) the explicit transformation for space–time–coordinates is then

$$x_1' \; = \; \frac{x_1 \, - \, v_1 t}{\sqrt{1 \, - \, v_1^2}} \, , \; t' \; = \; \frac{t \, - \, v_1 x_1}{\sqrt{1 \, - \, v_1^2}} \, , \; x_2' \; = \; x_2 \, , \; x_3' \; = \; x_3 \tag{10.63}$$

which agrees with (10.1).
The range of the parameters w_k can now be specified. v_k defined in (10.59) for the case $k = 1$ corresponds to the relative velocity of two frames of reference. We expect that v_k can only assume values less than the velocity of light c which in the present units is $c = 1$. Accordingly, we can state $v_k \in]-1, 1[$. This property is, in fact, consistent with (10.59). From (10.59) follows, however, for w_k

$$w_k \; \in]-\infty, \infty[\, . \tag{10.64}$$

We note that the range of w_k-values is not a compact set even though the range of v_k-values is compact. This property of the w_k-values contrasts with the property of the parameters ϑ_k specifying rotational angles which assume only values in a compact range.

10.2 Scalars, 4–Vectors and Tensors

In this Section we define quantities according to their behaviour under Lorentz transformations. Such quantities appear in the description of physical systems and statements about transformation properties are often extremely helpful and usually provide important physical insight. We have encountered examples in connection with rotational transformations, namely, scalars like $r = \sqrt{x_1^2 + x_2^2 + x_3^2}$, vectors like $\vec{r} = (x_1, x_2, x_3)^T$, spherical harmonics $Y_{\ell m}(\hat{r})$, total angular momentum states of composite systems like $\mathcal{Y}_{\ell m}(\ell_1, \ell_2 | \hat{r}_1, \hat{r}_2)$ and, finally, tensor operators T_{km}. Some of these quantities were actually defined with respect to representations of the rotation group in function spaces, not in the so-called natural representation associated with the 3–dimensional Euclidean space $\mathbb{E}^3$.
Presently, we have not yet defined representations of Lorentz transformations beyond the 'natural' representation acting in the 4–dimensional space of position– and time–coordinates. Hence, our definition of quantities with special properties under Lorentz transformations presently is confined to the natural representation. Nevertheless, we will encounter an impressive example of physical properties.

Scalars The quantities with the simplest transformation behaviour are so-called scalars $f \in \mathbb{R}$ which are invariant under transformations, i.e.,

$$f' \;=\; f \;. \tag{10.65}$$

An example is s^2 defined in (10.4), another example is the rest mass m of a particle. However, not any physical property $f \in \mathbb{R}$ is a scalar. Counterexamples are the energy, the charge density, the z–component x_3 of a particle, the square of the electric field $|\vec{E}(\vec{r}, t)|^2$ or the scalar product $\vec{r}_1 \cdot \vec{r}_2$ of two particle positions. We will see below how true scalars under Lorentz transformations can be constructed.

4-Vectors The quantities with the transformation behaviour like that of the position–time vector x^μ defined in (10.3) are the so-called 4–vectors a^μ. These quantites always come as four components $(a^0, a^1, a^2, a^3)^T$ and transform according to

$$a'^\mu \;=\; L^\mu{}_\nu a^\nu \;. \tag{10.66}$$

Examples of 4-vectors beside x^μ are the momentum 4-vector

$$p^\mu \;=\; (E,\, \vec{p}) \;,\; E \;=\; \frac{m}{\sqrt{1 \,-\, \vec{v}^{\,2}}} \;,\; \vec{p} \,=\, \frac{m\,\vec{v}}{\sqrt{1 \,-\, \vec{v}^{\,2}}} \tag{10.67}$$

the transformation behaviour of which we will demonstrate further below. A third 4-vector is the so-called current vector

$$J^\mu \;=\; (\rho,\, \vec{J}) \tag{10.68}$$

where $\rho(\vec{r}, t)$ and $\vec{J}(\vec{r}, t)$ are the charge density and the current density, respectively, of a system of charges. Another example is the potential 4-vector

$$A^\mu \;=\; (V,\, \vec{A}) \tag{10.69}$$

where $V(\vec{r}, t)$ and $\vec{A}(\vec{r}, t)$ are the electrical and the vector potential of an electromagnetic field. The 4-vector character of J^μ and of A^μ will be demonstrated further below.

Scalar Product 4-vectors allow one to construct scalar quantities. If a^μ and b^μ are 4-vectors then

$$a^\mu g_{\mu\nu} b^\nu \tag{10.70}$$

is a scalar. This property follows from (10.66) together with (10.12)

$$a'^\mu g_{\mu\nu} b'^\nu \;=\; L^\mu{}_\rho g_{\mu\nu} L^\nu{}_\sigma a^\rho b^\sigma \;=\; a^\rho g_{\rho\sigma} b^\sigma \tag{10.71}$$

Contravariant and Covariant 4-Vectors It is convenient to define a second class of 4-vectors. The respective vectors a_μ are associated with the 4-vectors a^μ, the relationship being

$$a_\mu \;=\; g_{\mu\nu} a^\nu \;=\; (a^0, -a^1, -a^2, -a^3) \tag{10.72}$$

where a^ν is a vector with transformation behaviour as stated in (10.66). One calls 4-vectors a_μ *covariant* and 4-vectors a^μ *contravariant*. Covariant 4-vectors transform like

$$a'_\mu \;=\; g_{\mu\nu} L^\nu{}_\rho g^{\rho\sigma} a_\sigma \tag{10.73}$$

where we defined

$$g^{\mu\nu} \;=\; g_{\mu\nu}\,. \tag{10.74}$$

We like to point out that from definition (10.72) of the covariant 4-vector follows $a^\mu = g^{\mu\nu} a_\nu$. In fact, one can employ the tensors $g^{\mu\nu}$ and $g_{\mu\nu}$ to raise and lower indices of $L^\mu{}_\nu$ as well. We do not establish here the consistency of the ensuing notation. In any case one can express (10.73)

$$a'_\mu \;=\; L_\mu{}^\sigma a_\sigma\,. \tag{10.75}$$

Note that according to (10.17) $L_\mu{}^\sigma$ is the transformation inverse to $L^\sigma{}_\mu$. In fact, one can express $[(\mathbf{L}^{-1})^T]^\mu{}_\nu = (L^{-1})^\nu{}_\mu$ and, accordingly, (10.17) can be written

$$(L^{-1})^\nu{}_\mu \;=\; L_\mu{}^\nu\,. \tag{10.76}$$

The 4-Vector ∂_μ An important example of a covariant 4-vector is the differential operator

$$\partial_\mu \;=\; \frac{\partial}{\partial x^\mu} \;=\; \left(\frac{\partial}{\partial t}, \nabla\right) \tag{10.77}$$

The transformed differential operator will be denoted by

$$\partial'_\mu \;\stackrel{\text{def}}{=}\; \frac{\partial}{\partial x'^\mu}\,. \tag{10.78}$$

To prove the 4-vector property of ∂_μ we will show that $g^{\mu\nu}\partial_\nu$ transforms like a contravariant 4-vector, i.e., $g^{\mu\nu}\partial'_\nu = L^\mu{}_\rho g^{\rho\sigma}\partial_\sigma$. We start from $x'^\mu = L^\mu{}_\nu x^\nu$. Multiplication (and summation) of $x'^\mu = L^\mu{}_\nu x^\nu$ by $L^\rho{}_\sigma g_{\rho\mu}$ yields, using (10.12), $g_{\sigma\nu} x^\nu = L^\rho{}_\sigma g_{\rho\mu} x'^\mu$ and $g^{\mu\sigma} g_{\sigma\nu} = \delta^\mu{}_\nu$,

$$x^\nu \;=\; g^{\nu\sigma} L^\rho{}_\sigma g_{\rho\mu} x'^\mu\,. \tag{10.79}$$

This is the inverse Lorentz transformation consistent with (10.16). We have duplicated the expression for the inverse of $L^\mu{}_\nu$ to obtain the correct notation in terms of covariant, i.e., lower, and

contravariant, i.e., upper, indices. (10.79) allows us to determine the connection between ∂_μ and ∂'_μ. Using the chain rule of differential calculus we obtain

$$\partial'_\mu \;=\; \sum_{\nu=0}^{3} \frac{\partial x^\nu}{\partial x'^\mu} \frac{\partial}{\partial x^\nu} \;=\; g^{\nu\sigma} L^\rho{}_\sigma g_{\rho\mu} \partial_\nu \;=\; L_\mu{}^\nu \, \partial_\nu \tag{10.80}$$

Multiplication by $g^{\lambda\mu}$ (and summation over μ) together with $g^{\lambda\mu} g_{\rho\mu} = \delta^\lambda{}_\rho$ yields

$$g^{\lambda\mu} \partial'_\mu \;=\; L^\lambda{}_\sigma g^{\sigma\nu} \partial_\nu \;, \tag{10.81}$$

i.e., ∂_μ does indeed transform like a covariant vector.

d'Alembert Operator We want to construct now a scalar differential operator. For this purpose we define first the contravariant differential operator

$$\partial^\mu \;=\; g^{\mu\nu} \partial_\nu \;=\; \left(\frac{\partial}{\partial t}, -\nabla \right) \;. \tag{10.82}$$

Then the operator

$$\partial_\mu \partial^\mu \;=\; \partial_t^2 \;-\; \nabla^2 \tag{10.83}$$

is a scalar under Lorentz transformations. In fact, this operator is equal to the d'Alembert operator which is known to be Lorentz-invariant.

Proof that p^μ is a 4-vector We will demonstrate now that the momentum 4-vector p^μ defined in (10.67) transforms like (10.66). For this purpose we consider the scalar differential

$$(d\tau)^2 \;=\; dx^\mu dx_\mu \;=\; (dt)^2 \;-\; (d\vec{r})^2 \tag{10.84}$$

It holds

$$\left(\frac{d\tau}{dt} \right)^2 \;=\; 1 \;-\; (\vec{v})^2 \tag{10.85}$$

from which follows

$$\frac{d}{d\tau} \;=\; \frac{1}{\sqrt{1 \;-\; \vec{v}^2}} \frac{d}{dt} \;. \tag{10.86}$$

One can write

$$p^0 \;=\; E \;=\; \frac{m}{\sqrt{1 \;-\; \vec{v}^2}} \;=\; \frac{m}{\sqrt{1 \;-\; \vec{v}^2}} \frac{dt}{dt} \;. \tag{10.87}$$

The remaining components of p^μ can be written, e.g.,

$$p^1 \;=\; \frac{m\, v^1}{\sqrt{1 \;-\; \vec{v}^2}} \;=\; \frac{m}{\sqrt{1 \;-\; \vec{v}^2}} \frac{dx^1}{dt} \;. \tag{10.88}$$

One can express then the momentum vector

$$p^\mu \;=\; \frac{m}{\sqrt{1 \;-\; \vec{v}^2}} \frac{dx^\mu}{dt} \;=\; m \frac{d}{d\tau} x^\mu \;. \tag{10.89}$$

The operator $m\frac{d}{d\tau}$ transforms like a scalar. Since x^μ transforms like a contravariant 4-vector, the r.h.s. of (10.89) alltogether transforms like a contravariant 4-vector, and, hence, p^μ on the l.h.s. of (10.89) must be a 4-vector.
The momentum 4-vector allows us to construct a scalar quantity, namely

$$p^\mu p_\mu \;=\; p^\mu g_{\mu\nu} p^\nu \;=\; E^2 \;-\; \vec{p}^{\,2} \tag{10.90}$$

Evaluation of the r.h.s. yields according to (10.67)

$$E^2 \;-\; \vec{p}^{\,2} \;=\; \frac{m^2}{1 \;-\; \vec{v}^{\,2}} \;-\; \frac{m^2 \vec{v}^{\,2}}{1 \;-\; \vec{v}^{\,2}} \;=\; m^2 \tag{10.91}$$

or

$$p^\mu p_\mu \;=\; m^2 \tag{10.92}$$

which, in fact, is a scalar. We like to rewrite the last result

$$E^2 \;=\; \vec{p}^{\,2} \;+\; m^2 \tag{10.93}$$

or

$$E \;=\; \pm\sqrt{\vec{p}^{\,2} \;+\; m^2}\,. \tag{10.94}$$

In the non-relativistic limit the rest energy m is the dominant contribution to E. Expansion in $\frac{1}{m}$ should then be rapidly convergent. One obtains

$$E \;=\; \pm m \;\pm\; \frac{\vec{p}^{\,2}}{2m} \mp \frac{(\vec{p}^{\,2})^2}{4m^3} \;+\; O\left(\frac{(\vec{p}^{\,2})^3}{4m^5}\right)\,. \tag{10.95}$$

This obviously describes the energy of a free particle with rest energy $\pm m$, kinetic energy $\pm\frac{\vec{p}^{\,2}}{2m}$ and relativistic corrections.

10.3 Relativistic Electrodynamics

In the following we summarize the Lorentz-invariant formulation of electrodynamics and demonstrate its connection to the conventional formulation as provided in Sect. 8.

Lorentz Gauge In our previous description of the electrodynamic field we had introduced the scalar and vector potential $V(\vec{r},t)$ and $\vec{A}(\vec{r},t)$, respectively, and had chosen the so-called Coulomb gauge (8.12), i.e., $\nabla\cdot\vec{A} = 0$, for these potentials. This gauge is not Lorentz-invariant and we will adopt here another gauge, namely,

$$\partial_t V(\vec{r},t) \;+\; \nabla\cdot\vec{A}(\vec{r},t) \;=\; 0\,. \tag{10.96}$$

The Lorentz-invariance of this gauge, the so-called *Lorentz gauge*, can be demonstrated readily using the 4-vector notation (10.69) for the electrodynamic potential and the 4-vector derivative (10.77) which allow one to express (10.96) in the form

$$\partial_\mu A^\mu \;=\; 0\,. \tag{10.97}$$

We have proven already that ∂_μ is a contravariant 4-vector. If we can show that A^μ defined in (10.69) is, in fact, a contravariant 4-vector then the l.h.s. in (10.97) and, equivalently, in (10.96) is a scalar and, hence, Lorentz-invariant. We will demonstrate now the 4-vector property of A^μ.

Transformation Properties of J^μ and A^μ

The charge density $\rho(\vec{r},t)$ and current density $\vec{J}(\vec{r},t)$ are known to obey the continuity property

$$\partial_t \rho(\vec{r},t) \; + \; \nabla \cdot \vec{J}(\vec{r},t) \; = \; 0 \tag{10.98}$$

which reflects the principle of charge conservation. This principle should hold in any frame of reference. Equation (10.98) can be written, using (10.77) and (10.68),

$$\partial_\mu J^\mu(x^\mu) \; = \; 0 \; . \tag{10.99}$$

Since this equation must be true in any frame of reference the right hand side must vanish in all frames, i.e., must be a scalar. Consequently, also the l.h.s. of (10.99) must be a scalar. Since ∂_μ transforms like a covariant 4-vector, it follows that J^μ, in fact, has to transform like a contravariant 4-vector.

We want to derive now the differential equations which determine the 4-potential A^μ in the Lorentz gauge (10.97) and, thereby, prove that A^μ is, in fact, a 4-vector. The respective equation for $A^0 \; = \; V$ can be obtained from Eq. (8.13). Using $\nabla \cdot \partial_t \vec{A}(\vec{r},t) \; = \; \partial_t \nabla \cdot \vec{A}(\vec{r},t)$ together with (10.96), i.e., $\nabla \cdot \vec{A}(\vec{r},t) \; = \; -\partial_t V(\vec{r},t)$, one obtains

$$\partial_t^2 V(\vec{r},t) \; - \; \nabla^2 V(\vec{r},t) \; = \; 4\pi\rho(\vec{r},t) \; . \tag{10.100}$$

Similarly, one obtains for $\vec{A}(\vec{r},t)$ from (8.17) using the identity (8.18) and, according to (10.96), $\nabla \cdot \vec{A}(\vec{r},t) \; = \; -\partial_t V(\vec{r},t)$

$$\partial_t^2 \vec{A}(\vec{r},t) \; - \; \nabla^2 \vec{A}(\vec{r},t) \; = \; 4\,\pi\,\vec{J}(\vec{r},t) \; . \tag{10.101}$$

Combining equations (10.100, 10.101), using (10.83) and (10.69), yields

$$\partial_\mu \partial^\mu \, A^\nu(x^\sigma) \; = \; 4\,\pi\,J^\nu(x^\sigma) \; . \tag{10.102}$$

In this equation the r.h.s. transforms like a 4-vector. The l.h.s. must transform likewise. Since $\partial_\mu \partial^\mu$ transforms like a scalar one can conclude that $A^\nu(x^\sigma)$ must transform like a 4-vector.

The Field Tensor

The electric and magnetic fields can be collected into an anti-symmetric 4×4 tensor

$$F^{\mu\nu} \; = \; \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix} \; . \tag{10.103}$$

Alternatively, this can be stated

$$F^{k0} \; = \; -F^{0k} \; = \; E^k \; , \quad F^{mn} \; = \; -\epsilon^{mn\ell} B^\ell \; , \quad k,\ell,m,n \; = \; 1,2,3 \tag{10.104}$$

where $\epsilon^{mn\ell} \; = \; \epsilon_{mn\ell}$ is the totally anti-symmetric three-dimensional tensor defined in (5.32).

One can readily verify, using (8.6) and (8.9), that $F^{\mu\nu}$ can be expressed through the potential A^μ in (10.69) and ∂^μ in (10.82) as follows

$$F^{\mu\nu} \; = \; \partial^{\,\mu} A^\nu \; - \; \partial^{\,\nu} A^\mu \; . \tag{10.105}$$

The relationships (10.103, 10.104) establishe the transformation behaviour of $\vec{E}(\vec{r},t)$ and $\vec{B}(\vec{r},t)$. In a new frame of reference holds

$$F'^{\mu\nu} \;=\; L^{\mu}{}_{\alpha}\, L^{\nu}{}_{\beta}\, F^{\alpha\beta} \tag{10.106}$$

In case that the Lorentz transformation $L^{\mu}{}_{\nu}$ is given by (10.62) or, equivalently, by (10.63), one obtains

$$F'^{\mu\nu} \;=\; \begin{pmatrix} 0 & -E_x & -\frac{E_y - v_1 B_z}{\sqrt{1-v_1^2}} & -\frac{E_z + v_1 B_y}{\sqrt{1-v_1^2}} \\ E_x & 0 & -\frac{B_z - v_1 E_y}{\sqrt{1-v_1^2}} & \frac{B_y + v_1 E_z}{\sqrt{1-v_1^2}} \\ \frac{E_y - v_1 B_z}{\sqrt{1-v_1^2}} & \frac{B_z - v_1 E_y}{\sqrt{1-v_1^2}} & 0 & -B_x \\ \frac{E_z + v_1 B_y}{\sqrt{1-v_1^2}} & -\frac{B_y + v_1 E_z}{\sqrt{1-v_1^2}} & B_x & 0 \end{pmatrix} \tag{10.107}$$

Comparision with

$$F'^{\mu\nu} \;=\; \begin{pmatrix} 0 & -E'_x & -E'_y & -E'_z \\ E'_x & 0 & -B'_z & B'_y \\ E'_y & B'_z & 0 & -B'_x \\ E'_z & -B'_y & B'_x & 0 \end{pmatrix} \tag{10.108}$$

yields then the expressions for the transformed fields $\vec{E}'$ and $\vec{B}'$. The results can be put into the more general form

$$\vec{E}'_{\parallel} \;=\; \vec{E}_{\parallel}\,, \quad \vec{E}'_{\perp} \;=\; \frac{\vec{E}_{\perp} \,+\, \vec{v}\times\vec{B}}{\sqrt{1 \,-\, \vec{v}^{\,2}}} \tag{10.109}$$

$$\vec{B}'_{\parallel} \;=\; \vec{B}_{\parallel}\,, \quad \vec{B}'_{\perp} \;=\; \frac{\vec{B}_{\perp} \,-\, \vec{v}\times\vec{E}}{\sqrt{1 \,-\, \vec{v}^{\,2}}} \tag{10.110}$$

where $\vec{E}_{\parallel}$, $\vec{B}_{\parallel}$ and $\vec{E}_{\perp}$, $\vec{B}_{\perp}$ are, respectively, the components of the fields parallel and perpendicular to the velocity $\vec{v}$ which determines the Lorentz transformation. These equations show that under Lorentz transformations electric and magnetic fields convert into one another.

Maxwell Equations in Lorentz-Invariant Form

One can express the Maxwell equations in terms of the tensor $F^{\mu\nu}$ in Lorentz-invariant form. Noting

$$\partial_\mu F^{\mu\nu} \;=\; \partial_\mu \partial^\mu A^\nu \;-\; \partial_\mu \partial^\nu A^\mu \;=\; \partial_\mu \partial^\mu A^\nu \;-\; \partial^\nu \partial_\mu A^\mu \;=\; \partial_\mu \partial^\mu A^\nu\,, \tag{10.111}$$

where we used (10.105) and (10.97), one can conclude from (10.102)

$$\partial_\mu F^{\mu\nu} \;=\; 4\pi\, J^\nu\,. \tag{10.112}$$

One can readily prove that this equation is equivalent to the two inhomogeneous Maxwell equations (8.1, 8.2). From the definition (10.105) of the tensor $F^{\mu\nu}$ one can conclude the property

$$\partial^\sigma F^{\mu\nu} \;+\; \partial^\mu F^{\nu\sigma} \;+\; \partial^\nu F^{\sigma\mu} \;=\; 0 \tag{10.113}$$

which can be shown to be equivalent to the two homogeneous Maxwell equations (8.3, 8.4).

Lorentz Force

One important property of the electromagnetic field is the Lorentz force acting on charged particles moving through the field. We want to express this force through the tensor $F^{\mu\nu}$. It holds for a particle with 4-momentum p^μ as given by (10.67) and charge q

$$\frac{dp^\mu}{d\tau} \;=\; \frac{q}{m}\, p_\nu\, F^{\mu\nu} \tag{10.114}$$

where $d/d\tau$ is given by (10.86). We want to demonstrate now that this equation is equivalent to the equation of motion (8.5) where $\vec{p} = m\vec{v}/\sqrt{1-v^2}$.
To avoid confusion we will employ in the following for the energy of the particle the notation $\mathcal{E} = m/\sqrt{1-v^2}$ [see (10.87)] and retain the definition $\vec{E}$ for the electric field. The $\mu = 0$ component of (10.114) reads then, using (10.104),

$$\frac{d\mathcal{E}}{d\tau} \;=\; \frac{q}{m}\,\vec{p}\cdot\vec{E} \tag{10.115}$$

or with (10.86)

$$\frac{d\mathcal{E}}{dt} \;=\; \frac{q}{\mathcal{E}}\,\vec{p}\cdot\vec{E}\,. \tag{10.116}$$

From this one can conclude, employing (10.93),

$$\frac{1}{2}\frac{d\mathcal{E}^2}{dt} \;=\; \frac{1}{2}\frac{d\vec{p}^{\,2}}{dt} \;=\; q\,\vec{p}\cdot\vec{E} \tag{10.117}$$

This equation follows, however, also from the equation of motion (8.5) taking the scalar product with $\vec{p}$

$$\vec{p}\cdot\frac{d\vec{p}}{dt} \;=\; q\vec{p}\cdot\vec{E} \tag{10.118}$$

where we exploited the fact that according to $\vec{p} = m\vec{v}/\sqrt{1-v^2}$ holds $\vec{p} \parallel \vec{v}$.
For the spatial components, e.g., for $\mu = 1$, (10.114) reads using (10.103)

$$\frac{dp_x}{d\tau} \;=\; \frac{q}{m}\,(\,\mathcal{E}E_x \;+\; p_yB_z \;-\; p_zB_y\,)\,. \tag{10.119}$$

Employing again (10.86) and (10.67), i.e., $\mathcal{E} = m/\sqrt{1-v^2}$, yields

$$\frac{dp_x}{dt} \;=\; q\,\Big[\,E_x \;+\; (\vec{v}\times\vec{B})_x\,\Big] \tag{10.120}$$

which is the x-component of the equation of motion (8.5). We have, hence, demonstrated that (10.114) is, in fact, equivalent to (8.5). The term on the r.h.s. of (10.120) is referred to as the Lorentz force. Equation (10.114), hence, provides an alternative description of the action of the Lorentz force.

10.4 Function Space Representation of Lorentz Group

In the following it will be required to decribe the transformation of wave functions under Lorentz transformations. In this section we will investigate the transformation properties of *scalar* functions $\psi(x^\mu)$, $\psi \in \mathbb{C}_\infty(4)$. For such functions holds in the transformed frame

$$\psi'(L^\mu{}_\nu x^\nu) \;=\; \psi(x^\mu) \tag{10.121}$$

which states that the function values $\psi'(x'^\mu)$ at each point x'^μ in the new frame are identical to the function values $\psi(x^\mu)$ in the old frame *taken at the same space–time point* x^μ, i.e., taken at the pairs of points $(x'^\mu = L^\mu{}_\nu x^\nu, x^\mu)$. We need to emphasize that (10.121) covers solely the transformation behaviour of *scalar* functions. Functions which represent 4-vectorsor other non-scalar entities, e.g., the charge-current density in case of Sect. 10.3 or the bi-spinor wave function of electron-positron pairs in Sect. 10.7, obey a different transformation law.
We like to express now $\psi'(x'^\mu)$ in terms of the old coordinates x^μ. For this purpose one replaces x^μ in (10.121) by $(L^{-1})^\mu{}_\nu x^\nu$ and obtains

$$\psi'(x^\mu) \;=\; \psi((L^{-1})^\mu{}_\nu x^\nu)\;. \tag{10.122}$$

This result gives rise to the definition of the function space representation $\rho(L^\mu{}_\nu)$ of the Lorentz group

$$(\rho(L^\mu{}_\nu)\psi)(x^\mu) \;\stackrel{\text{def}}{=}\; \psi((L^{-1})^\mu{}_\nu x^\nu)\;. \tag{10.123}$$

This definition corresponds closely to the function space representation (5.42) of SO(3). In analogy to the situation for SO(3) we seek an expression for $\rho(L^\mu{}_\nu)$ in terms of an exponential operator and transformation parameters $\vec{\vartheta}$, $\vec{w}$, i.e., we seek an expression which corresponds to (10.51) for the natural representation of the Lorentz group. The resulting expression should be a generalization of the function space representation (5.48) of SO(3), in as far as SO(3,1) is a generalization (rotation + boosts) of the group SO(3). We will denote the intended representation by

$$\mathcal{L}(\vec{\vartheta},\,\vec{w}) \;\stackrel{\text{def}}{=}\; \rho(L^\mu{}_\nu(\vec{\vartheta},\,\vec{w})) \;=\; \rho\left(\,e^{\vec{\vartheta}\cdot\vec{J}\,+\,\vec{w}\cdot\vec{K}}\,\right) \tag{10.124}$$

which we present in the form

$$\mathcal{L}(\vec{\vartheta},\,\vec{w}) \;=\; \exp\left(\,\vec{\vartheta}\cdot\vec{\mathcal{J}} \;+\; \vec{w}\cdot\vec{\mathcal{K}}\,\right)\;. \tag{10.125}$$

In this expression $\vec{\mathcal{J}} = (\mathcal{J}_1, \mathcal{J}_2, \mathcal{J}_3)$ and $\vec{\mathcal{K}} = (\mathcal{K}_1, \mathcal{K}_2, \mathcal{K}_3)$ are the generators of $\mathcal{L}(\vec{\vartheta},\,\vec{w})$ which correspond to the generators J_k and K_k in (10.47), and which can be constructed following the procedure adopted for the function space representation of SO(3). However, in the present case we exclude the factor '$-i$' [cf. (5.48) and (10.125)]. Accordingly, one can evaluate $\mathcal{J}_k$ as follows

$$\mathcal{J}_k \;=\; \lim_{\vartheta_k\to 0}\,\frac{1}{\vartheta_1}\,\left[\,\rho\left(e^{\vartheta_k J_k}\right) \;-\; \mathbb{1}\,\right] \tag{10.126}$$

and $\mathcal{K}_k$

$$\mathcal{K}_k \;=\; \lim_{w_k\to 0}\,\frac{1}{w_1}\,\left[\,\rho\left(e^{w_k K_k}\right) \;-\; \mathbb{1}\,\right]\;. \tag{10.127}$$

One obtains

$$\begin{aligned} \mathcal{J}_1 &= x^3\partial_2 - x^2\partial_3\,; \qquad \mathcal{K}_1 = x^0\partial_1 + x^1\partial_0 \\ \mathcal{J}_2 &= x^1\partial_3 - x^3\partial_1\,; \qquad \mathcal{K}_2 = x^0\partial_2 + x^2\partial_0 \\ \mathcal{J}_3 &= x^2\partial_1 - x^1\partial_2\,; \qquad \mathcal{K}_3 = x^0\partial_3 + x^3\partial_0 \end{aligned} \tag{10.128}$$

which we like to demonstrate for $\mathcal{J}_1$ and $\mathcal{K}_1$.
In order to evaluate (10.126) for $\mathcal{J}_1$ we consider first

$$\left(e^{\vartheta_1 J_1}\right)^{-1} = e^{-\vartheta_1 J_1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\vartheta_1 & \sin\vartheta_1 \\ 0 & 0 & -\sin\vartheta_1 & \cos\vartheta_1 \end{pmatrix} \tag{10.129}$$

which yields for small ϑ_1

$$\begin{aligned} \rho\left(e^{\vartheta_1 J_1}\right)\psi(x^\mu) &= \psi(x^0, x^1, \cos\vartheta_1\, x^2 + \sin\vartheta_1\, x^3, -\sin\vartheta_1\, x^2 + \cos\vartheta_1\, x^3) \\ &= \psi(x^\mu) + \vartheta_1\,(x^3\partial_2 - x^2\partial_3)\,\psi(x^\mu) + O(\vartheta_1^2)\,. \end{aligned} \tag{10.130}$$

This result, obviously, reproduces the expression for $\mathcal{J}_1$ in (10.128).
One can determine similarly $\mathcal{K}_1$ starting from

$$\left(e^{w_1 K_1}\right)^{-1} = e^{-w_1 K_1} = \begin{pmatrix} \cosh w_1 & \sinh w_1 & 0 & 0 \\ \sinh w_1 & \cosh w_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{10.131}$$

This yields for small w_1

$$\begin{aligned} \rho\left(e^{w_1 K_1}\right)\psi(x^\mu) &= \psi(\cosh w_1\, x^0 + \sinh w_1\, x^1, \sinh w_1\, x^0 + \cosh w_1\, x^1, x^2, x^3) \\ &= \psi(x^\mu) + w_1\,(x^1\partial_0 + x^0\partial_1)\,\psi(x^\mu) + O(w_1^2) \end{aligned} \tag{10.132}$$

and, obviously, the expression for $\mathcal{K}_1$ in (10.126).
The generators $\vec{\mathcal{J}}$, $\vec{\mathcal{K}}$ obey the same Lie algebra (10.49) as the generators of the natural representation, i.e.

$$\begin{aligned} [\,\mathcal{J}_k,\, \mathcal{J}_\ell\,] &= \epsilon_{k\ell m}\,\mathcal{J}_m \\ [\,\mathcal{K}_k,\, \mathcal{K}_\ell\,] &= -\,\epsilon_{k\ell m}\,\mathcal{J}_m \\ [\,\mathcal{J}_k,\, \mathcal{K}_\ell\,] &= \epsilon_{k\ell m}\,\mathcal{K}_m\,. \end{aligned} \tag{10.133}$$

We demonstrate this for three cases, namely $[\mathcal{J}_1, \mathcal{J}_2] = \mathcal{J}_3$, $[\mathcal{K}_1, \mathcal{K}_2] = -\mathcal{J}_3$, and $[\mathcal{J}_1, \mathcal{K}_2] = \mathcal{K}_3$:

$$\begin{aligned} [\,\mathcal{J}_1,\, \mathcal{J}_2\,] &= [x^3\partial_2 - x^2\partial_3,\, x^1\partial_3 - x^3\partial_1] \\ &= [x^3\partial_2,\, x^1\partial_3] - [x^2\partial_3,\, x^3\partial_1] \\ &= -x^1\partial_2 + x^2\partial_1 = \mathcal{J}_3\,, \end{aligned} \tag{10.134}$$

$$\begin{aligned}[\mathcal{K}_1, \mathcal{K}_2] &= [x^0\partial_1 + x^1\partial_0, x^0\partial_2 + x^2\partial_0] \\ &= [x^0\partial_1, x^2\partial_0] - [x^1\partial_0, x^0\partial_2] \\ &= -x^2\partial_1 + x^1\partial_2 = -\mathcal{J}_3 , \end{aligned} \tag{10.135}$$

$$\begin{aligned}[\mathcal{J}_1, \mathcal{K}_2] &= [x^3\partial_2 - x^2\partial_3, x^0\partial_2 + x^2\partial_0] \\ &= [x^3\partial_2, x^2\partial_0] - [x^2\partial_3, x^0\partial_2] \\ &= x^3\partial_0 + x^0\partial_3 = \mathcal{K}_3 . \end{aligned} \tag{10.136}$$

One-Dimensional Function Space Representation

The exponential operator (10.125) in the case of a one-dimensional transformation of the type

$$\mathcal{L}(w^3) = \exp\left(w^3\mathcal{K}_3 \right) , \tag{10.137}$$

where $\mathcal{K}_3$ is given in (10.128), can be simplified considerably. For this purpose one expresses $\mathcal{K}_3$ in terms of hyperbolic coordinates R, Ω which are connected with x^0, x^3 as follows

$$x^0 = R\cosh\Omega , \quad x^3 = R\sinh\Omega \tag{10.138}$$

a relationship which can also be stated

$$R = \begin{cases} +\sqrt{(x^0)^2 - (x^3)^2} & \text{if } x^0 \geq 0 \\ -\sqrt{(x^0)^2 - (x^3)^2} & \text{if } x^0 < 0 \end{cases} \tag{10.139}$$

and

$$\tanh\Omega = \frac{x^3}{x^0} , \quad \coth\Omega = \frac{x^0}{x^3} . \tag{10.140}$$

The transformation to hyperbolic coordinates closely resembles the transformation to radial coordinates for the generators of SO(3) in the function space representation [cf. Eqs. (5.85-5.87)]. In both cases the radial coordinate is the quantity conserved under the transformations, i.e., $\sqrt{x_1^2 + x_2^2 + x_3^2}$ in the case of SO(3) and $\sqrt{(x^0)^2 - (x^3)^2}$ in case of transformation (10.137).
In the following we consider solely the case $x^0 \geq 0$. The relationships (10.139, 10.140) allow one to express the derivatives ∂_0, ∂_3 in terms of $\frac{\partial}{\partial R}$, $\frac{\partial}{\partial\Omega}$. We note

$$\frac{\partial R}{\partial x^0} = \frac{x^0}{R} , \quad \frac{\partial R}{\partial x^3} = -\frac{x^0}{R} \tag{10.141}$$

and

$$\begin{aligned}\frac{\partial\Omega}{\partial x^3} &= \frac{\partial\Omega}{\partial\tanh\Omega}\frac{\partial\tanh\Omega}{\partial x^3} = \cosh^2\Omega\,\frac{1}{x^0} \\ \frac{\partial\Omega}{\partial x^0} &= \frac{\partial\Omega}{\partial\coth\Omega}\frac{\partial\coth\Omega}{\partial x^0} = -\sinh^2\Omega\,\frac{1}{x^3} . \end{aligned} \tag{10.142}$$

The chain rule yields then

$$\begin{aligned}\partial_0 &= \frac{\partial R}{\partial x^0}\frac{\partial}{\partial R} + \frac{\partial\Omega}{\partial x^0}\frac{\partial}{\partial\Omega} = \frac{x^0}{R}\frac{\partial}{\partial R} - \sinh^2\Omega\,\frac{1}{x^3}\frac{\partial}{\partial\Omega} \\ \partial_3 &= \frac{\partial R}{\partial x^3}\frac{\partial}{\partial R} + \frac{\partial\Omega}{\partial x^3}\frac{\partial}{\partial\Omega} = -\frac{x^3}{R}\frac{\partial}{\partial R} + \cosh^2\Omega\,\frac{1}{x^0}\frac{\partial}{\partial\Omega} . \end{aligned} \tag{10.143}$$

Inserting these results into the definition of $\mathcal{K}_3$ in (10.128) yields

$$\mathcal{K}_3 \;=\; x^0\partial_3 \;+\; x^3\partial_0 \;=\; \frac{\partial}{\partial\Omega}\,. \tag{10.144}$$

The action of the exponential operator (10.137) on a function $f(\Omega) \in \mathbb{C}_\infty(1)$ is then that of a shift operator

$$\mathcal{L}(w^3)\, f(\Omega) \;=\; \exp\left(w^3 \frac{\partial}{\partial\Omega} \right) f(\Omega) \;=\; f(\Omega + w^3)\,. \tag{10.145}$$

10.5 Klein–Gordon Equation

In the following Sections we will provide a heuristic derivation of the two most widely used quantum mechanical descriptions in the relativistic regime, namely the Klein–Gordon and the Dirac equations. We will provide a 'derivation' of these two equations which stem from the historical development of relativistic quantum mechanics. The historic route to these two equations, however, is not very insightful, but certainly is short and, therefore, extremely useful. Further below we will provide a more systematic, representation theoretic treatment.

Free Particle Case

A quantum mechanical description of a relativistic free particle results from applying the *correspondence principle*, which allows one to replace classical observables by quantum mechanical operators acting on wave functions. In the position representation the correspondence principle states

$$\begin{aligned} E \;&\Longrightarrow\; \hat{E} \;=\; -\frac{\hbar}{i}\partial_t \\ \vec{p} \;&\Longrightarrow\; \hat{\vec{p}} \;=\; \frac{\hbar}{i}\nabla \end{aligned} \tag{10.146}$$

which, in 4-vector notation reads

$$p_\mu \;\Longrightarrow \hat{p}_\mu \;=\; i\hbar(\partial_t, \nabla) \;=\; i\hbar\partial_\mu\,; \quad p^\mu \;\Longrightarrow \hat{p}^\mu \;=\; i(\partial_t, -\nabla) \;=\; i\hbar\partial^\mu\,. \tag{10.147}$$

Applying the correspondence principle to (10.92) one obtains the wave equation

$$-\,\hbar^2\partial^\mu\partial_\mu\,\psi(x^\nu) \;=\; m^2\,\psi(x^\nu) \tag{10.148}$$

or

$$\left(\,\hbar^2\,\partial^\mu\partial_\mu \;+\; m^2\,\right)\,\psi(x^\nu) \;=\; 0\,. \tag{10.149}$$

where $\psi(x^\mu)$ is a *scalar*, complex-valued function. The latter property implies that upon change of reference frame $\psi(x^\mu)$ transforms according to (10.121, 10.122). The partial differential equation (10.151) is called the *Klein-Gordon equation.*

In the following we will employ so-called *natural units* $\hbar = c = 1$. In these units the quantities energy, momentum, mass, $(\text{length})^{-1}$, and $(\text{time})^{-1}$ all have the same dimension. In natural units the Klein–Gordon equation (10.151) reads

$$\left(\,\partial_\mu\partial^\mu \;+\; m^2\,\right)\,\psi(x^\mu) \;=\; 0 \tag{10.150}$$

or

$$\left(\partial_t^2 \; - \; \nabla^2 \; + \; m^2 \right) \psi(x^\mu) \; = \; 0 \, . \tag{10.151}$$

One can notice immeadiately that (10.150) is invariant under Lorentz transformations. This follows from the fact that $\partial_\mu \partial^\mu$ and m^2 are scalars, and that (as postulated) $\psi(x^\mu)$ is a scalar. Under Lorentz transformations the free particle Klein–Gordon equation (10.150) becomes

$$\left(\partial'_\mu \partial'^\mu + \; m^2 \right) \psi'(x'^\mu) \; = \; 0 \tag{10.152}$$

which has the same form as the Klein–Gordon equation in the original frame.

Current 4-Vector Associated with the Klein-Gordon Equation

As is well-known the Schrödinger equation of a free particle

$$i\partial_t \psi(\vec{r}, t) \; = \; -\frac{1}{2m} \nabla^2 \, \psi(\vec{r}, t) \tag{10.153}$$

is associated with a conservation law for particle probability

$$\partial_t \rho_S(\vec{r}, t) \; + \; \nabla \cdot \vec{j}_S(\vec{r}, t) \; = \; 0 \tag{10.154}$$

where

$$\rho_S(\vec{r}, t) \; = \; \psi^*(\vec{r}, t) \, \psi(\vec{r}, t) \tag{10.155}$$

describes the *positive definite* probability to detect a particle at position $\vec{r}$ at time t and where

$$\vec{j}_S(\vec{r}, t) \; = \; \frac{1}{2mi} \left[\, \psi^*(\vec{r}, t) \nabla \psi(\vec{r}, t) \; - \; \psi(\vec{r}, t) \nabla \psi^*(\vec{r}, t) \, \right] \tag{10.156}$$

describes the current density connected with motion of the particle probability distribution. To derive this conservation law one rewrites the Schrödinger equation in the form $(i\partial_t - \frac{1}{2m} \nabla^2)\psi \; = \; 0$ and considers

$$\mathrm{Im} \left[\psi^* \left(i\partial_t \; - \; \frac{1}{2m} \nabla^2 \right) \psi \right] \; = \; 0 \tag{10.157}$$

which is equivalent to (10.154).
In order to obtain the conservation law connected with the Klein–Gordon equation (10.150) one considers

$$\mathrm{Im} \left[\psi^* \left(\partial_\mu \partial^\mu \; + \; m^2 \right) \psi \right] \; = \; 0 \tag{10.158}$$

which yields

$$\begin{aligned} \psi^* \partial_t^2 \psi \; - \; \psi \partial_t^2 \psi^* \; - \; \psi^* \nabla^2 \psi \; + \; \psi \nabla^2 \psi^* \qquad &= \\ \partial_t \left(\psi^* \partial_t \psi \; - \; \psi \partial_t \psi^* \right) \; + \; \nabla \cdot \left(\psi \nabla \psi^* \; - \; \psi^* \nabla \psi \right) \quad &= \; 0 \end{aligned} \tag{10.159}$$

which corresponds to

$$\partial_t \rho_{KG}(\vec{r}, t) \; + \; \nabla \cdot \vec{j}_{KG}(\vec{r}, t) \; = \; 0 \tag{10.160}$$

where

$$\begin{aligned} \rho_{KG}(\vec{r}, t) \quad &= \; \frac{i}{2m} \left(\, \psi^*(\vec{r}, t) \partial_t \psi(\vec{r}, t) \; - \; \psi(\vec{r}, t) \partial_t \psi^*(\vec{r}, t) \, \right) \\ \vec{j}_{KG}(\vec{r}, t) \quad &= \; \frac{1}{2mi} \left(\, \psi^*(\vec{r}, t) \nabla \psi(\vec{r}, t) \; - \; \psi(\vec{r}, t) \nabla \psi^*(\vec{r}, t) \, \right) \, . \end{aligned} \tag{10.161}$$

This conservation law differs in one important aspect from that of the Schrödinger equation (10.154), namely, in that the expression for ρ_{KG} is *not* positive definite. When the Klein-Gordon equation had been initially suggested this lack of positive definiteness worried physicists to a degree that the Klein–Gordon equation was rejected and the search for a Lorentz–invariant quantum mechanical wave equation continued. Today, the Klein-Gordon equation is considered as a suitable equation to describe spin–0 particles, for example pions. The proper interpretation of $\rho_{KG}(\vec{r},t)$, it had been realized later, is actually that of a charge density, not of particle probability.

Solution of the Free Particle Klein–Gordon Equation

Solutions of the free particle Klein–Gordon equation are

$$\psi(x^\mu) \;=\; N\, e^{-ip_\mu x^\mu} \;=\; N\, e^{i(\vec{p}_0\cdot\vec{r}\,-\,E_o t)}\;. \tag{10.162}$$

Inserting this into the Klein–Gordon equation (10.151) yields

$$\left(\, E_o^2 \;-\; \vec{p}_0^{\,2} \;-\; m^2 \,\right)\, \psi(\vec{r},t) \;=\; 0 \tag{10.163}$$

which results in the expected [see (10.93] dispersion relationship connecting E_0, $\vec{p}_0$, m

$$E_0^2 \;=\; m^2 \;+\; \vec{p}_o^{\,2}\;. \tag{10.164}$$

The corresponding energy is

$$E_o(\vec{p}_o,\pm) \;=\; \pm\sqrt{m^2 \;+\; \vec{p}_o^{\,2}} \tag{10.165}$$

This result together with (10.162) shows that the solutions of the free particle Klein-Gordon equation (10.150) are actually determined by $\vec{p}_o$ and by the choice of sign $\pm$. We denote this by summarizing the solutions as follows

$$\begin{aligned} \left(\,\partial_\mu\partial^\mu \;+\; m^2\,\right)\psi_o(\vec{p},\lambda|x^\mu) \qquad &=\; 0 \\ \psi_o(\vec{p},\lambda|x^\mu) \;=\; N_{\lambda,p}\, e^{i(\vec{p}\cdot\vec{r}\,-\,\lambda E_o(\vec{p})t)} \quad & E_o(\vec{p}) \;=\; \sqrt{m^2 \;+\; \vec{p}_o^{\,2}}\,,\lambda \;=\pm \end{aligned} \tag{10.166}$$

The spectrum of the Klein–Gordon equation (10.150) is a continuum of positive energies $E \geq m$, corresponding to $\lambda = +$, and of negative energies $E \leq -m$, corresponding to $\lambda = -$. The density $\rho_{KG}(\vec{p},\lambda)$ associated with the corresponding wave functions $\psi_o(\vec{p},\lambda|x^\mu)$ according to (10.161) and (10.166) is

$$\rho_{KG}(\vec{p},\lambda) \;=\; \lambda\,\frac{E_o(\vec{p})}{m}\,\psi_o^*(\vec{p},\lambda|x^\mu)\psi_o(\vec{p},\lambda|x^\mu) \tag{10.167}$$

which is positive for $\lambda = +$ and negative for $\lambda = -$. The proper interpretation of the two cases is that the Klein–Gordon equation describes particles as well as anti-particles; the *anti-particles* carry a charge opposite to that of the associated *particles*, and the density $\rho_{KG}(\vec{p},\lambda)$ actually describes charge density rather than probability.

Generating a Solution Through Lorentz Transformation

A particle at rest, i.e., with $\vec{p} \;=\; 0$, according to (**??**) is decribed by the $\vec{r}$–independent wave function

$$\psi_o(\vec{p}=0,\lambda|x^\mu) \;=\; N\, e^{-i\lambda mt}\;, \quad \lambda \;=\; \pm\,. \tag{10.168}$$

We want to demonstrate now that the wave functions for $\vec{p} \neq 0$ in (**??**) can be obtained through appropriate Lorentz transformation of (10.168). For this purpose we consider the wave function for a particle moving with momentum velocity v in the direction of the x^3–axis. Such wave function should be generated by applying the Lorentz transformation in the function space representation (10.145) choosing $\frac{p}{m} = \sinh w^3$. This yields, in fact, for the wave function (10.168), using (10.138) to replace $t = x^0$ by hyperbolic coordinates R, Ω,

$$\begin{aligned} \mathcal{L}(w^3)\psi_o(\vec{p}=0,\lambda|x^\mu) &= \exp\left(w^3 \frac{\partial}{\partial\Omega} \right) N\, e^{-i\lambda m R \cosh\Omega} \\ &= N\, e^{-i\lambda m R \cosh(\Omega + w^3)} \; . \end{aligned} \tag{10.169}$$

The addition theorem of hyperbolic functions $\cosh(\Omega + w^3) = \cosh\Omega \cosh w^3 + \sinh\Omega \sinh w^3$ allows us to rewrite the exponent on the r.h.s. of (10.169)

$$-i\lambda\,(\,m\cosh w^3\,)\,(\,R\cosh\Omega\,) \;-\; i\lambda\,(\,m\sinh w^3\,)\,(\,R\sinh\Omega\,) \; . \tag{10.170}$$

The coordinate transformation (10.138) and the relationships (10.61) yield for this expression

$$-i\lambda \frac{m}{\sqrt{1\,-\,v^2}}\, x^0 \;-\; i\lambda \frac{m\,v}{\sqrt{1\,-\,v^2}}\, x^3 \; . \tag{10.171}$$

One can interpret then for $\lambda = +$, i.e., for positive energy solutions,

$$p \;=\; -\,mv/\sqrt{1\,-\,v^2} \tag{10.172}$$

as the momentum of the particle relative to the moving frame and

$$\frac{m}{\sqrt{1\,-\,v^2}} \;=\; \sqrt{\frac{m^2}{1\,-\,v^2}} \;=\; \sqrt{m^2 \,+\, \frac{m^2 v^2}{1\,-\,v^2}} \;=\; \sqrt{m^2 \,+\, p^2} \;=\; E_o(p) \tag{10.173}$$

as the energy [c.f. (10.166)] of the particle. In case of $\lambda = +$ one obtains finally

$$\mathcal{L}(w^3)\psi_o(\vec{p}=0,\lambda=+|x^\mu) \;=\; N\, e^{i(px^3 \,-\, E_o(p)x^0} \tag{10.174}$$

which agrees with the expression given in (10.166). In case of $\lambda = -$, i.e., for negative energy solutions, one has to interprete

$$p \;=\; mv/\sqrt{1\,-\,v^2} \tag{10.175}$$

as the momentum of the particle and one obtains

$$\mathcal{L}(w^3)\psi_o(\vec{p}=0,\lambda=-|x^\mu) \;=\; N\, e^{i(px^3 \,+\, E_o(p)x^0} \; . \tag{10.176}$$

10.6 Klein–Gordon Equation for Particles in an Electromagnetic Field

We consider now the quantum mechanical wave equation for a spin–0 particle moving in an electromagnetic field described by the 4-vector potential

$$A^\mu(x^\mu) \;=\; (V(\vec{r},t), \vec{A}(\vec{r},t)) \; ; \quad A_\mu(x^\mu) \;=\; (V(\vec{r},t), -\vec{A}(\vec{r},t)) \tag{10.177}$$

	free classical particle	classical particle in field $(V, \vec{A})$	free quantum particle	quantum particle in field $(V, \vec{A})$
energy	E	$E - qV$	$i\partial_t$	$i\partial_t - qV$
momentum	$\vec{p}$	$\vec{p} - q\vec{A}$	$\hat{\vec{p}} = -i\nabla$	$\hat{\vec{p}} - q\vec{A} = \hat{\vec{\pi}}$
4-vector	p^μ	$p^\mu - qA^\mu$	$i\partial_\mu$	$i\partial_\mu - qA_\mu = \pi_\mu$

Table 10.1:
Coupling of a particle of charge q to an electromagnetic field described by the 4-vector potential $A^\mu = (V, \vec{A})$ or $A_\mu = (V, -\vec{A})$. According to the so-called minimum coupling principle the presence of the field is accounted for by altering energy, momenta for classical particles and the respective operators for quantum mechanical particles in the manner shown. See also Eq. (10.147).

To obtain the appropriate wave equation we follow the derivation of the free particle Klein–Gordon equation above and apply again the correspondence principle to (10.93), albeit in a form, which couples a particle of charge q to an electromagnetic field described through the potential $A_\mu(x^\nu)$. According to the principle of minimal coupling [see (10.69)] one replaces the quantum mechanical operators, i.e., $i\partial_t$ and $-i\nabla$ in (10.150), according to the rules shown in Table 10.1. For this purpose one writes the Klein-Gordon equation (10.150)

$$\left(-g^{\mu\nu}(i\partial_\mu)(i\partial_\nu) + m^2\right)\psi(x^\mu) = 0\,. \tag{10.178}$$

According to the replacements in Table 10.1 this becomes

$$g^{\mu\nu}(i\partial_\mu - qA_\mu)(i\partial_\nu - A_\nu)\,\psi(x^\mu) = m^2\,\psi(x^\mu) \tag{10.179}$$

which can also be written

$$g^{\mu\nu}\pi_\mu\pi_\nu - ; m^2\,)\,\psi(x^\mu) = 0\,. \tag{10.180}$$

In terms of space-time derivatives this reads

$$(i\partial_t - qV(\vec{r},t))^2\,\psi(\vec{r},t) = \left[\left(-i\nabla - q\vec{A}(\vec{r},t)\right)^2 + m^2\right]\psi(\vec{r},t)\,. \tag{10.181}$$

Non-Relativistic Limit of Free Particle Klein–Gordon Equation

In order to consider further the interpretation of the positive and negative energy solutions of the Klein–Gordon equation one can consider the non-relativistic limit. For this purpose we split-off a factor $\exp(-imt)$ which describes the oscillations of the wave function due to the rest energy, and focus on the remaining part of the wave function, i.e., we define

$$\psi(\vec{r},t) = e^{-imt}\,\Psi(\vec{r},t)\,, \tag{10.182}$$

and seek an equation for $\Psi(\vec{r},t)$. We will also assume, in keeping withnthe non-relativistic limit, that the mass m of the particle, i.e., it's rest energy, is much larger than all other energy terms, in

particular, larger than $|i\partial_t\Psi/\Psi|$ and alrger than qV, i.e.,

$$|\frac{i\partial_t\Psi}{\Psi}| \;<<\; m\,, \quad |q\,V| \;<<\; m\,. \tag{10.183}$$

The term on the l.h.s. of (10.181) can then be approximated as follows:

$$\begin{aligned}(i\partial_t \,-\, qV)^2 e^{-imt}\Psi \;&=\; (i\partial_t \,-\, qV)\,(me^{-imt}\Psi \,+\, e^{-imt}i\partial_t\Psi \,-\, qVe^{-imt}\Psi)\\ &=\; m^2e^{-imt}\Psi \,+\, me^{-imt}i\partial_t\Psi \,-\, qVe^{-imt}\Psi\\ &\quad +me^{-imt}i\partial_t\Psi \,-\, e^{-imt}\partial^2\Psi \,-\, qVe^{-imt}i\partial_t\Psi\\ &\quad -me^{-imt}qV\Psi \,-\, e^{-imt}i\partial_t qV\Psi \,+\, q^2V^2e^{-imt}\Psi\\ &\approx\; m^2e^{-imt}\Psi \,-\, 2mqVe^{-imt}\Psi \,-\, 2me^{-imt}i\partial_t\Psi\end{aligned} \tag{10.184}$$

where we neglected all terms which did not contain factors m. The approximation is justified on the ground of the inequalities (10.183). The Klein-Gordon equation (10.181) reads then

$$i\,\partial_t\,\Psi(\vec{r},t) \;=\; \left[\frac{[\hat{\vec{p}} \,-\, q\vec{A}(\vec{r},t)]^2}{2m} \;+\; qV(\vec{r},t)\right]\Psi(\vec{r},t) \tag{10.185}$$

This is, however, identical to the Schrödinger equation (10.2) of a non-relativistic spin-0 particle moving in an electromagnetic field.

Pionic Atoms

To apply the Klein–Gordon equation (10.181) to a physical system we consider pionic atoms, i.e., atoms in which one or more electrons are replaced by π^- mesons. This application demonstrates that the Klein–Gordon equation describes spin zero particles, e.g., spin-0 mesons.
To 'manufacture' pionic atoms, π^- mesons are generated through inelastic proton–proton scattering

$$p \;+\; p \;\longrightarrow\; p \;+\; p \;+\; \pi^- \;+\; \pi^+\,, \tag{10.186}$$

then are slowed down, filtered out of the beam and finally fall as slow pions onto elements for which a pionic variant is to be studied. The process of π^- meson capture involves the so-called Auger effect, the binding of a negative charge (typically an electron) while at the same time a lower shell electron is being emitted

$$\pi^- \;+\; atom \;\longrightarrow\; (atom \,-\, e^- \,+\, \pi^-) \;+\; e^-\,. \tag{10.187}$$

We want to investigate in the following a description of a *stationary state* of a pionic atom involving a nucleus with charge $+Ze$ and a π^- meson. A stationary state of the Klein–Gordon equation is described by a wave function

$$\psi(x^\mu) \;=\; \varphi(\vec{r})\,e^{-i\epsilon t}\,. \tag{10.188}$$

Inserting this into (10.181) yields (we assume now that the Klein–Gordon equation describes a particle with mass m_π and charge $-e$) for $qV(\vec{r},t) \;=\; -\frac{Ze^2}{r}$ and $\vec{A}(\vec{r},t) \;\equiv\; 0$

$$\left[\left(\epsilon \,+\, \frac{Ze^2}{r}\right)^2 \;+\; \nabla^2 \;-\; m_\pi^2\right]\varphi(\vec{r}) \;=\; 0\,. \tag{10.189}$$

Because of the radial symmetry of the Coulomb potential we express this equation in terms of spherical coordinates r, θ, ϕ. The Laplacian is

$$\nabla^2 \;=\; \frac{1}{r}\partial_r^2 r \;+\; \frac{1}{r^2 \sin^2\theta}\partial_\theta \sin\theta \partial_\theta \;+\; \frac{1}{r^2\sin^2\theta}\partial_\phi^2 \;=\; \frac{1}{r}\partial_r^2 r \;-\; \frac{\hat{L}^2}{r^2}\,. \tag{10.190}$$

With this expression and after expanding $(\epsilon + \frac{Ze^2}{r})^2$ one obtains

$$\left(\frac{d^2}{dr^2} \;-\; \frac{\hat{L}^2 \;-\; Z^2e^4}{r^2} \;+\; \frac{2\epsilon Z e^2}{r} \;+\; \epsilon^2 \;-\; m_\pi^2\right) r\,\phi(\vec{r}) \;=\; 0\,. \tag{10.191}$$

The operator $\hat{L}^2$ in this equation suggests to choose a solution of the type

$$\varphi(\vec{r}) \;=\; \frac{R_\ell(r)}{r}\, Y_{\ell m}(\theta,\phi) \tag{10.192}$$

where the functions $Y_{\ell m}(\theta,\phi)$ are spherical harmonics, i.e., the eigenfunctions of the operator $\hat{L}^2$ in (10.191)

$$\hat{L}^2\, Y_{\ell m}(\theta,\phi) \;=\; \ell\,(\ell \,+\, 1)\, Y_{\ell m}(\theta,\phi)\,. \tag{10.193}$$

(10.192) leads then to the ordinary differential equation

$$\left(\frac{d^2}{dr^2} \;-\; \frac{\ell(\ell+1) \;-\; Z^2e^4}{r^2} \;+\; \frac{2\epsilon Z e^2}{r} \;+\; \epsilon^2 \;-\; m_\pi^2\right) R_\ell(r) \;=\; 0\,. \tag{10.194}$$

Bound state solutions can be obtained readily noticing that this equation is essentially identical to that posed by the Coulomb problem (potential $-\frac{Ze^2}{r}$) for the Schrödinger equation

$$\left(\frac{d^2}{dr^2} \;-\; \frac{\ell(\ell+1)}{r^2} \;+\; \frac{2m_\pi Z e^2}{r} \;+\; 2m_\pi E\right) R_\ell(r) \;=\; 0 \tag{10.195}$$

The latter problem leads to the well-known spectrum

$$E_n \;=\; -\,\frac{m_\pi\,(Ze^2)^2}{2n^2}\,; n \;=\; 1,2,\ldots\,; \ell \;=\; 0,1,\ldots n-1\,. \tag{10.196}$$

In this expression the number n' defined through

$$n' \;=\; n \;-\; \ell \;-1 \tag{10.197}$$

counts the number of nodes of the wave function, i.e., this quantity definitely must be an integer. The similarity of (10.194) and (10.195) can be made complete if one determines λ such that

$$\lambda(\ell)\,(\lambda(\ell) \,+\, 1) \;=\; \ell\,(\ell \,+\, 1) \;-\; Z^2e^4\,. \tag{10.198}$$

The suitable choice is

$$\lambda(\ell) \;=\; -\frac{1}{2} \;+\; \sqrt{(\ell \,+\, \frac{1}{2})^2 \;-\; Z^2e^4} \tag{10.199}$$

and one can write (10.194)

$$\left(\frac{d^2}{dr^2} \; - \; \frac{\lambda(\ell)\,(\,\lambda(\ell) \; + \; 1\,)}{r^2} \; + \; \frac{2\epsilon Z e^2}{r} \; + \; \epsilon^2 \; - \; m_\pi^2 \right) R_\ell(r) \; = \; 0\,. \tag{10.200}$$

The bound state solutions of this equation should correspond to ϵ values which can be obtained from (10.196) if one makes the replacement

$$E \; \longrightarrow \; \frac{\epsilon^2 \; - \; m_\pi^2}{2m_\pi}\,, \qquad \ell \; \longrightarrow \; \lambda(\ell)\,, \qquad e^2 \; \longrightarrow \; e^2\,\frac{\epsilon}{m_\pi}\,. \tag{10.201}$$

One obtains

$$\frac{\epsilon^2 \; - \; m_\pi^2}{2m_\pi} \; = \; -\frac{m_\pi\, Z^2 e^4 \frac{\epsilon^2}{m_\pi^2}}{2\,(n' \; + \; \lambda(\ell) \; + \; 1)^2}\,. \tag{10.202}$$

Solving this for ϵ (choosing the root which renders $\epsilon \; \leq \; m_\pi$, i.e., which corresponds to a bound state) yields

$$\epsilon \; = \; \frac{m_\pi}{\sqrt{1 \; + \; \frac{Z^2 e^4}{(n' + \lambda(\ell) + 1)^2}}} \quad ; \; n' \; = \; 0, 1, \ldots \; ; \; \ell \; = \; 0, 1, \ldots\,. \tag{10.203}$$

Using (10.197, 10.199) and defining $E_{\rm KG} \; = \; \epsilon$ results in the spectrum

$$E_{\rm KG}(n, \ell, m) \; = \; \frac{m_\pi}{\sqrt{1 \; + \; \frac{Z^2 e^4}{(\,n - \ell - \frac{1}{2} + \sqrt{(\ell + \frac{1}{2})^2 - Z^2 e^4}\,)^2}}} \qquad \begin{array}{l} n \; = \; 1, 2, \ldots \\ \ell \; = \; 0, 1, \ldots, n-1 \\ m \; = \; -\ell, -\ell+1, \ldots, +\ell \end{array} \tag{10.204}$$

In order to compare this result with the spectrum of the non-relativistic hydrogen-like atom we expand in terms of the fine structure constant e^2 to order $O(\epsilon^8)$. Introducing $\alpha \; = \; Z^2 e^4$ and $\beta \; = \; \ell + \frac{1}{2}$ (10.204) reads

$$\frac{1}{\sqrt{1 \; + \; \frac{\alpha}{(n - \beta + \sqrt{\beta^2 - \alpha})^2}}} \tag{10.205}$$

and one obtains the series of approximations

$$\begin{aligned} & \frac{1}{\sqrt{1 \; + \; \frac{\alpha}{(n - \beta + \sqrt{\beta^2 - \alpha})^2}}} \\ \approx \; & \frac{1}{\sqrt{1 \; + \; \frac{\alpha}{(n - \frac{\alpha}{2\beta} + O(\alpha^2))^2}}} \\ \approx \; & \frac{1}{\sqrt{1 \; + \; \frac{\alpha}{n^2 - \frac{\alpha}{\beta} n + O(\alpha^2)}}} \\ \approx \; & \frac{1}{\sqrt{1 \; + \; \frac{\alpha}{n^2} \; + \; \frac{\alpha^2}{\beta n^3} \; + \; O(\alpha^3)}} \\ \approx \; & \frac{1}{1 \; + \; \frac{\alpha}{2n^2} \; + \; \frac{\alpha^2}{2\beta n^3} \; - \; \frac{\alpha^2}{8n^4} \; + \; O(\alpha^3)} \end{aligned}$$

$$\approx \quad 1 - \frac{\alpha}{2n^2} - \frac{\alpha^2}{2\beta n^3} + \frac{\alpha^2}{8n^4} + \frac{\alpha^2}{4n^4} + O(\alpha^3)\,. \tag{10.206}$$

From this results for (10.204)

$$E_{\rm KG}(n,\ell,m) \approx m - \frac{mZ^2e^4}{2n^2} - \frac{mZ^4e^8}{2n^3}\left[\frac{1}{\ell+\frac{1}{2}} - \frac{3}{4n}\right] + O(Z^6e^{12})\,. \tag{10.207}$$

Here the first term represents the rest energy, the second term the non-relativistic energy, and the third term gives the leading relativistic correction. The latter term agrees with observations of pionic atoms, however, it does not agree with observations of the hydrogen spectrum. The latter spectrum shows, for example, a splitting of the six $n = 2, \ell = 1$ states into groups of two and four degenerate states. In order to describe electron spectra one must employ the Lorentz-invariant wave equation for spin-$\frac{1}{2}$ particles, i.e., the Dirac equation introduced below.
It must be pointed out here that ϵ does not denote energy, but in the present case rather the negative of the energy. Also, the π^- meson is a pseudoscalar particle, i.e., the wave function changes sign under reflection.

10.7 The Dirac Equation

Historically, the Klein–Gordon equation had been rejected since it did not yield a positive-definite probability density, a feature which is connected with the 2nd order time derivative in this equation. This derivative, in turn, arises because the Klein–Gordon equation, through the correspondence principle, is related to the equation $E^2 = m^2 + \vec{p}^2$ of the classical theory which involves a term E^2. A more satisfactory Lorentz–invariant wave equation, i.e., one with a positive-definite density, would have only a first order time derivative. However, because of the equivalence of space and time coordinates in the Minkowski space such equation necessarily can only have then first order derivatives with respect to spatial coordinates. It should feature then a differential operator of the type $\mathcal{D} = i\gamma^\mu\partial_\mu$.

Heuristic Derivation Starting from the Klein-Gordon Equation

An obvious starting point for a Lorentz-invariant wave equation with only a first order time derivative is $E = \pm\sqrt{m^2 + \vec{p}^2}$. Application of the correspondence principle (10.146) leads to the wave equation

$$i\partial_t\,\Psi(\vec{r},t) \;=\; \pm\sqrt{m^2 - \nabla^2}\,\Psi(\vec{r},t)\,. \tag{10.208}$$

These two equation can be combined

$$\left(i\partial_t + \sqrt{m^2 - \nabla^2}\right)\left(i\partial_t - \sqrt{m^2 - \nabla^2}\right)\Psi(\vec{r},t) \tag{10.209}$$

which, in fact, is identical to the two equations (10.208). Equations (10.208, 10.209), however, are unsatisfactory since expansion of the square root operator involves all powers of the Laplace operator, but not an operator $i\vec{\gamma}\cdot\nabla$ as suggested by the principle of relativity (equivalence of space and time). Many attempts were made by theoretical physicists to 'linearize' the square root operator in (10.208, 10.209), but for a long time to no avail. Finally, Dirac succeeded. His solution

to the problem involved an ingenious step, namely, the realization that the linearization can be carried out only if one assumes a 4-dimensional representation of the coefficients γ^μ.
Initially, it was assumed that the 4-dimensional space introduced by Dirac could be linked to 4-vectors, i.e., quantities with the transformation law (10.66). However, this was not so. Instead, the 4-dimensionsional representation discovered by Dirac involved new physical properties, spin-$\frac{1}{2}$ and anti-particles. The discovery by Dirac, achieved through a beautiful mathematical theory, strengthens the believe of many theoretical physicists today that the properties of physical matter ultimately derive from a, yet to be discovered, beautiful mathematical theory and that, therefore, one route to important discoveries in physics is the creation of new mathematical descriptions of nature, these descriptions ultimately merging with the true theory of matter.

Properties of the Dirac Matrices

Let us now trace Dirac's steps in achieving the linearization of the 'square root operator' in (10.208). Starting point is to boldly factorize, according to (10.209), the operator of the Klein–Gordon equation

$$\partial_\mu \partial^\mu \; + \; m^2 \; = \; -(\, P \; + \; m \,) \, (\, P \; - \; m \,) \tag{10.210}$$

where

$$P \; = \; i \gamma^\mu \partial_\mu \; . \tag{10.211}$$

Obviously, this would lead to the two wave equations $(P \, - \, m)\Psi \, = \, 0$ and $(P \, + \, m)\Psi \, = \, 0$ which have a first order time derivative and, therefore, are associated with a positive-definite particle density. We seek to identify the coefficients γ^μ. Inserting (10.211) into (10.210) yields

$$\begin{aligned} & -g^{\mu\nu} \partial_\mu \partial_\nu \; - \; m^2 \; = \; (i \gamma^\mu \partial_\mu \; + \; m)(i \gamma^\mu \partial_\mu \; - \; m) \\ = \quad & -\gamma^\mu \gamma^\nu \partial_\mu \partial_\nu \; - \; m^2 \; = \; -\frac{1}{2} \, (\, \gamma^\mu \gamma^\nu \partial_\mu \partial_\nu \; + \; \gamma^\nu \gamma^\mu \partial_\nu \partial_\mu \,) \; - \; m^2 \\ = \quad & -\frac{1}{2} \, (\, \gamma^\mu \gamma^\nu \; + \; \gamma^\nu \gamma^\mu \,) \, \partial_\mu \partial_\nu \; - \; m^2 \end{aligned} \tag{10.212}$$

where we have changed 'dummy' summation indices, exploited $\partial_\mu \partial_\nu \; = \; \partial_\nu \partial_\mu$, but did not commute the, so far, unspecified algebraic objects γ^μ and γ^ν. Comparing the left-most and the right-most side of the equations above one can conclude the following property of γ_μ

$$\gamma^\mu \gamma^\nu \; + \; \gamma^\nu \gamma^\mu \; = \; [\, \gamma^\mu , \, \gamma^\nu \,]_+ \; = \; 2 \, g^{\mu\nu} \tag{10.213}$$

We want to determine now the simplest algebraic realization of γ^μ. It turns out that no 4-vector of real or complex coefficients can satisfy these conditions. In fact, the quantities γ^0, γ^1, γ^2, γ^3 can only be realized by $d \times d$–matrices requiring that the wave function $\Psi(x^\mu)$ is actually a d–dimensional vector of functions $\psi_1(x^\mu)$, $\psi_2(x^\mu)$, $\ldots$ $\psi_d(x^\mu)$.
For $\mu \; = \; \nu$ condition (10.213) reads

$$(\gamma^\mu)^2 \; = \; \left\{ \begin{array}{ll} 1 & \mu \; = \; 0 \\ -1 & \mu \; = \; 1, 2, 3 \end{array} \right. \; . \tag{10.214}$$

From this follows that γ^0 has real eigenvalues ± 1 and γ^j, $j = 1, 2, 3$ has imaginary eigenvalues $\pm i$. Accordingly, one can impose the condition

$$\gamma^0 \quad \text{is hermitian} \qquad ; \gamma^j \; , \; j = 1, 2, 3 \quad \text{are anti-hermitian} \; . \tag{10.215}$$

For $\mu \neq \nu$ (10.213) reads

$$\gamma^\mu \gamma^\nu \;=\; -\gamma^\nu \gamma^\mu \;, \tag{10.216}$$

i.e., the γ^μ are anti-commuting. From this one can conclude for the determinants of γ^μ

$$\det(\gamma^\mu \gamma^\nu) \;=\; \det(-\gamma^\nu \gamma^\mu) \;=\; (-1)^d \det(\gamma^\nu \gamma^\mu) \;=\; (-1)^d \det(\gamma^\mu \gamma^\nu) \;. \tag{10.217}$$

Obviously, as long as $\det(\gamma^\mu) \neq 0$ the dimension d of the square matrices γ^μ must be even so that $(-1)^d = 1$.
For $d = 2$ there exist only three anti-commuting matrices, namely the Pauli matrices $\sigma^1, \sigma^2, \sigma^3$ for which, in fact, holds

$$\left(\sigma^j\right)^2 \;=\; \mathbb{1} \;; \quad \sigma^j \sigma^k \;=\; -\,\sigma^k \sigma^j \quad \text{for } j \neq k \;. \tag{10.218}$$

The Pauli matrices allow one, however, to construct four matrices γ^μ for the next possible dimension $d = 4$. A proper choice is

$$\gamma^0 \;=\; \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \quad ; \quad \gamma^j \;=\; \begin{pmatrix} 0 & \sigma^j \\ -\sigma^j & 0 \end{pmatrix} \;, \tag{10.219}$$

Using property (10.218) of the Pauli matrices one can readily prove that condition (10.213) is satisfied. We will argue further below that the choice f γ^μ, except for similarity trasnformations, is unique.

The Dirac Equation

Altogether we have shown that the Klein–Gordon equation can be factorized formally

$$\left(i\gamma^\mu \partial_\mu \;+\; m \right) \left(i\gamma^\mu \partial_\mu \;-\; m \right) \Psi(x^\mu) \;= 0 \tag{10.220}$$

where $\Psi(x^\mu)$ represents a 4-dimensional wave function, rather than a scalar wave function. From this equation one can conclude that also the following should hold

$$\left(i\gamma^\mu \partial_\mu \;-\; m \right) \Psi(x^\mu) \;= 0 \tag{10.221}$$

which is the celebrated *Dirac equation*.

The Adjoint Dirac Equation

The adjoint equation is

$$\Psi^\dagger(x^\mu) \left(i(\gamma^\mu)^\dagger \overleftarrow{\partial_\mu} \;+ m \right) \;= 0 \tag{10.222}$$

where we have defined $\Psi^\dagger \;=\; (\psi_1^*, \psi_2^*, \psi_3^*, \psi_4^*)$ and where $\overleftarrow{\partial_\mu}$ denotes the differential operator ∂^μ operating to the left side, rather than to the right side. One can readily show using the hermitian property of the Pauli matrices $(\gamma^0)^\dagger \;=\; \gamma^0$ and $(\gamma^j)^\dagger \;=\; -\gamma^j$ for $j = 1, 2, 3$ which, in fact, is implied by (10.215). This property can also be written

$$(\gamma^\mu)^\dagger \;=\; \gamma^0 \gamma^\mu \gamma^0 \;. \tag{10.223}$$

Inserting this into (10.222) and multiplication from the right by γ^0 yields the *adjoint Dirac equation*

$$\Psi^\dagger(x^\mu)\, \gamma^0 \left(i\gamma^\mu \overleftarrow{\partial_\mu} \;+ m \right) \;= 0 \;. \tag{10.224}$$

Similarity Transformations of the Dirac Equation - Chiral Representation

The Dirac equation can be subject to any similarity transformation defined through a non-singular 4×4–matrix S. Defining a new representation of the wave function $\tilde{\Psi}(x^\mu)$

$$\tilde{\Psi}(x^\mu) \;=\; S\,\Psi(x^\mu) \tag{10.225}$$

leads to the 'new' Dirac equation

$$(\, i\tilde{\gamma}^\mu \partial_\mu \;-\; m\,)\;\tilde{\Psi}(x^\mu) \;= 0 \tag{10.226}$$

where

$$\tilde{\gamma}^\mu \;=\; S\,\gamma^\mu S^{-1} \tag{10.227}$$

A representation often adopted beside the one given by (10.222, 10.219) is the socalled *chiral representation* defined through

$$\tilde{\Psi}(x^\mu) \;=\; S\,\Psi(x^\mu)\,; \quad S \;=\; \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbb{1} & \mathbb{1} \\ \mathbb{1} & -\mathbb{1} \end{pmatrix} \tag{10.228}$$

and

$$\tilde{\gamma}^0 \;=\; \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \quad ; \quad \tilde{\gamma}^j \;=\; \begin{pmatrix} 0 & -\sigma^j \\ \sigma^j & 0 \end{pmatrix}\,,\; j = 1,2,3\,. \tag{10.229}$$

The similarity transformation (10.227) leaves the algebra of the Dirac matrices unaffected and commutation property (10.213) still holds, i.e.,

$$[\tilde{\gamma}^\mu,\, \tilde{\gamma}^\nu]_+ \;=\; 2\,g^{\mu\nu}\,. \tag{10.230}$$

Exercise 10.7.1: Derive (refeq:Dirac-intro20a) from (10.213), (10.227).

Schrödinger Form of the Dirac Equation

Another form in which the Dirac equation is used often results from multiplying (10.221) from the left by γ^0

$$\left(\, i\,\partial_t \;+\; i\,\hat{\vec{\alpha}} \cdot \nabla \;-\; \hat{\beta}\, m \right)\,\Psi(\vec{r},t) \;=\; 0 \tag{10.231}$$

where$\hat{\vec{\alpha}}$ has the three components α_j, $j = 1,2,3$ and

$$\hat{\beta} \;=\; \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}\;; \quad \hat{\alpha}_j \;=\; \begin{pmatrix} 0 & \sigma^j \\ \sigma^j & 0 \end{pmatrix}\,,\; j = 1,2,3\,. \tag{10.232}$$

This form of the Dirac equation is called the Schrödinger form since it can be written in analogy to the time-dependent Schrödinger equation

$$i\partial_t \Psi(x^\mu) \;=\; \mathcal{H}_o \Psi(x^\mu)\,; \quad \mathcal{H}_o \;=\; \hat{\vec{\alpha}} \cdot \hat{\vec{p}} \;+\; \hat{\beta}\, m\,. \tag{10.233}$$

The eigenstates and eigenvalues of $\mathcal{H}$ correspond to the stationary states and energies of the particles described by the Dirac equation.

Clifford Algebra and Dirac Matrices

The matrices defined through

$$d_j \;=\; i\gamma^j \;, j \;=\; 1,2,3\;; \quad d_4 \;=\; \gamma^0 \tag{10.234}$$

satisfy the anti-commutation property

$$d_j d_k \;+\; d_k d_j \;=\; \begin{cases} 2 & \text{for } j \;=\; k \\ 0 & \text{for } j \;\neq\; k \end{cases} \tag{10.235}$$

as can be readily verified from (10.213). The associative algebra generated by $d_1 \ldots d_4$ is called a *Clifford algebra* C_4. The three Pauli matrices also obey the property (10.235) and, hence, form a Clifford algebra C_3. The representations of Clifford algebras C_m are well established. For example, in case of C_4, a representation of the d_j's is

$$\begin{aligned} d_1 \;&=\; \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \bigotimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \;;\; d_2 \;=\; \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \bigotimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ d_3 \;&=\; \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \bigotimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \;;\; d_4 \;=\; \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \bigotimes \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \end{aligned} \tag{10.236}$$

where '$\otimes$' denotes the Kronecker product between matrices, i.e., the matrix elements of $C \;=\; A \otimes B$ are $C_{jk,\ell m} \;=\; A_{j\ell} B_{km}$.
The Clifford algebra C_4 entails a subgroup G_4 of elements

$$\pm\, d_{j_1} d_{j_2} \cdots d_{j_s} \;, \quad j_1 < j_2 < \cdots < j_s \; s \leq 4 \tag{10.237}$$

which are the ordered products of the operators $\pm \mathbb{1}$ and d_1, d_2, d_3, d_4. Obviously, any product of the d_j's can be brought to the form (10.237) by means of the property (10.235). There are (including the different signs) 32 elements in G_4 which we define as follows

$$\begin{aligned} &\Gamma_{\pm 1} \;=\; \pm \mathbb{1} \\ &\Gamma_{\pm 2} \;=\; \pm d_1 \;,\; \Gamma_{\pm 3} \;=\; \pm d_2 \;,\; \Gamma_{\pm 4} \;=\; \pm d_3 \;,\; \Gamma_{\pm 5} \;=\; \pm d_4 \\ &\Gamma_{\pm 6} \;=\; \pm d_1 d_2 \;,\; \Gamma_{\pm 7} \;=\; \pm d_1 d_3 \;,\; \Gamma_{\pm 8} \;=\; \pm d_1 d_4 \;,\; \Gamma_{\pm 9} \;=\; \pm d_2 d_3 \\ &\Gamma_{\pm 10} \;=\; \pm d_2 d_4 \;,\; \Gamma_{\pm 11} \;=\; \pm d_3 d_4 \\ &\Gamma_{\pm 12} \;=\; \pm d_1 d_2 d_3 \;,\; \Gamma_{\pm 13} \;=\; \pm d_1 d_2 d_4 \;,\; \Gamma_{\pm 14} \;=\; \pm d_1 d_3 d_4 \;,\; \Gamma_{\pm 15} \;=\; \pm d_2 d_3 d_4 \\ &\Gamma_{\pm 16} \;=\; \pm d_1 d_2 d_3 d_4 \end{aligned} \tag{10.238}$$

These elements form a group since obviously any product of two Γ_r's can be expressed in terms of a third Γ_r. The representations of this group are given by a set of 32 4×4–matrices which are equivalent with respect to similarity transformations. Since the Γ_j are hermitian the similarity transformations are actually given in terms of unitary transformations. One can conclude then that also any set of 4×4–matrices obeying (10.235) can differ only with respect to unitary similarity transformations. This property extends then to 4×4–matrices which obey (10.213), i.e., to Dirac matrices.
To complete the proof in this section the reader may consult Miller 'Symmetry Groups and their Application' Chapter 9.6 and R.H.Good, Rev.Mod.Phys. 27, (1955), page 187. The reader may

also want to establish the unitary transformation which connects the Dirac matrices in the form (10.236) with the Dirac representation (10.219).

Exercise 7.2:
Demonstrate the anti-commutation relationships (10.218) of the Pauli matrices σ^j.
Exercise 7.3:
Demonstrate the anti-commutation relationships (10.218) of the Dirac matrices γ^μ.
Exercise 7.4:
Show that from (10.214) follows that γ^0 has real eigenvalues ± 1 and can be represented by a hermitian matrix, and γ^j, $j = 1, 2, 3$ has imaginary eigenvalues $\pm i$ and can be represented by an anti-hermitian matrix.

10.8 Lorentz Invariance of the Dirac Equation

We want to show now that the Dirac equation is invariant under Lorentz transformations, i.e., the form of the Dirac equation is identical in equivalent frames of reference, i.e., in frames connected by Lorentz transformations. The latter transformations imply that coordinates transform according to (10.6), i.e., $x' = L^\mu{}_\nu x^\nu$, and derivatives according to (10.80). Multiplication and summation of (10.80) by $L^\mu{}_\rho$ and using (10.76) yields $\partial_\rho = L^\nu{}_\rho \partial'_\nu$, a result one could have also obtained by applying the chain rule to (10.6). We can, therefore, transform coordinates and derivatives of the Dirac equation. However, we do not know yet how to transform the 4-dimensional wave function Ψ and the Dirac matrices γ^μ.

Lorentz Transformation of the Bispinor State

Actually, we will approach the proof of the Lorentz invariance of the Dirac equation by testing if there exists a transformation of the bispinor wave function Ψ and of the Dirac matrices γ^μ which together with the transformations of coordinates and derivatives leaves the form of the Dirac equation invariant, i.e., in a moving frame should hold

$$\left(i\gamma'^\mu \partial'_\mu \; - \; m \right) \Psi'(x'^\mu) \; = 0 \, . \tag{10.239}$$

Form invariance implies that the matrices γ'^μ should have the same properties as γ^μ, namely, (10.213, 10.215). Except for a similarity transformation, these properties determine the matrices γ'^μ uniquely, i.e., it must hold $\gamma'^\mu = \gamma^\mu$. Hence, in a moving frame holds

$$\left(i\gamma^\mu \partial'_\mu \; - \; m \right) \Psi'(x'^\mu) \; = 0 \, . \tag{10.240}$$

Infinitesimal Bispinor State Transformation

We want to show now that a suitable transformation of $\Psi(x^\mu)$ does, in fact, exist. The transformation is assumed to be linear and of the form

$$\Psi'(x'^\mu) \; = \; \mathcal{S}(L^\mu{}_\nu) \, \Psi(x^\mu) \quad ; \, x'^\mu \; = \; L^\mu{}_\nu x^\nu \tag{10.241}$$

where $\mathcal{S}(L^{\mu}{}_{\nu})$ is a non-singular 4×4–matrix, the coefficients of which depend on the matrix $L^{\mu}{}_{\nu}$ defining the Lorentz transformation in such a way that $\mathcal{S}(L^{\mu}{}_{\nu}) = \mathbb{1}$ for $L^{\mu}{}_{\nu} = \delta^{\mu}{}_{\nu}$ holds. Obviously, the transformation (10.241) implies a similarity transformation $\mathcal{S}\gamma^{\mu}\mathcal{S}^{-1}$. One can, hence, state that the Dirac equation (10.221) upon Lorentz transformation yields

$$\left(i\mathcal{S}(L^{\eta}{}_{\xi})\gamma^{\mu}\mathcal{S}^{-1}(L^{\eta}{}_{\xi})\, L^{\nu}{}_{\mu}\, \partial'_{\nu} \; - \; m \right) \Psi'(x'^{\mu}) \; = \; 0 \; . \tag{10.242}$$

The form invariance of the Dirac equation under this transformation implies then the condition

$$\mathcal{S}(L^{\eta}{}_{\xi})\gamma^{\mu}\mathcal{S}^{-1}(L^{\eta}{}_{\xi})\, L^{\nu}{}_{\mu} \; = \; \gamma^{\nu} \; . \tag{10.243}$$

We want to determine now the 4×4–matrix $\mathcal{S}(L^{\eta}{}_{\xi})$ which satisfies this condition.
The proper starting point for a constructiuon of $\mathcal{S}(L^{\eta}{}_{\xi})$ is actually (10.243) in a form in which the Lorentz transformation in the form $L^{\mu\nu}$ is on the r.h.s. of the equation. For this purpose we exploit (10.12) in the form $L^{\nu}{}_{\mu} g_{\nu\sigma} L^{\sigma}{}_{\rho} \; = \; g_{\mu\rho} \; = \; g_{\rho\mu}$. Multiplication of (10.243) from the left by $L^{\sigma}{}_{\rho} g_{\sigma\nu}$ yields

$$\mathcal{S}(L^{\eta}{}_{\xi})\gamma^{\mu}\mathcal{S}^{-1}(L^{\eta}{}_{\xi})\, g_{\rho\mu} \; = \; L^{\sigma}{}_{\rho} g_{\sigma\nu} \gamma_{\sigma} \; . \tag{10.244}$$

from which, using $g_{\rho\mu}\gamma^{\mu} \; = \; \gamma_{\rho}$, follows

$$\mathcal{S}(L^{\eta}{}_{\xi})\gamma_{\rho}\mathcal{S}^{-1}(L^{\eta}{}_{\xi}) \; = \; L^{\sigma}{}_{\rho}\gamma_{\sigma} \; . \tag{10.245}$$

One can finally conclude multiplying both sides by $g^{\rho\mu}$

$$\mathcal{S}(L^{\eta}{}_{\xi})\gamma^{\mu}\mathcal{S}^{-1}(L^{\eta}{}_{\xi}) \; = \; L^{\nu\mu}\gamma_{\nu} \; . \tag{10.246}$$

The construction of $\mathcal{S}(L^{\eta}{}_{\xi})$ will proceed using the avenue of infinitesimal transformations. We had introduced in (10.38) the infinitesimal Lorentz transformations in the form $L^{\mu}{}_{\nu} \; = \; \delta^{\mu}{}_{\nu} + \epsilon^{\mu}{}_{\nu}$ where the infinitesimal operator $\epsilon^{\mu}{}_{\nu}$ obeyed $\epsilon^{T} \; = \; - \, \mathbf{g}\epsilon\mathbf{g}$. Multiplication of this property by $\mathbf{g}$ from the right yields $(\epsilon\mathbf{g})^{T} \; = \; - \, \epsilon\mathbf{g}$, i.e., $\epsilon\mathbf{g}$ is an anti-symmetric matrix. The elements of $\epsilon\mathbf{g}$ are, however, $\epsilon^{\mu}{}_{\rho} g^{\rho\nu} \; = \; \epsilon^{\mu\nu}$ and, hence, in the expression of the infinitesimal transformation

$$L^{\mu\nu} \; = \; g^{\mu\nu} \; + \; \epsilon^{\mu\nu} \tag{10.247}$$

the infinitesimal matrix $\epsilon^{\mu\nu}$ is anti-symmetric.
The infinitesimal transformation $\mathcal{S}(L^{\rho}{}_{\sigma})$ which corresponds to (10.247) can be expanded

$$\mathcal{S}(\epsilon^{\mu\nu}) \; = \; \mathbb{1} - \frac{i}{4}\sigma_{\mu\nu}\epsilon^{\mu\nu} \tag{10.248}$$

Here $\sigma_{\mu\nu}$ denote 4×4–matrices operating in the 4-dimensional space of the wave functions Ψ. $\mathcal{S}(\epsilon^{\mu\nu})$ should not change its value if one replaces in its argument $\epsilon^{\mu\nu}$ by $-\epsilon^{\nu\mu}$. It holds then

$$\mathcal{S}(\epsilon^{\mu\nu}) \; = \; \mathbb{1} - \frac{i}{4}\sigma_{\mu\nu}\epsilon^{\mu\nu} \; = \; \mathcal{S}(-\epsilon^{\nu\mu}) \; = \; \mathbb{1} + \frac{i}{4}\sigma_{\mu\nu}\epsilon^{\nu\mu} \tag{10.249}$$

from which we can conclude $\sigma_{\mu\nu}\epsilon^{\mu\nu} \; = \; -\sigma_{\mu\nu}\epsilon^{\nu\mu} \; = \; -\sigma_{\nu\mu}\epsilon^{\mu\nu}$, i.e., it must hold

$$\sigma_{\mu\nu} \; = \; -\sigma_{\nu\mu} \; . \tag{10.250}$$

One can readily show expanding $\mathcal{S}\mathcal{S}^{-1} = \mathbb{1}$ to first order in $\epsilon^{\mu\nu}$ that for the inverse infinitesimal transformation holds

$$\mathcal{S}^{-1}(\epsilon^{\mu\nu}) \;=\; \mathbb{1} + \frac{i}{4}\sigma_{\mu\nu}\epsilon^{\mu\nu} \tag{10.251}$$

Inserting (10.248, 10.251) into (10.246) results then in a condition for the generators $\sigma_{\mu\nu}$

$$-\frac{i}{4}\,(\,\sigma_{\alpha\beta}\gamma^{\mu} \;-\; \gamma^{\mu}\sigma_{\alpha\beta}\,)\,\epsilon^{\alpha\beta} \;=\; \epsilon^{\nu\mu}\gamma_{\nu}\;. \tag{10.252}$$

Since six of the coefficients $\epsilon^{\alpha\beta}$ can be chosen independently, this condition can actually be expressed through six independent conditions. For this purpose one needs to express formally the r.h.s. of (10.252) also as a sum over both indices of $\epsilon^{\alpha\beta}$. Furthermore, the expression on the r.h.s., like the expression on the l.h.s., must be symmetric with respect to interchange of the indices α and β. For this purpose we express

$$\begin{aligned}\epsilon^{\nu\mu}\gamma_{\nu} \;=\; \frac{1}{2}\epsilon^{\alpha\mu}\gamma_{\alpha} \;+\; \frac{1}{2}\epsilon^{\beta\mu}\gamma_{\beta} \;&=\; \frac{1}{2}\epsilon^{\alpha\beta}\,\delta^{\mu}{}_{\beta}\gamma_{\alpha} \;+\; \frac{1}{2}\epsilon^{\alpha\beta}\,\delta^{\mu}{}_{\alpha}\gamma_{\beta} \\ &=\; \frac{1}{2}\epsilon^{\alpha\beta}\,(\,\delta^{\mu}{}_{\beta}\gamma_{\alpha} \;-\; \delta^{\mu}{}_{\alpha}\gamma_{\beta}\,)\;.\end{aligned} \tag{10.253}$$

Comparing this with the l.h.s. of (10.252) results in the condition for each α, β

$$[\,\sigma_{\alpha\beta},\,\gamma^{\mu}\,]_{-} \;=\; 2i\,(\,\delta^{\mu}{}_{\beta}\gamma_{\alpha} \;-\; \delta^{\mu}{}_{\alpha}\gamma_{\beta}\,)\;. \tag{10.254}$$

The proper $\sigma_{\alpha\beta}$ must be anti-symmetric in the indices α, β and operate in the same space as the Dirac matrices. In fact, a solution of condition (10.254) is

$$\sigma_{\alpha\beta} \;=\; \frac{i}{2}\,[\,\gamma_{\alpha},\,\gamma_{\beta}\,]_{-} \tag{10.255}$$

which can be demonstrated using the properties (10.213, 10.216) of the Dirac matrices.

Exercise 7.5:
Show that the $\sigma_{\alpha\beta}$ defined through (10.255) satisfy condition (10.254).

Algebra of Generators of Bispinor Transformation

We want to construct the bispinor Lorentz transformation by exponentiating the generators $\sigma_{\mu\nu}$. For this purpose we need to verify that the algebra of the generators involving addition and multiplication is closed. For this purpose we inspect the properties of the generators in a particular representation, namely, the chiral representation introduced above in Eqs. (10.228, 10.229). In this representation the Dirac matrices $\tilde{\gamma}_{\mu} = (\tilde{\gamma}^{0}, -\vec{\tilde{\gamma}})$ are

$$\tilde{\gamma}_{0} \;=\; \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \quad ; \quad \tilde{\gamma}_{j} \;=\; \begin{pmatrix} 0 & \sigma^{j} \\ -\sigma^{j} & 0 \end{pmatrix}\;,\; j = 1,2,3\;. \tag{10.256}$$

One can readily verify that the non-vanishing generators $\tilde{\sigma}_{\mu\nu}$ are given by (note $\tilde{\sigma}_{\mu\nu} = -\tilde{\sigma}_{\nu\mu}$, i.e. only six generators need to be determined)

$$\tilde{\sigma}_{0j} \;=\; \frac{i}{2}\,[\tilde{\gamma}_{0}, \tilde{\gamma}_{j}] \;=\; \begin{pmatrix} -i\sigma^{j} & 0 \\ 0 & i\sigma^{j} \end{pmatrix}\;;\; \tilde{\sigma}_{jk} \;=\; [\tilde{\gamma}_{j}, \tilde{\gamma}_{k}] \;=\; \epsilon_{jk\ell}\begin{pmatrix} \sigma^{\ell} & 0 \\ 0 & \sigma^{\ell} \end{pmatrix}\;. \tag{10.257}$$

Obviously, the algebra of these generators is closed under addition and multiplication, since both operations convert block-diagonal operators

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \tag{10.258}$$

again into block-diagonal operators, and since the algebra of the Pauli matrices is closed.
We can finally note that the closedness of the algebra of the generators $\sigma_{\mu\nu}$ is not affected by similarity transformations and that, therefore, any representation of the generators, in particular, the representation (10.255) yields a closed algebra.

Finite Bispinor Transformation

The closedness of the algebra of the generators $\sigma_{\mu\nu}$ defined through (10.248) allows us to write the transformation $\mathcal{S}$ for any, i.e., not necessarily infinitesimal, $\epsilon^{\mu\nu}$ in the exponential form

$$\mathcal{S} \;=\; \exp\left(-\frac{i}{4}\,\sigma_{\mu\nu}\epsilon^{\mu\nu}\right)\,. \tag{10.259}$$

We had stated before that the transformation $\mathcal{S}$ is actually determined through the Lorentz transformation $L^{\mu}{}_{\nu}$. One should, therefore, be able to state $\mathcal{S}$ in terms of the same parameters $\vec{w}$ and $\vec{\vartheta}$ as the Lorentz transformation in (10.51). In fact, one can express the tensor $\epsilon^{\mu\nu}$ through $\vec{w}$ and $\vec{\vartheta}$ using $\epsilon^{\mu\nu} \;=\; \epsilon^{\mu}{}_{\rho}g^{\rho\nu}$ and the expression (10.44)

$$\epsilon^{\mu\nu} \;=\; \begin{pmatrix} 0 & -w_1 & -w_2 & -w_3 \\ w_1 & 0 & \vartheta_3 & -\vartheta_2 \\ w_2 & -\vartheta_3 & 0 & \vartheta_1 \\ w_3 & \vartheta_2 & -\vartheta_1 & 0 \end{pmatrix} \tag{10.260}$$

Inserting this into (10.259) yields the desired connection between the Lorentz transformation (10.51) and $\mathcal{S}$.
In order to construct an explicit expression of $\mathcal{S}$ in terms of $\vec{w}$ and $\vec{\vartheta}$ we employ again the chiral representation. In this representation holds

$$\begin{aligned} -\tfrac{i}{4}\tilde{\sigma}_{\mu\nu}\epsilon^{\mu\nu} \;&=\; -\frac{i}{2}\,(\tilde{\sigma}_{01}\epsilon^{01} \,+\, \tilde{\sigma}_{02}\epsilon^{02} \,+\, \tilde{\sigma}_{03}\epsilon^{03} \,+\, \tilde{\sigma}_{12}\epsilon^{12} \,+\, \tilde{\sigma}_{13}\epsilon^{13} \,+\, \tilde{\sigma}_{23}\epsilon^{23}) \\ &=\; \frac{1}{2}\begin{pmatrix} (\vec{w} \,-\, i\vec{\vartheta})\cdot\vec{\sigma} & 0 \\ 0 & -(\vec{w} \,+\, i\vec{\vartheta})\cdot\vec{\sigma} \end{pmatrix}\,. \end{aligned} \tag{10.261}$$

We note that this operator is block-diagonal. Since such operator does not change its block-diagonal form upon exponentiation the bispinor transformation (10.259) becomes in the chiral representation

$$\tilde{\mathcal{S}}(\vec{w},\,\vec{\vartheta}) \;=\; \begin{pmatrix} e^{\frac{1}{2}(\vec{w}\,-\,i\vec{\vartheta})\cdot\vec{\sigma}} & 0 \\ 0 & e^{-\frac{1}{2}(\vec{w}\,+\,i\vec{\vartheta})\cdot\vec{\sigma}} \end{pmatrix} \tag{10.262}$$

This expression allows one to transform according to (10.241) bispinor wave functions from one frame of reference into another frame of reference.

Current 4-Vector Associated with Dirac Equation

We like to derive now an expression for the current 4-vector j^μ associated with the Dirac equation which satisfies the conservation law

$$\partial_\mu j^\mu \;=\; 0\,. \tag{10.263}$$

Starting point are the Dirac equation in the form (10.221) and the adjoint Dirac equation (10.224). Multiplying (10.221) from the left by $\Psi^\dagger(x^\mu)\gamma^0$, (10.224) from the right by $\Psi(x^\mu)$, and addition yields

$$\Psi^\dagger(x^\mu)\gamma^0 \left(i\gamma^\mu\partial_\mu \;+\; i\gamma^\mu \overleftarrow{\partial_\mu} \right) \Psi(x^\mu) \;=\; 0\,. \tag{10.264}$$

The last result can be written

$$\partial_\mu\, \Psi^\dagger(x^\nu)\gamma^0\gamma^\mu\Psi(x^\nu) \;=\; 0\,, \tag{10.265}$$

i.e., the conservation law (10.263) does hold, in fact, for

$$j^\mu(x^\mu) \;=\; (\rho, \vec{j}) \;=\; \Psi^\dagger(x^\mu)\gamma^0\gamma^\mu\Psi(x^\mu)\,. \tag{10.266}$$

The time-like component ρ of j^μ

$$\rho(x^\mu) \;=\; \Psi^\dagger(x^\mu)\Psi(x^\mu) \;=\; \sum_{s=1}^{4} |\psi_s(x^\mu)|^2 \tag{10.267}$$

has the desired property of being positive definite.
The conservation law (10.263) allows one to conclude that j^μ must transform like a contravariant 4-vector as the notation implies. The reason is that the r.h.s. of (10.263) obviously is a scalar under Lorentz transformations and that the left hand side must then also transform like a scalar. Since ∂_μ transforms like a covariant 4-vector, j^μ must transform like a contravariant 4-vector. This transformation behaviour can also be deduced from the transformation properties of the bispinor wave function $\Psi(x_\mu)$. For this purpose we prove first the relationship

$$\mathcal{S}^{-1} \;=\; \gamma^0\,\mathcal{S}^\dagger\gamma^0\,. \tag{10.268}$$

We will prove this property in the chiral representation. Obviously, the property applies then in any representation of $\mathcal{S}$.
For our proof we note first

$$\tilde{\mathcal{S}}^{-1}(\vec{w},\vec{\vartheta}) \;=\; \tilde{\mathcal{S}}(-\vec{w},-\vec{\vartheta}) \;=\; \begin{pmatrix} e^{-\frac{1}{2}(\vec{w}-i\vec{\vartheta})\cdot\vec{\sigma}} & 0 \\ 0 & e^{\frac{1}{2}(\vec{w}+i\vec{\vartheta})\cdot\vec{\sigma}} \end{pmatrix} \tag{10.269}$$

One can readily show that the same operator is obtained evaluating

$$\tilde{\gamma}^0\tilde{\mathcal{S}}^\dagger(\vec{w},\vec{\vartheta})\tilde{\gamma}^0 \;=\; \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \begin{pmatrix} e^{\frac{1}{2}(\vec{w}+i\vec{\vartheta})\cdot\vec{\sigma}} & 0 \\ 0 & e^{-\frac{1}{2}(\vec{w}-i\vec{\vartheta})\cdot\vec{\sigma}} \end{pmatrix} \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}\,. \tag{10.270}$$

We conclude that (10.268) holds for the bispinor Lorentz transformation.
We will now determine the relationship between the flux

$$j'^\mu \;=\; \Psi'^\dagger(x'^\mu)\gamma^0\gamma^\mu\Psi'(x') \tag{10.271}$$

in a moving frame of reference and the flux j^μ in a frame at rest. Note that we have assumed in (10.271) that the Dirac matrices are independent of the frame of reference. One obtains using (10.268)

$$j'^\mu \;=\; \Psi^\dagger(x^\mu)\mathcal{S}^\dagger\gamma^0\gamma^\mu\mathcal{S}\Psi(x^\mu) \;=\; \Psi^\dagger(x^\mu)\gamma^0\mathcal{S}^{-1}\gamma^\mu\mathcal{S}\Psi(x^\mu)\;. \tag{10.272}$$

With $\mathcal{S}^{-1}(L^\eta{}_\xi) = \mathcal{S}((L^{-1})^\eta{}_\xi)$ one can restate (10.246)

$$\mathcal{S}^{-1}(L^\eta{}_\xi)\gamma^\mu\mathcal{S}^(L^\eta{}_\xi) \;=\; (L^{-1})^{\nu\mu}\gamma_\nu \;=\; (L^{-1})_\nu{}^\mu\gamma^\nu \;=\; L^\mu{}_\nu\gamma_\nu. \tag{10.273}$$

where we have employed (10.76). Combining this with (10.272) results in the expected transformation behaviour

$$j'^\mu \;=\; L^\mu{}_\nu j^\nu\;. \tag{10.274}$$

10.9 Solutions of the Free Particle Dirac Equation

We want to determine now the wave functions of free particles described by the Dirac equation. Like in non-relativistic quantum mechanics the free particle wave function plays a central role, not only as the most simple demonstration of the theory, but also as providing a basis in which the wave functions of interacting particle systems can be expanded and characterized. The solutions provide also a complete, orthonormal basis and allows one to quantize the Dirac field $\Psi(x^\mu)$ much like the classical electromagnetic field is quantized through creation and annihilation operators representing free electromagnetic waves of fixed momentum and frequency.

In case of non-relativistic quantum mechanics the free particle wave function has a single component $\psi(\vec{r},t)$ and is determined through the momentum $\vec{p} \in \mathbb{R}^3$. In relativistic quantum mechanics a Dirac particle can also be characterized through a momentum, however, the wave function has four components which invite further characterization of the free particle state. In the following we want to provide this characterization, specific for the Dirac free particle.

We will start from the Dirac equation in the Schrödinger form (10.231, 10.232, 10.233)

$$i\partial_t\Psi(x^\mu) \;=\; \mathcal{H}_o\Psi(x^\mu)\;. \tag{10.275}$$

The free particle wave function is an eigenfunction of $\mathcal{H}_o$, a property which leads to the energy–momentum (dispersion) relationship of the Dirac particle. The additional degrees of freedom described by the four components of the bispinor wave function require, as just mentioned, additional characterizations, i.e., the identification of observables and their quantum mechanical operators, of which the wave functions are eigenfunctions as well. As it turns out, only two degrees of freedom of the bispinor four degrees of freedom are independent [c.f. (10.282, 10.283)]. The independent degrees of freedom allow one to choose the states of the free Dirac particle as eigenstates of the 4-momentum operator $\hat{p}_\mu$ and of the helicity operator $\Gamma \sim \vec{\sigma}\cdot\hat{\vec{p}}/|\hat{\vec{p}}|$ introduced below. These operators, as is required for the mentioned property, commute with each other. The operators commute also with $\mathcal{H}_o$ in (10.233).

Like for the free particle wave functions of the non-relativistic Schrödinger and the Klein–Gordon equations one expects that the space–time dependence is governed by a factor $\exp[i(\vec{p}\cdot\vec{r} - \epsilon t)]$. As pointed out, the Dirac particles are described by 4-dimensional, bispinor wave functions and we need to determine corresponding components of the wave function. For this purpose we consider

the following form of the free Dirac particle wave function

$$\Psi(x^\mu) \;=\; \left(\begin{array}{c} \phi(x^\mu) \\ \chi(x^\mu) \end{array}\right) \;=\; \left(\begin{array}{c} \phi_o \\ \chi_o \end{array}\right) e^{i(\vec{p}\cdot\vec{r}\,-\,\epsilon t)} \tag{10.276}$$

where $\vec{p}$ and ϵ together represent four real constants, later to be identified with momentum and energy, and ϕ_o, χ_o each represent a constant, two-dimensional spinor state. Inserting (10.276) into (10.231, 10.232) leads to the 4-dimensional eigenvalue problem

$$\left(\begin{array}{cc} m & \vec{\sigma}\cdot\vec{p} \\ \vec{\sigma}\cdot\vec{p} & -m \end{array}\right) \left(\begin{array}{c} \phi_o \\ \chi_o \end{array}\right) \;=\; \epsilon \left(\begin{array}{c} \phi_o \\ \chi_o \end{array}\right) . \tag{10.277}$$

To solve this problem we write (10.277) explicitly

$$\begin{aligned} (\epsilon \,-\, m)\,\mathbb{1}\!\!\mathrm{l}\,\phi_o \,-\, \vec{\sigma}\cdot\vec{p}\,\chi_o \quad &=\; 0 \\ -\,\vec{\sigma}\cdot\vec{p}\,\phi_o \,+\, (\epsilon \,+\, m)\,\mathbb{1}\!\!\mathrm{l}\,\chi_o \quad &=\; 0\,. \end{aligned} \tag{10.278}$$

Multiplication of the 1st equation by $(\epsilon + m)\mathbb{1}\!\!\mathrm{l}$ and of the second equation by $-\vec{\sigma}\cdot\vec{p}$ and subtraction of the results yields the 2-dimensional equation

$$\left[\,(\epsilon^2 \,-\, m^2)\,\mathbb{1}\!\!\mathrm{l} \,-\, (\vec{\sigma}\cdot\vec{p})^2\,\right]\,\phi_o \;=\; 0\,. \tag{10.279}$$

According to the property (5.234) of Pauli matrices holds $(\vec{\sigma}\cdot\vec{p})^2 \;=\; \vec{p}^{\,2}\mathbb{1}\!\!\mathrm{l}$. One can, hence, conclude from (10.279) the well-known relativistic dispersion relationship

$$\epsilon^2 \;=\; m^2 \,+\, \vec{p}^{\,2} \tag{10.280}$$

which has a positive and a negative solution

$$\epsilon \;=\; \pm\, E(\vec{p})\,, \quad E(\vec{p}) \;=\; \sqrt{m^2 \,+\, \vec{p}^{\,2}}\,. \tag{10.281}$$

Obviously, the Dirac equation, like the Klein–Gordon equation, reproduce the classical relativistic energy–momentum relationships (10.93, 10.94)
Equation (10.278) provides us with information about the components of the bispinor wave function (10.276), namely ϕ_o and χ_o are related as follows

$$\phi_0 \quad = \quad \frac{\vec{\sigma}\cdot\vec{p}}{\epsilon - m}\,\chi_0 \tag{10.282}$$

$$\chi_o \quad = \quad \frac{\vec{\sigma}\cdot\vec{p}}{\epsilon \,+\, m}\,\phi_0\,, \tag{10.283}$$

where ϵ is defined in (10.281). These two relationships are consistent with each other. In fact, one finds using (5.234) and (10.280)

$$\frac{\vec{\sigma}\cdot\vec{p}}{\epsilon + m}\,\phi_0 \;=\; \frac{(\vec{\sigma}\cdot\vec{p})^2}{(\epsilon + m)\,(\epsilon + m)}\,\chi_0 \;=\; \frac{\vec{p}^{\,2}}{\epsilon^2 - m^2}\,\chi_0 \;=\; \chi_0\,. \tag{10.284}$$

The relationships (10.282, 10.283) imply that the bispinor part of the wave function allows only two degrees of freedom to be chosen independently. We want to show now that these degrees of freedom correspond to a spin-like property, the socalled helicity of the particle.

For our further characterization we will deal with the positive and negative energy solutions [cf. (10.281)] separately. For the positive energy solution, i.e., the solution for $\epsilon = +E(\vec{p})$, we present ϕ_o through the normalized vector

$$\phi_o \;=\; \left(\begin{array}{c} u_1 \\ u_2 \end{array} \right) \;=\; \mathbf{u} \;\in\; \mathbb{C}^2 \;,\; \mathbf{u}^\dagger \mathbf{u} \;=\; |u_1|^2 \,+\, |u_2|^2 \;=\; 1 \;. \tag{10.285}$$

The corresponding free Dirac particle is then described through the wave function

$$\Psi(\vec{p}, +|x^\mu) \;=\; \mathcal{N}_+(\vec{p}) \left(\begin{array}{c} \mathbf{u} \\ \frac{\vec{\sigma}\cdot\vec{p}}{E(\vec{p})+m}\,\mathbf{u} \end{array} \right) e^{i(\vec{p}\cdot\vec{r} \,-\, \epsilon t)} \;, \quad \epsilon \;=\; +E(\vec{p}) \;. \tag{10.286}$$

Here $\mathcal{N}_+(\vec{p})$ is a constant which will be chosen to satisfy the normalization condition

$$\Psi^\dagger(\vec{p}, +)\, \gamma^0\, \Psi^\dagger(\vec{p}, +) \;=\; 1 \;, \tag{10.287}$$

the form of which will be justified further below. Similarly, we present the negative energy solution, i.e., the solution for $\epsilon = -E(\vec{p})$, through χ_o given by

$$\chi_o \;=\; \left(\begin{array}{c} u_1 \\ u_2 \end{array} \right) \;=\; \mathbf{u} \;\in\; \mathbb{C}^2 \;,\; \mathbf{u}^\dagger \mathbf{u} \;=\; |u_1|^2 \,+\, |u_2|^2 \;=\; 1 \;. \tag{10.288}$$

corresponding to the wave function

$$\Psi(\vec{p}, -|x^\mu) \;=\; \mathcal{N}_-(\vec{p}) \left(\begin{array}{c} \frac{-\vec{\sigma}\cdot\vec{p}}{E(\vec{p})+m}\,\mathbf{u} \\ \mathbf{u} \end{array} \right) e^{i(\vec{p}\cdot\vec{r} \,-\, \epsilon t)} \;, \quad \epsilon \;=\; -E(\vec{p}) \;. \tag{10.289}$$

Here $\mathcal{N}_-(\vec{p})$ is a constant which will be chosen to satisfy the normalization condition

$$\Psi^\dagger(\vec{p}, +)\, \gamma^0\, \Psi^\dagger(\vec{p}, +) \;=\; -1 \;, \tag{10.290}$$

which differs from the normalization condition (10.287) in the minus sign on the r.h.s. The form of this condition and of (10.287) will be justified now.
First, we demonstrate that the product $\Psi^\dagger(\vec{p}, \pm)\gamma^0\Psi(\vec{p}, \pm)$, i.e., the l.h.s. of (10.287, 10.290), is invariant under Lorentz transformations. One can see this as follows: Let $\Psi(\vec{p}, \pm)$ denote the solution of a free particle moving with momentum $\vec{p}$ in the laboratory frame, and let $\Psi(0, \pm)$ denote the corresponding solution of a particle in its rest frame. The connection between the solutions, according to (10.241), is $\Psi(\vec{p}, \pm) \;=\; S\, \Psi(0, \pm)$, where S is given by (10.259). Hence,

$$\begin{aligned} \Psi^\dagger(\vec{p}, \pm)\, \gamma^0\, \Psi(\vec{p}, \pm) \quad &= \quad \left(\Psi^\dagger(0, \pm)\, S^\dagger \right) \gamma^0\, S\, \Psi(0, \pm) \\ &= \quad \Psi^\dagger(0, \pm)\, \left(\gamma^0\, S^{-1}\, \gamma^0 \right) \gamma^0\, S\, \Psi(0, \pm) \\ &= \quad \Psi^\dagger(0, \pm)\, \gamma^0\, \Psi(0, \pm). \end{aligned} \tag{10.291}$$

Note that we have used that, according to (10.268), $S^{-1} = \gamma^0 S^\dagger \gamma^0$ and, hence, $S^\dagger \;=\; \gamma^0 S^{-1} \gamma^0$.
We want to demonstrate now that the signs on the r.h.s. of (10.287, 10.290) should differ. For this purpose we consider first the positive energy solution. Employing (10.286) for $\vec{p} = 0$ yields, using γ^0 as given in (10.219) and $\mathbf{u}^\dagger \mathbf{u} \;=\; 1$ [c.f. (10.285)],

$$\Psi^\dagger(0, +)\, \gamma^0\, \Psi(0, +) \;=\; |\mathcal{N}_+(0)|^2\, (\mathbf{u}^\dagger, \quad 0)\, \gamma^0 \left(\begin{array}{c} \mathbf{u} \\ 0 \end{array} \right) \;=\; |\mathcal{N}_+(0)|^2 \;. \tag{10.292}$$

The same calculation for the negative energy wave function as given in (10.289) yields

$$\Psi^\dagger(0,-)\,\gamma^0\,\Psi(0,-) \;=\; |\mathcal{N}_-(0)|^2\,(0,\quad \mathbf{u}^\dagger)\,\gamma^0\begin{pmatrix}0\\ \mathbf{u}\end{pmatrix} \;=\; -\,|\mathcal{N}_-(0)|^2\;. \tag{10.293}$$

Obviously, this requires the choice of a negative side on the r.h.s. of (10.290) to assign a positive value to $|\mathcal{N}_-(0)|^2$. We can also conclude from our derivation

$$\mathcal{N}_\pm(0) \;=\; 1\;. \tag{10.294}$$

We want to determine now $\mathcal{N}_\pm(\vec{p})$ for arbitrary $\vec{p}$. We consider first the positive energy solution. Condition (10.287) written explicitly using (10.286) is

$$\mathcal{N}_+^2(\vec{p})\,\left((\mathbf{u}^*)^T,\;\left[\frac{\vec{\sigma}\cdot\vec{p}}{E(\vec{p})\;+\;m}\,\mathbf{u}^*\right]^T\right)\gamma^o\begin{pmatrix}\mathbf{u}\\ \frac{\vec{\sigma}\cdot\vec{p}}{E(\vec{p})+m}\,\mathbf{u}\end{pmatrix} \;=\; 1 \tag{10.295}$$

Evaluating the l.h.s. using γ^0 as given in (10.219) yields

$$\mathcal{N}_+^2(\vec{p})\,\left[\mathbf{u}^\dagger\mathbf{u}\;-\;\mathbf{u}^\dagger\,\frac{(\vec{\sigma}\cdot\vec{p})^2}{(E(\vec{p})\;+\;m)^2}\,\mathbf{u}\right] \;=\; 1\;. \tag{10.296}$$

Replacing $(\vec{\sigma}\cdot\vec{p})^2$ by $\vec{p}^2$ [c.f. (5.234)] and using the normalization of $\mathbf{u}$ in (10.285) results in

$$\mathcal{N}_+^2(\vec{p})\,\left[1\;-\;\frac{\vec{p}^2}{(E(\vec{p})\;+\;m)^2}\right] \;=\; 1 \tag{10.297}$$

from which follows

$$\mathcal{N}_+(\vec{p}) \;=\; \sqrt{\frac{(\,m\;+\;E(\vec{p})\,)^2}{(\,m\;+\;E(\vec{p})\,)^2\;-\;\vec{p}^2}}\;. \tag{10.298}$$

Noting

$$(\,m\;+\;E(\vec{p})\,)^2\;-\;\vec{p}^2 \;=\; m^2\;-\;\vec{p}^2\;+\;2mE(\vec{p})\;+\;E^2(\vec{p}) \;=\; 2(m\;+\;E(\vec{p}))\,m \tag{10.299}$$

the normalization coefficient (10.298) becomes

$$\mathcal{N}_+(\vec{p}) \;=\; \sqrt{\frac{m\;+\;E(\vec{p})}{2m}}\;. \tag{10.300}$$

This result completes the expression for the wave function (10.286).

Exercise 7.6: Show that the normalization condition

$$\mathcal{N}_+'^2(\vec{p})\,\left((\mathbf{u}^*)^T,\;\left[\frac{\vec{\sigma}\cdot\vec{p}}{E(\vec{p})\;+\;m}\,\mathbf{u}^*\right]^T\right)\begin{pmatrix}\mathbf{u}\\ \frac{\vec{\sigma}\cdot\vec{p}}{E(\vec{p})+m}\,\mathbf{u}\end{pmatrix} \;=\; 1 \tag{10.301}$$

yields the normalization coefficient

$$\mathcal{N}_+'(\vec{p}) \;=\; \sqrt{\frac{m\;+\;E(\vec{p})}{2\;E(\vec{p})}}\;. \tag{10.302}$$

We consider now the negative energy solution. Condition (10.290) written explicitly using (10.289) is

$$\mathcal{N}_-^2(\vec{p}) \left(\left[\frac{-\vec{\sigma}\cdot\vec{p}}{E(\vec{p}) + m}\, \mathbf{u}^* \right]^T , \, (\mathbf{u}^*)^T \right) \gamma^o \begin{pmatrix} \frac{-\vec{\sigma}\cdot\vec{p}}{E(\vec{p})+m}\, \mathbf{u} \\ \mathbf{u} \end{pmatrix} = -1 \tag{10.303}$$

Evaluating the l.h.s. yields

$$\mathcal{N}_-^2(\vec{p}) \left[\mathbf{u}^\dagger \frac{(\vec{\sigma}\cdot\vec{p})^2}{(E(\vec{p}) + m)^2}\, \mathbf{u} - \mathbf{u}^\dagger \mathbf{u} \right] = -1 \, . \tag{10.304}$$

This condition is, however, identical to the condition (10.296) for the normalization constant $\mathcal{N}_+(\vec{p})$ of the positive energy solution. We can, hence, conclude

$$\mathcal{N}_-(\vec{p}) = \sqrt{\frac{m + E(\vec{p})}{2m}} \tag{10.305}$$

and, thereby, have completed the determination for the wave function (10.289).
The wave functions (10.286, 10.289, 10.300) have been constructed to satisfy the free particle Dirac equation (10.275). Inserting (10.286) into (10.275) yields

$$\mathcal{H}_o\, \Psi(\vec{p}, \lambda | x^\mu) = \lambda\, E(\vec{p})\, \Psi(\vec{p}, \lambda | x^\mu) \, , \tag{10.306}$$

i.e., the wave functions constructed represent eigenstates of $\mathcal{H}_o$. The wave functions are also eigenstates of the momentum operator $i\partial_\mu$, i.e.,

$$i\partial_\mu \Psi(\vec{p}, \lambda | x^\mu) = p_\mu\, \Psi(\vec{p}, \lambda | x^\mu) \tag{10.307}$$

where $p_\mu = (\epsilon, -\vec{p})$. This can be verified expressing the space–time factor of $\Psi(\vec{p}, \lambda | x^\mu)$ in 4-vector notation, i.e., $\exp[i(\vec{p}\cdot\vec{r} - \epsilon t)] = \exp(ip_\mu x^\mu)$.

Helicity

The free Dirac particle wave functions (10.286, 10.289) are not completely specified, the two components of $\mathbf{u}$ indicate another degree of freedom which needs to be defined. This degree of freedom describes a spin–$\frac{1}{2}$ attribute. This attribute is the so-called helicity, defined as the component of the particle spin along the direction of motion. The corresponding operator which measures this observable is

$$\Lambda = \frac{1}{2}\, \sigma \cdot \frac{\vec{\hat{p}}}{|\vec{\hat{p}}|} \, . \tag{10.308}$$

Note that $\vec{\hat{p}}$ represents here an operator, not a constant vector. Rather than considering the observable (10.307) we investigate first the observable due to the simpler operator $\vec{\sigma}\cdot\vec{\hat{p}}$. We want to show that this operator commutes with $\mathcal{H}_o$ and $\vec{\hat{p}}$ to ascertain that the free particle wave function can be simultaneously an eigenvector of all three operators. The commutation property $[\vec{\sigma}\cdot\vec{\hat{p}}, \hat{p}_j] = 0\,, j = 1,2,3$ is fairly obvious. The property $[\vec{\sigma}\cdot\vec{\hat{p}}, \mathcal{H}_o] = 0$ follows from (10.233) and from the two identities

$$\begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix} \cdot \vec{\hat{p}} - \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix} \cdot \vec{\hat{p}} \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} = 0 \tag{10.309}$$

and

$$\begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix} \cdot \vec{\hat{p}} \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix} \cdot \vec{\hat{p}} - \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix} \cdot \vec{\hat{p}} \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix} \cdot \vec{\hat{p}}$$
$$= \begin{pmatrix} 0 & (\vec{\sigma} \cdot \vec{\hat{p}})^2 \\ (\vec{\sigma} \cdot \vec{\hat{p}})^2 & 0 \end{pmatrix} - \begin{pmatrix} 0 & (\vec{\sigma} \cdot \vec{\hat{p}})^2 \\ (\vec{\sigma} \cdot \vec{\hat{p}})^2 & 0 \end{pmatrix} = 0 \,. \tag{10.310}$$

We have shown altogether that the operators $\vec{\hat{p}}$, $\mathcal{H}_o$ and $\vec{\sigma} \cdot \vec{\hat{p}}$ commute with each other and, hence, can be simultaneously diagonal. States which are simultaneously eigenvectors of these three operators are also simulteneously eigenvectors of the three operators $\vec{\hat{p}}$, $\mathcal{H}_o$ and Λ defined in (10.308) above. The condition that the wave functions (10.286) are eigenfunctions of Λ as well will specify now the vectors $\mathbf{u}$.

Since helicity is defined relative to the direction of motion of a particle the characterization of $\mathbf{u}$ as an eigenvector of the helicity operator, in principle, is independent of the direction of motion of the particle. We consider first the simplest case that particles move along the x_3–direction, i.e., $\vec{p} = (0, 0, p_3)$. In this case $\Lambda = \frac{1}{2}\sigma^3$.

We assume first particles with positive energy, i.e., $\epsilon = +E(\vec{p})$. According to the definition (5.224) of σ_3 the two $\mathbf{u}$ vectors $(1, 0)^T$ and $(0, 1)^T$ are eigenstates of $\frac{1}{2}\sigma^3$ with eigenvalues $\pm\frac{1}{2}$. Therefore, the wave functions which are eigenstates of the helicity operator, are

$$\begin{aligned} \Psi(p\hat{e}_3, +, +\tfrac{1}{2}|\vec{r}, t) &= \mathcal{N}_p \begin{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \frac{p}{m + E_p} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix} e^{i(px^3 - E_p t)} \\ \Psi(p\hat{e}_3, +, -\tfrac{1}{2}|\vec{r}, t) &= \mathcal{N}_p \begin{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \frac{-p}{m + E_p} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{pmatrix} e^{i(px^3 - E_p t)} \end{aligned} \tag{10.311}$$

where $\hat{e}_3$ denotes the unit vector in the x_3-direction and where

$$E_p = \sqrt{m^2 + p^2} \,; \quad \mathcal{N}_p = \sqrt{\frac{m + E_p}{2m}} \,. \tag{10.312}$$

We assume now particles with negative energy, i.e., $\epsilon = -E(\vec{p})$. The wave functions which are eigenfunctions of the helicity operator are in this case

$$\begin{aligned} \Psi(p\hat{e}_3, -, +\tfrac{1}{2}|\vec{r}, t) &= \mathcal{N}_p \begin{pmatrix} \frac{-p}{m + E_p} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix} e^{i(px^3 + E_p t)} \\ \Psi(p\hat{e}_3, -, -\tfrac{1}{2}|\vec{r}, t) &= \mathcal{N}_p \begin{pmatrix} \frac{p}{m + E_p} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{pmatrix} e^{i(px^3 + E_p t)} \end{aligned} \tag{10.313}$$

where E_p and $\mathcal{N}_p$ are defined in (10.312).

To obtain free particle wave functions for arbitrary directions of $\vec{p}$ one can employ the wave functions (10.311, 10.313) except that the states $(1,\,0)^T$ and $(0,\,1)^T$ have to be replaced by eigenstates $\mathbf{u}_\pm(\vec{p})$ of the spin operator along the direction of $\vec{p}$. These eigenstates are obtained through a rotational transformation (5.222, 5.223) as follows

$$\mathbf{u}(\vec{p},+\frac{1}{2}) \;=\; \exp\left(-\frac{i}{2}\,\vec{\vartheta}(\vec{p})\cdot\vec{\sigma}\right)\begin{pmatrix}1\\0\end{pmatrix}\,, \tag{10.314}$$

$$\mathbf{u}(\vec{p},-\frac{1}{2}) \;=\; \exp\left(-\frac{i}{2}\,\vec{\vartheta}(\vec{p})\cdot\vec{\sigma}\right)\begin{pmatrix}0\\1\end{pmatrix} \tag{10.315}$$

where

$$\vec{\vartheta}(\vec{p}) \;=\; \frac{\hat{e}_3\times\vec{p}}{|\vec{p}|}\,\angle(\hat{e}_3,\,\vec{p}) \tag{10.316}$$

describes a rotation which aligns the x_3–axis with the direction of $\vec{p}$. [One can also express the rotation through Euler angles α, β, γ , in which case the transformation is given by (5.220).] The corresponding free particle wave functions are then

$$\Psi(\vec{p},+,+\tfrac{1}{2}|\vec{r},t) \;=\; \mathcal{N}_p\begin{pmatrix}\mathbf{u}(\vec{p},+\frac{1}{2})\\ \frac{p}{m+E_p}\mathbf{u}(\vec{p},+\frac{1}{2})\end{pmatrix}e^{i(\vec{p}\cdot\vec{r}-E_pt)} \tag{10.317}$$

$$\Psi(\vec{p},+,-\tfrac{1}{2}|\vec{r},t) \;=\; \mathcal{N}_p\begin{pmatrix}\mathbf{u}(\vec{p},-\frac{1}{2})\\ \frac{-p}{m+E_p}\mathbf{u}(\vec{p},-\frac{1}{2})\end{pmatrix}e^{i(\vec{p}\cdot\vec{r}-E_pt)} \tag{10.318}$$

$$\Psi(\vec{p},-,+\tfrac{1}{2}|\vec{r},t) \;=\; \mathcal{N}_{\surd}\begin{pmatrix}\frac{-p}{m+E_p}\mathbf{u}(\vec{p},+\frac{1}{2})\\ \mathbf{u}(\vec{p},+\frac{1}{2})\end{pmatrix}e^{i(\vec{p}\cdot\vec{r}+E_pt)} \tag{10.319}$$

$$\Psi(\vec{p},-,-\tfrac{1}{2}|\vec{r},t) \;=\; \mathcal{N}_p\begin{pmatrix}\frac{p}{m+E_p}\mathbf{u}(\vec{p},-\frac{1}{2})\\ \mathbf{u}(\vec{p},-\frac{1}{2})\end{pmatrix}e^{i(\vec{p}\cdot\vec{r}+E_pt)} \tag{10.320}$$

where E_p and $\mathcal{N}$ are again given by (10.312).

Generating Solutions Through Lorentz Transformation

The solutions (10.311, 10.312) can be obtained also by means of the Lorentz transformation (10.262) for the bispinor wave function and the transformation (10.123). For this purpose one starts from the solutions of the Dirac equation in the chiral representation (10.226, 10.229), denoted by $\tilde{}$, for an $\vec{r}$–independent wave function, i.e., a wave function which represents free particles at rest. The corresponding wave functions are determined through

$$\left(\,i\tilde{\gamma}^0\partial_t\;-\;m\,\right)\,\tilde{\Psi}(t)\;=\;0\,. \tag{10.321}$$

and are

$$\tilde{\Psi}(p=0,\,+,\;\tfrac{1}{2}|t)\;=\;\tfrac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\1\\0\end{pmatrix}e^{-imt}\,,$$

$$\tilde{\Psi}(p=0,\,+,\,-\tfrac{1}{2}|t) \;=\; \tfrac{1}{\sqrt{2}}\begin{pmatrix}0\\1\\0\\1\end{pmatrix} e^{-imt}\,,$$

$$\tilde{\Psi}(p=0,\,-,\;\tfrac{1}{2}|t) \;=\; \tfrac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\-1\\0\end{pmatrix} e^{+imt}\,,$$

$$\tilde{\Psi}(p=0,\,-,\,-\tfrac{1}{2}|t) \;=\; \tfrac{1}{\sqrt{2}}\begin{pmatrix}0\\1\\0\\-1\end{pmatrix} e^{+imt}\,. \tag{10.322}$$

The reader can readily verify that transformation of these solutions to the Dirac representationsas defined in (10.228) yields the corresponding solutions (10.311, 10.313) in the $p \to 0$ limit. This correspondence justifies the characterization $\pm, \pm\frac{1}{2}$ of the wave functions stated in (10.322).
The solutions (10.322) can be written in spinor form

$$\frac{1}{\sqrt{2}}\begin{pmatrix}\phi_o\\ \chi_o\end{pmatrix} e^{\mp imt}\,, \quad \phi_o\,,\,\chi_o \in \left\{\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}0\\1\end{pmatrix}\right\} \tag{10.323}$$

Transformation (10.262) for a boost in the x_3–direction, i.e., for $\vec{w} = (0,0,w_3)$, yields for the exponential space–time dependence according to (10.174, 10.176)

$$\mp imt \;\to i\,(\,p_3 x^3 \,\mp Et\,) \tag{10.324}$$

and for the bispinor part according to (10.262)

$$\begin{pmatrix}\phi_o\\ \chi_o\end{pmatrix} \to \begin{pmatrix} e^{\frac{1}{2}w_3\sigma^3} & 0\\ 0 & e^{-\frac{1}{2}w_3\sigma^3}\end{pmatrix}\begin{pmatrix}\phi_o\\ \chi_o\end{pmatrix} = \begin{pmatrix} e^{\frac{1}{2}w_3\sigma^3}\phi_o\\ e^{-\frac{1}{2}w_3\sigma^3}\chi_o\end{pmatrix}\,. \tag{10.325}$$

One should note that ϕ_o, χ_o are eigenstates of σ^3 with eigenvalues ± 1. Applying (10.324, 10.325) to (10.323) should yield the solutions for non-vanishing momentum p in the x_3–direction. For the resulting wave functions in the chiral representation one can use then a notation corresponding to that adopted in (10.311)

$$\tilde{\Psi}(p(w_3)\hat{e}_3,\,+,\,+\tfrac{1}{2}|\vec{r},t) \;=\; \frac{1}{\sqrt{2}}\begin{pmatrix} e^{\frac{1}{2}w_3}\begin{pmatrix}1\\0\end{pmatrix}\\ e^{-\frac{1}{2}w_3}\begin{pmatrix}1\\0\end{pmatrix}\end{pmatrix} e^{i(px^3 - E_p t)}$$

$$\tilde{\Psi}(p(w_3)\hat{e}_3,\,+,\,-\tfrac{1}{2}|\vec{r},t) \;=\; \frac{1}{\sqrt{2}}\begin{pmatrix} e^{-\frac{1}{2}w_3}\begin{pmatrix}0\\1\end{pmatrix}\\ e^{\frac{1}{2}w_3}\begin{pmatrix}0\\1\end{pmatrix}\end{pmatrix} e^{i(px^3 - E_p t)}$$

$$\tilde{\Psi}(p(w_3)\hat{e}_3,\,-,\,+\tfrac{1}{2}|\vec{r},t) \;=\; \frac{1}{\sqrt{2}}\begin{pmatrix} e^{\frac{1}{2}w_3}\begin{pmatrix}1\\0\end{pmatrix} \\ -e^{-\frac{1}{2}w_3}\begin{pmatrix}1\\0\end{pmatrix}\end{pmatrix} e^{i(px^3+E_pt)}$$

$$\tilde{\Psi}(p(w_3)\hat{e}_3,\,-,\,-\tfrac{1}{2}|\vec{r},t) \;=\; \frac{1}{\sqrt{2}}\begin{pmatrix} e^{-\frac{1}{2}w_3}\begin{pmatrix}0\\1\end{pmatrix} \\ -e^{\frac{1}{2}w_3}\begin{pmatrix}0\\1\end{pmatrix}\end{pmatrix} e^{i(px^3+E_pt)} \tag{10.326}$$

where according to (10.61) $p(w_3) = m\sinh w_3$. Transformation to the Dirac representation by means of (10.228) yields

$$\Psi(p(w_3)\hat{e}_3,\,+,\,+\tfrac{1}{2}|\vec{r},t) \;=\; \begin{pmatrix} \cosh\frac{w_3}{2}\begin{pmatrix}1\\0\end{pmatrix} \\ \sinh\frac{w_3}{2}\begin{pmatrix}1\\0\end{pmatrix}\end{pmatrix} e^{i(px^3-E_pt)}$$

$$\Psi(p(w_3)\hat{e}_3,\,+,\,-\tfrac{1}{2}|\vec{r},t) \;=\; \begin{pmatrix} \cosh\frac{w_3}{2}\begin{pmatrix}0\\1\end{pmatrix} \\ -\sinh\frac{w_3}{2}\begin{pmatrix}0\\1\end{pmatrix}\end{pmatrix} e^{i(px^3-E_pt)}$$

$$\Psi(p(w_3)\hat{e}_3,\,-,\,+\tfrac{1}{2}|\vec{r},t) \;=\; \begin{pmatrix} \sinh\frac{w_3}{2}\begin{pmatrix}1\\0\end{pmatrix} \\ \cosh\frac{w_3}{2}\begin{pmatrix}1\\0\end{pmatrix}\end{pmatrix} e^{i(px^3+E_pt)}$$

$$\Psi(p(w_3)\hat{e}_3,\,-,\,-\tfrac{1}{2}|\vec{r},t) \;=\; \begin{pmatrix} -\sinh\frac{w_3}{2}\begin{pmatrix}0\\1\end{pmatrix} \\ \cosh\frac{w_3}{2}\begin{pmatrix}0\\1\end{pmatrix}\end{pmatrix} e^{i(px^3+E_pt)} \tag{10.327}$$

Employing the hyperbolic function properties

$$\cosh\frac{x}{2} \;=\; \sqrt{\frac{\cosh x\,+\,1}{2}}\,, \quad \sinh\frac{x}{2} \;=\; \sqrt{\frac{\cosh x\,-\,1}{2}}\,, \tag{10.328}$$

the relationship (10.61) between the parameter w_3 and boost velocity v_3, and the expression (10.311) for E_p one obtains

$$\begin{aligned}\cosh\tfrac{w_3}{2} \;&=\; \frac{1}{\sqrt{2}}\sqrt{\sqrt{\frac{1}{1\,-\,v_3^2}}\,+\,1} \;=\; \frac{1}{\sqrt{2}}\sqrt{\sqrt{1\,+\,\frac{v_3^2}{1\,-\,v_3^2}}\,+\,1} \\ &=\; \sqrt{\frac{\sqrt{m^2\,+\,\frac{m^2v_3^2}{1-v_3^2}}\,+\,m}{2\,m}} \;=\; \sqrt{\frac{E_p\,+\,m}{2\,m}}\end{aligned} \tag{10.329}$$

and similarly

$$\sinh\frac{w_3}{2} \;=\; \sqrt{\frac{E_p \,-\, m}{2\,m}} \;=\; \frac{p}{\sqrt{2m\,(E_p \,+\, m)}} \tag{10.330}$$

Inserting expressions (10.329, 10.330) into (10.327), indeed, reproduces the positive energy wave functions (10.311) as well as the negative energy solutions (10.313) for $-p$. The change of sign for the latter solutions had to be expected as it was already noted for the negative energy solutions of the Klein–Gordon equation (10.168–10.176).

Invariance of Dirac Equation Revisited

At this point we like to provide a variation of the derivation of (10.243), the essential property stating the Lorentz–invariance of the Dirac equation. Actually, we will derive this equation only for infinitesimal transformations, which however, is sufficient since (1) it must hold then for any finite transformation, and since (2) the calculations following (10.243) considered solely the limit of infinitesimal transformations anyway.
The reason why we provide another derivation of (10.243) is to familiarize ourselves with a formulation of Lorentz transformations of the bispinor wave finction $\Psi(x^\mu)$ which treats the spinor and the space-time part of the wave function on the same footing. Such description will be essential for the formal description of Lorentz invariant wave equations for arbitray spin further below.
In the new derivation we consider the particle described by the wave function transformed, but not the observer. This transformation, refered to as the active transformation, expresses the system in the old coordinates. The transformation is

$$\Psi'(x^\mu) \;=\; S(L^\eta{}_\xi)\,\rho(L^\eta{}_\xi)\,\Psi(x^\mu) \tag{10.331}$$

where $S(L^\eta{}_\xi)$ denotes again the transformation acting on the bispinor character of the wave function $\Psi(x^\mu)$ and where $\rho(L^\eta{}_\xi)$ denotes the transformation acting on the space-time character of the wave function $\Psi(x^\mu)$. $\rho(L^\eta{}_\xi)$ has been defined in (10.123) above and characterized there. Such transformation had been applied by us, of course, when we generated the solutions $\Psi(\vec{p},\lambda,\Lambda|x^\mu)$ from the solutions describing particles at rest $\Psi(\vec{p}=0,\lambda,\Lambda|t)$. We expect, in general, that if $\Psi(x^\mu)$ is a solution of the Dirac equation that $\Psi'(x^\mu)$ as given in (10.331) is a solution as well. Making this expectation a postulate allows one to derive the condition (10.243) and, thereby, the proper transformation $S(L^\eta{}_\xi)$.
To show this we rewrite the Dirac equation (10.221) using (10.331)

$$\left(\, i\,S(L^\eta{}_\xi)\gamma^\mu S^{-1}(L^\eta{}_\xi)\,\rho(L^\eta{}_\xi)\partial_\mu\rho^{-1}(L^\eta{}_\xi) \;-\; m\,\right)\,\Psi'(x^\mu) \;=\; 0 \tag{10.332}$$

Here we have made use of the fact that $S(L^\eta{}_\xi)$ commutes with ∂_μ and $\rho(L^\eta{}_\xi)$ commutes with γ^μ. The fact that any such $\Psi'(x^\mu)$ is a solution of the Dirac equation allows us to conclude

$$S(L^\eta{}_\xi)\gamma^\mu S^{-1}(L^\eta{}_\xi)\,\rho(L^\eta{}_\xi)\partial_\mu\rho^{-1}(L^\eta{}_\xi) \;=\; \gamma^\nu\partial_\nu \tag{10.333}$$

which is satisfied in case that the following conditions are met

$$\begin{aligned} \rho(L^\eta{}_\xi)\partial_\mu\rho^{-1}(L^\eta{}_\xi) \quad &=\; L^\nu{}_\mu\,\partial_\nu\;; \\ \mathcal{S}(L^\eta{}_\xi)\gamma^\mu\mathcal{S}^{-1}(L^\eta{}_\xi)\,L^\nu{}_\mu \quad &=\; \gamma^\nu\;. \end{aligned} \tag{10.334}$$

We will demonstrate now that the first condition is satisfied by $\rho(L^{\eta}{}_{\xi})$. The second condition is identical to (10.243) and, of course, it is met by $\mathcal{S}(L^{\eta}{}_{\xi})$ as given in the chiral representation by (10.262).
As mentioned already we will show condition (10.334) for infinitesimal Lorentz transformations $L^{\eta}{}_{\xi}$. We will proceed by employing the generators (10.128) to express $\rho(L^{\eta}{}_{\xi})$ in its infinitesimal form and evaluate the expression

$$\begin{aligned}\left(\mathbb{1} + \epsilon\vec{\vartheta}\cdot\vec{\mathcal{J}} + \epsilon\vec{w}\cdot\vec{\mathcal{K}} \right) \partial_\mu \left(\mathbb{1} - \epsilon\vec{\vartheta}\cdot\vec{\mathcal{J}} - \epsilon\vec{w}\cdot\vec{\mathcal{K}} \right) \\ = \partial_\mu + \epsilon M^{\nu}{}_{\mu}\partial_\nu + O(\epsilon^2)\end{aligned} \tag{10.335}$$

The result will show that the matrix $M^{\nu}{}_{\mu}$ is identical to the generators of $L^{\nu}{}_{\mu}$ for the six choices $\vec{\vartheta} = (1,0,0)$, $\vec{w} = (0,0,0)$, $\vec{\vartheta} = (0,1,0)$, $\vec{w} = (0,0,0)$, $\ldots$, $\vartheta = (0,0,0)$, $\vec{w} = (0,0,1)$. Inspection of (10.335) shows that we need to demonstrate

$$[\mathcal{J}_\ell, \partial_\mu] = (J_\ell)^{\nu}{}_{\mu}\partial_\nu \; ; \quad [\mathcal{K}_\ell, \partial_\mu] = (K_\ell)^{\nu}{}_{\mu}\partial_\nu \; . \tag{10.336}$$

We will proceed with this task considering all six cases:

$$[\mathcal{J}_1, \partial_\mu] = [x^3\partial_2 - x^2\partial_3, \partial_\mu] = \begin{cases} 0 & \mu = 0 \\ 0 & \mu = 1 \\ \partial_3 & \mu = 2 \\ -\partial_2 & \mu = 3 \end{cases} \tag{10.337}$$

$$[\mathcal{J}_2, \partial_\mu] = [x^1\partial_3 - x^3\partial_1, \partial_\mu] = \begin{cases} 0 & \mu = 0 \\ -\partial_3 & \mu = 1 \\ 0 & \mu = 2 \\ \partial_1 & \mu = 3 \end{cases} \tag{10.338}$$

$$[\mathcal{J}_3, \partial_\mu] = [x^2\partial_1 - x^1\partial_2, \partial_\mu] = \begin{cases} 0 & \mu = 0 \\ \partial_2 & \mu = 1 \\ -\partial_1 & \mu = 2 \\ 0 & \mu = 3 \end{cases} \tag{10.339}$$

$$[\mathcal{K}_1, \partial_\mu] = [x^0\partial_1 + x^1\partial_0, \partial_\mu] = \begin{cases} -\partial_1 & \mu = 0 \\ -\partial_0 & \mu = 1 \\ 0 & \mu = 2 \\ 0 & \mu = 3 \end{cases} \tag{10.340}$$

$$[\mathcal{K}_2, \partial_\mu] = [x^0\partial_2 + x^2\partial_0, \partial_\mu] = \begin{cases} -\partial_2 & \mu = 0 \\ 0 & \mu = 1 \\ -\partial_0 & \mu = 2 \\ 0 & \mu = 3 \end{cases} \tag{10.341}$$

$$[\mathcal{K}_3, \partial_\mu] = [x^0\partial_3 + x^3\partial_0, \partial_\mu] = \begin{cases} -\partial_3 & \mu = 0 \\ 0 & \mu = 1 \\ 0 & \mu = 2 \\ -\partial_0 & \mu = 3 \end{cases} \tag{10.342}$$

One can readily convince oneself that these results are consistent with (10.336). We have demonstrated, therefore, that any solution $\Psi(x^\mu)$ transformed according to (10.331) is again a solution of the Dirac equation, i.e., the Dirac equation is invariant under *active* Lorentz transformations.

10.10 Dirac Particles in Electromagnetic Field

We like to provide now a description for particles governed by the Dirac equation which includes the coupling to an electromagnetic field in the minimum coupling description. Following the respective procedure developed for the Klein-Gordon equation in Sect. 10.6 we assume that the field is described through the 4-vector potential A^μ and, accordingly, we replace in the Dirac equation the momentum operator $\hat{p}_\mu = i\partial_\mu$ by $i\partial_\mu - qA_\mu$ where q is the charge of the respective particles (see Table 10.1 in Sect 10.6 above). Equivalently, we replace the operator ∂_μ by $\partial_\mu + iqA_\mu$. The Dirac equation (10.221) reads then

$$[\, i\gamma^\mu(\partial_\mu + iqA_\mu) - m\,]\,\Psi(x^\nu) = 0 \tag{10.343}$$

One may also include the electromagnetic field in the Dirac equation given in the Schrödinger form (10.233) by replacing $i\partial_t$ by (see Table 10.1) $i\partial_t - qV$ and $\hat{\vec{p}}$ by

$$\hat{\vec{\pi}} = \hat{\vec{p}} - q\vec{A}\,. \tag{10.344}$$

The Dirac equation in the Schrödinger form reads then

$$i\partial_t\Psi(x^\mu) = \left(\hat{\vec{\alpha}}\cdot\hat{\vec{\pi}} + qV + \hat{\beta}\,m\right)\Psi(x^\mu) \tag{10.345}$$

where $\hat{\vec{\alpha}}$ and $\hat{\beta}$ are defined in (10.232).

Non-Relativistic Limit

We want to consider now the Dirac equation (10.345) in the so-called non-relativistic limit in which all energies are much smaller than m, e.g., for the scalar field V in (10.345) holds

$$|qV| << m\,. \tag{10.346}$$

For this purpose we choose the decomposition

$$\Psi(x^\mu) = \begin{pmatrix} \phi(x^\mu) \\ \chi(x^\mu) \end{pmatrix}. \tag{10.347}$$

Using the notation $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)^T$ one obtains then

$$i\partial_t\phi = \vec{\sigma}\cdot\hat{\vec{\pi}}\,\chi + qV\,\phi + m\phi \tag{10.348}$$

$$i\partial_t\chi = \vec{\sigma}\cdot\hat{\vec{\pi}}\,\phi + qV\,\chi - m\chi\,. \tag{10.349}$$

We want to focus on the stationary positive energy solution. This solution exhibits a time-dependence $\exp[-i(m+\epsilon)t]$ where for ϵ holds in the non-relativistic limit $|\epsilon| << m$. Accordingly, we define

$$\phi(x^\mu) = e^{-imt}\,\Phi(x^\mu) \tag{10.350}$$

$$\chi(x^\mu) = e^{-imt}\,\mathcal{X}(x^\mu) \tag{10.351}$$

and assume that for the time-derivative of Φ and $\mathcal{X}$ holds

$$\left|\frac{\partial_t \Phi}{\Phi}\right| << m\,, \quad \left|\frac{\partial_t \mathcal{X}}{\mathcal{X}}\right| << m\,. \tag{10.352}$$

Using (10.350, 10.351) in (10.348, 10.349) yields

$$i\partial_t \Phi = \vec{\sigma}\cdot\hat{\vec{\pi}}\,\mathcal{X} + qV\,\Phi \tag{10.353}$$

$$i\partial_t \mathcal{X} = \vec{\sigma}\cdot\hat{\vec{\pi}}\,\Phi + qV\,\mathcal{X} - 2m\mathcal{X}\,. \tag{10.354}$$

The properties (10.346, 10.352) allow one to approximate (10.354)

$$0 \approx \vec{\sigma}\cdot\hat{\vec{\pi}}\,\Phi - 2m\,\mathcal{X}\,. \tag{10.355}$$

and, accordingly, one can replace $\mathcal{X}$ in (10.353) by

$$\mathcal{X} \approx \frac{\vec{\sigma}\cdot\hat{\vec{\pi}}}{2m}\,\Phi \tag{10.356}$$

to obtain a closed equation for Φ

$$i\partial_t \Phi \approx \frac{\left(\vec{\sigma}\cdot\hat{\vec{\pi}}\right)^2}{2m}\,\Phi + qV\,\Phi\,. \tag{10.357}$$

Equation (10.356), due to the m^{-1} factor, identifies $\mathcal{X}$ as the small component of the bi-spinor wave function which, henceforth, does not need to be considered anymore.
Equation (10.357) for Φ can be reformulated by expansion of $(\vec{\sigma}\cdot\hat{\vec{\pi}})^2$. For this purpose we employ the identity (5.230), derived in Sect. 5.7, which in the present case states

$$(\vec{\sigma}\cdot\hat{\vec{\pi}})^2 = \hat{\vec{\pi}}^2 + i\,\vec{\sigma}\cdot(\hat{\vec{\pi}}\times\hat{\vec{\pi}})\,. \tag{10.358}$$

For the components of $\hat{\vec{\pi}}\times\hat{\vec{\pi}}$ holds

$$\left(\hat{\vec{\pi}}\times\hat{\vec{\pi}}\right)_\ell = \epsilon_{jk\ell}\,(\,\pi_j\pi_k - \pi_k\pi_j\,) = \epsilon_{jk\ell}\,[\pi_j,\,\pi_\ell]\,. \tag{10.359}$$

We want to evaluate the latter commutator. One obtains

$$\begin{aligned}
[\pi_j,\pi_k] &= [\frac{1}{i}\partial_j + qA_j,\,\frac{1}{i}\partial_k + qA_k] \\
&= \underbrace{[\frac{1}{i}\partial_j,\,\frac{1}{i}\partial_k]}_{=0} + q\,[A_j,\,\frac{1}{i}\partial_k] + q\,[\frac{1}{i}\partial_j,\,A_k] + q^2\,\underbrace{[A_j,\,A_k]}_{=0} \\
&= \frac{q}{i}\,[A_j,\,\partial_k] + \frac{q}{i}\,[\partial_j,\,A_k]\,.
\end{aligned} \tag{10.360}$$

For an arbitrary function $f(\vec{r})$ holds

$$(\,[A_j,\,\partial_k] + [\partial_j,\,A_k]\,)\,f = (\,\partial_j A_k - A_k\partial_j + A_j\partial_k - \partial_k A_j\,)\,f\,. \tag{10.361}$$

Using

$$\partial_j A_k\, f \;=\; ((\partial_j A_k))\, f \;+\; A_k \partial_j\, f$$
$$\partial_k A_j\, f \;=\; ((\partial_k A_j))\, f \;+\; A_j \partial_k\, f$$

where $((\partial_j \cdots))$ denotes confinement of the differential operator to within the brackets $((\cdots))$, one obtains

$$(\,[A_j,\, \partial_k] \;+\; [\partial_j,\, A_k]\,)\, f \;=\; [\,((\partial_j A_k)) \;-\; ((\partial_k A_j))\,]\, f \tag{10.362}$$

or, using (10.360) and $A_\mu \;=\; (V,\, -\vec{A})$,

$$[\pi_j, \pi_k] \;=\; \frac{q}{i}\,((\,\partial_j A_k \;-\; \partial_k A_j\,)) \;=\; -\frac{q}{i}\,\left(\nabla \times \vec{A}\right)_\ell \epsilon_{jk\ell} \;=\; -\frac{q}{i} B_\ell\, \epsilon_{jk\ell} \tag{10.363}$$

where we employed $\vec{B}(\vec{r}, t) \;=\; \nabla \times \vec{A}(\vec{r}, t)$ [see (8.6)]. Equations (10.344, 10.358, 10.359, 10.363) allow us to write (10.357) in the final form

$$i\partial_t \Phi(\vec{r}, t) \;\approx\; \left[\, \frac{[\hat{\vec{p}} \;-\; q\vec{A}(\vec{r}, t)]^2}{2m} \;-\; \frac{q}{2m}\, \vec{\sigma} \cdot \vec{B}(\vec{r}, t) \;+\; q\, V(\vec{r}, t) \right] \Phi(\vec{r}, t) \tag{10.364}$$

which is referred to as the *Pauli equation*.
Comparision of (10.364) governing a two-dimensional wave function $\Phi \in \mathbb{C}^2$ with the corresponding non-relativistic Schrödinger equation (10.2) governing a one-dimensional wave function $\psi \in \mathbb{C}$, reveals a stunning feature: the Pauli equation does justice to its two-dimensional character; while agreeing in all other respects with the non-relativistic Schrödinger equation (10.2) it introduces the extra term $q\vec{\sigma} \cdot \vec{B}\, \Phi$ which describes the well-known interaction of a spin-$\frac{1}{2}$ particle with a magnetic field $\vec{B}$. In other words, the spin-$\frac{1}{2}$ which emerged in the Lorentz-invariant theory as an algebraic necessity, does not leave the theory again when one takes the non-relativistic limit, but rather remains as a steady "guest" of non-relativistic physics with the proper interaction term.
Let us consider briefly the consequences of the interaction of a spin-$\frac{1}{2}$ with the magnetic field. For this purpose we disregard the spatial degrees of freedom and assume the Schrödinger equation

$$i\partial_t \Phi(t) \;=\; q\, \vec{\sigma} \cdot \vec{B}\, \Phi(t)\,. \tag{10.365}$$

The formal solution of this equation is

$$\Phi(t) \;=\; e^{-iqt\vec{B}\cdot\vec{\sigma}}\, \Phi(0)\,. \tag{10.366}$$

Comparision of this expression with (5.222, 5.223) shows that the propagator in (10.366) can be interpreted as a rotation around the field $\vec{B}$ by an angle qtB, i.e., the interaction $q\vec{\sigma} \cdot \vec{B}$ induces a precession of the spin-$\frac{1}{2}$ around the magnetic field.

Dirac Particle in Coulomb Field - Spectrum

We want to describe now the spectrum of a relativistic electron ($q \;=\; -e$) in the Coulomb field of a nucleus with charge Ze. The respective bispinor wave function $\Psi(x^\mu) \in \mathbb{C}^4$ is described as the stationary solution of the Dirac equation (10.343) for the vector potential

$$A_\mu \;=\; (-\frac{Ze^2}{r},\, 0,\, 0,\, 0)\,. \tag{10.367}$$

For the purpose of the solution we assume the *chiral representation*, i.e, we solve

$$[\, i\tilde{\gamma}^{\mu}(\partial_{\mu} \,+\, iqA_{\mu}) \;-\; m\,]\;\tilde{\Psi}(x^{\mu}) \;=0 \tag{10.368}$$

where $\tilde{\Psi}(x^{\mu})$ and $\tilde{\gamma}^{\mu}$ are defined in (10.228) and in (10.229), respectively. Employing π_{μ} as defined in Table 10.1 one can write (10.368)

$$(\,\tilde{\gamma}^{\mu}\,\pi_{\mu} \;-\; m\,)\;\tilde{\Psi}(x^{\mu}) \;=0\,. \tag{10.369}$$

For our solution we will adopt presently a strategy which follows closely that for the spectrum of pionic atoms in Sect. 10.6. For this purpose we 'square' the Dirac equation, multiplying (10.369) from the left by $\gamma^{\nu}\pi_{\nu} + m$. This yields

$$\begin{aligned}[\, i\tilde{\gamma}^{\mu}(\partial_{\mu} \,+\, iqA_{\mu}) \;+\; m\,]\,[\, i\tilde{\gamma}^{\mu}(\partial_{\mu} \,+\, iqA_{\mu}) \;-\; m\,]\;\tilde{\Psi}(x^{\mu}) \\ =\; (\tilde{\gamma}^{\mu}\hat{\pi}_{\mu}\,\tilde{\gamma}^{\nu}\hat{\pi}_{\nu} \;-\; m^2)\,\tilde{\Psi}(x^{\mu}) \;=\; 0\,.\end{aligned} \tag{10.370}$$

Any solution of (10.368) is also a solution of (10.370), but the converse is not necessarily true. However, once a solution $\tilde{\Psi}(x^{\mu})$ of (10.370) is obtained then

$$[\, i\tilde{\gamma}^{\mu}(\partial_{\mu} \,+\, iqA_{\mu}) \;+\; m\,]\;\tilde{\Psi}(x^{\mu}) \tag{10.371}$$

is a solution of (10.369). This follows from

$$\begin{aligned}[\, i\tilde{\gamma}^{\mu}(\partial_{\mu} \,+\, iqA_{\mu}) \;+\; m\,]\,[\, i\tilde{\gamma}^{\mu}(\partial_{\mu} \,+\, iqA_{\mu}) \;-\; m\,] \\ =\; [\, i\tilde{\gamma}^{\mu}(\partial_{\mu} \,+\, iqA_{\mu}) \;-\; m\,]\,[\, i\tilde{\gamma}^{\mu}(\partial_{\mu} \,+\, iqA_{\mu}) \;+\; m\,]\end{aligned} \tag{10.372}$$

according to which follows from (10.370)

$$[\, i\tilde{\gamma}^{\mu}(\partial_{\mu} \,+\, iqA_{\mu}) \;-\; m\,]\,[\, i\tilde{\gamma}^{\mu}(\partial_{\mu} \,+\, iqA_{\mu}) \;+\; m\,]\;\tilde{\Psi}(x^{\mu}) \;=\; 0 \tag{10.373}$$

such that we can conclude that (10.371), indeed, is a solution of (10.369).
Equation (10.370) resembles closely the Klein-Gordon equation (10.180), but differs from it in an essential way. The difference arises from the term $\tilde{\gamma}^{\mu}\hat{\pi}_{\mu}\,\tilde{\gamma}^{\nu}\hat{\pi}_{\nu}$ in (10.370) for which holds

$$\tilde{\gamma}^{\mu}\hat{\pi}_{\mu}\,\tilde{\gamma}^{\nu}\hat{\pi}_{\nu} \;=\; \sum_{\mu=0}^{3}\,(\tilde{\gamma}^{\mu})^2\,\hat{\pi}_{\mu}^2 \;+\; \sum_{\substack{\mu,\nu=1\\ \mu\,\neq\nu}}\,\tilde{\gamma}^{\mu}\tilde{\gamma}^{\nu}\hat{\pi}_{\mu}\hat{\pi}_{\nu}\,. \tag{10.374}$$

The first term on the r.h.s. can be rewritten using, according to (10.230), $(\tilde{\gamma}^0)^2 = \mathbb{1}$ and $(\tilde{\gamma}^j)^2 = -\mathbb{1}$, $j = 1, 2, 3$,

$$\sum_{\mu=0}^{3}\,(\tilde{\gamma}^{\mu})^2\,\hat{\pi}_{\mu}^2 \;=\; \hat{\pi}_0^2 \;-\; \hat{\vec{\pi}}^{\,2}\,. \tag{10.375}$$

Following the algebra that connected Eqs. (5.231), (5.232) in Sect. 5.7 one can write the second term in (10.374), noting from (10.230) $\tilde{\gamma}^{\mu}\tilde{\gamma}^{\nu} = -\tilde{\gamma}^{\nu}\tilde{\gamma}^{\mu}$, $\mu \neq \nu$ and altering 'dummy' summation

indices,

$$\begin{aligned}\sum_{\substack{\mu,\nu=1\\ \mu\neq\nu}} \tilde{\gamma}^\mu\tilde{\gamma}^\nu\hat{\pi}_\mu\hat{\pi}_\nu &= \frac{1}{2}\sum_{\substack{\mu,\nu=1\\ \mu\neq\nu}} (\,\tilde{\gamma}^\mu\tilde{\gamma}^\nu\hat{\pi}_\mu\hat{\pi}_\nu \,+\, \tilde{\gamma}^\nu\tilde{\gamma}^\mu\hat{\pi}_\nu\hat{\pi}_\mu\,)\\ &= \frac{1}{4}\sum_{\substack{\mu,\nu=1\\ \mu\neq\nu}} (\,\tilde{\gamma}^\mu\tilde{\gamma}^\nu\hat{\pi}_\mu\hat{\pi}_\nu \,-\, \tilde{\gamma}^\nu\tilde{\gamma}^\mu\hat{\pi}_\mu\hat{\pi}_\nu \,+\, \tilde{\gamma}^\nu\tilde{\gamma}^\mu\hat{\pi}_\nu\hat{\pi}_\mu \,-\, \tilde{\gamma}^\mu\tilde{\gamma}^\nu\hat{\pi}_\nu\hat{\pi}_\mu\,)\\ &= \frac{1}{4}\sum_{\substack{\mu,\nu=1\\ \mu\neq\nu}} [\tilde{\gamma}^\mu,\,\tilde{\gamma}^\nu]\,[\hat{\pi}_\mu,\,\hat{\pi}_\nu] \end{aligned} \tag{10.376}$$

This expression can be simplified due to the special form (10.367) of A_μ, i.e., due to $\vec{A} = 0$. Since

$$[\hat{\pi}_\mu,\,\hat{\pi}_\nu] \;=\; 0 \quad \text{for } \mu,\nu \,=\, 1,2,3 \tag{10.377}$$

which follows readily from the definition (10.344), it holds

$$\begin{aligned}&\frac{1}{4}\sum_{\substack{\mu,\nu=1\\ \mu\neq\nu}} [\tilde{\gamma}^\mu,\,\tilde{\gamma}^\nu]\,[\hat{\pi}_\mu,\,\hat{\pi}_\nu]\\ &= \frac{1}{4}\sum_{j=1}^{3}[\tilde{\gamma}^0,\,\tilde{\gamma}^j]\,[\hat{\pi}_0,\,\hat{\pi}_j] \,+\, \frac{1}{4}\sum_{j=1}^{3}[\tilde{\gamma}^j,\,\tilde{\gamma}^0]\,[\hat{\pi}_j,\,\hat{\pi}_0]\\ &= \frac{1}{2}\sum_{j=1}^{3}[\tilde{\gamma}^0,\,\tilde{\gamma}^j]\,[\hat{\pi}_0,\,\hat{\pi}_j]\,. \end{aligned} \tag{10.378}$$

According to the definition (10.229), the commutators $[\tilde{\gamma}^0,\,\tilde{\gamma}^j]$ are

$$\begin{aligned}[\tilde{\gamma}^0,\,\tilde{\gamma}^j] &= \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}\begin{pmatrix} 0 & -\sigma^j \\ \sigma^j & 0 \end{pmatrix} - \begin{pmatrix} 0 & -\sigma^j \\ \sigma^j & 0 \end{pmatrix}\begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}\\ &= 2\begin{pmatrix} \sigma^j & 0 \\ 0 & -\sigma^j \end{pmatrix} \end{aligned} \tag{10.379}$$

The commutators $[\hat{\pi}_0,\,\hat{\pi}_j]$ in (10.378) can be evaluated using (10.367) and the definition (10.344)

$$\begin{aligned}[\hat{\pi}_0,\,\hat{\pi}_j] &= (-i\partial_t \,+\, qA_0,\,-i\partial_j] \;=\; -\,[\,(\partial_t \,+\, iqA_0)\,\partial_j \;-\; \partial_j\,(\partial_t \,+\, iqA_0)\,]\,f\\ &= i\,((\partial_j qA_0))\,f \end{aligned} \tag{10.380}$$

where $f \;=\; f(\vec{r},t)$ is a suitable test function and where $((\cdots))$ denotes the range to which the derivative is limited. Altogether, one can summarize (10.376–10.380)

$$\sum_{\substack{\mu,\nu=1\\ \mu\neq\nu}} \tilde{\gamma}^\mu\tilde{\gamma}^\nu\hat{\pi}_\mu\hat{\pi}_\nu \;=\; i\begin{pmatrix} \vec{\sigma} & 0 \\ 0 & -\vec{\sigma} \end{pmatrix}\cdot((\nabla qA_0)) \tag{10.381}$$

According to (10.367) holds

$$\nabla qA_0 \;=\; \hat{r}\,\frac{Ze^2}{r^2}\,. \tag{10.382}$$

where $\hat{r} = \vec{r}/|\vec{r}|$ is a unit vector. Combining this result with (10.381), (10.374), (10.375) the 'squared' Dirac equation (10.368) reads

$$\left[-\left(\partial_t - i\frac{Ze^2}{r}\right)^2 + \nabla^2 + i \begin{pmatrix} \vec{\sigma}\cdot\hat{r} & 0 \\ 0 & -\vec{\sigma}\cdot\hat{r} \end{pmatrix} \frac{Ze^2}{r^2} - m^2 \right] \Psi(x^\mu) = 0 \tag{10.383}$$

We seek stationary solutions of this equation. Such solutions are of the form

$$\tilde{\Psi}(x^\mu) = \tilde{\Phi}(\vec{r})\, e^{-i\epsilon t} . \tag{10.384}$$

ϵ can be interpreted as the energy of the stationary state and, hence, it is this quantity that we want to determine. Insertion of (10.384) into (10.383) yields the purely spatial four-dimensional differential equation

$$\left[\left(\epsilon + \frac{Ze^2}{r}\right)^2 + \nabla^2 + i \begin{pmatrix} \vec{\sigma}\cdot\hat{r} & 0 \\ 0 & -\vec{\sigma}\cdot\hat{r} \end{pmatrix} \frac{Ze^2}{r^2} - m^2 \right] \Phi(\vec{r}) = 0 . \tag{10.385}$$

We split the wave function into two spin-$\frac{1}{2}$ components

$$\tilde{\Psi}(\vec{r}) = \begin{pmatrix} \tilde{\phi}_+(\vec{r}) \\ \tilde{\phi}_-(\vec{r}) \end{pmatrix} \tag{10.386}$$

and obtain for the separate components $\phi_\pm(\vec{r})$

$$\left[\left(\epsilon + \frac{Ze^2}{r}\right)^2 + \nabla^2 \pm i\,\vec{\sigma}\cdot\hat{r}\,\frac{Ze^2}{r^2} - m^2 \right] \phi_\pm(\vec{r}) = 0 . \tag{10.387}$$

The expression (10.189) for the Laplacian and expansion of the term $(\cdots)^2$ result in the two-dimensional equation

$$\left[\partial_r^2 - \frac{\hat{L}^2 - Z^2e^4 \mp i\vec{\sigma}\cdot\hat{r}\, Ze^2}{r^2} + \frac{2Ze^2\epsilon}{r} + \epsilon^2 - m^2 \right] r\,\phi_\pm(\vec{r}) = 0 . \tag{10.388}$$

Except for the term $i\vec{\sigma}\cdot\hat{r}$ this equation is identical to that posed by the one-dimensional Klein-Gordon equation for pionic atoms (10.191) solved in Sect. 10.6. In the latter case, a solution of the form $\sim Y_{\ell m}(\hat{r})$ can be obtained. The term $i\vec{\sigma}\cdot\hat{r}$, however, is genuinely two-dimensional and, in fact, couples the orbital angular momentum of the electron to its spin-$\frac{1}{2}$. Accordingly, we express the solution of (10.388) in terms of states introduced in Sect. 6.5 which describe the coupling of orbital angular momentum and spin

$$\begin{aligned} &\{(\,\mathcal{Y}_{jm}(j-\tfrac{1}{2},\tfrac{1}{2}|\hat{r}), \mathcal{Y}_{jm}(j+\tfrac{1}{2},\tfrac{1}{2}|\hat{r})\,), \\ &\qquad j = \tfrac{1}{2}, \tfrac{3}{2} \ldots ; m = -j, -j+1, \ldots + j \,\} \end{aligned} \tag{10.389}$$

According to the results in Sect. 6.5 the operator $i\vec{\sigma}\cdot\hat{r}$ is block-diagonal in this basis such that only the states for identical j, m values are coupled, i.e., only the two states $\{\mathcal{Y}_{jm}(j-\frac{1}{2},\frac{1}{2}|\hat{r}), \mathcal{Y}_{jm}(j+\frac{1}{2},\frac{1}{2}|\hat{r})\}$ as given in (6.147, 6.148). We note that these states are also eigenstates of the angular

momentum operator $\hat{L}^2$ [cf. (6.151]. We select, therefore, a specific pair of total spin-orbital angular momentum quantum numbers j, m and expand

$$\phi_{\pm}(\vec{r}) \;=\; \frac{h_{\pm}(r)}{r}\,\mathcal{Y}_{jm}(j-\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \;+\; \frac{g_{\pm}(r)}{r}\,\mathcal{Y}_{jm}(j+\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \tag{10.390}$$

Using

$$\vec{\sigma}\cdot\hat{r}\,\mathcal{Y}_{jm}(j\pm\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \;=\; -\,\mathcal{Y}_{jm}(j\mp\tfrac{1}{2},\tfrac{1}{2}|\hat{r}) \tag{10.391}$$

derived in Sect. 6.5 [c.f. (6.186)], property (6.151), which states that the states $\mathcal{Y}_{jm}(j\pm\frac{1}{2},\frac{1}{2}|\hat{r})$ are eigenfunctions of $\hat{L}^2$, together with the orthonormality of these two states leads to the coupled differential equation

$$\left[\left(\partial_r^2 \;+\; \frac{2Ze^2\epsilon}{r} \;+\; \epsilon^2 \;-\; m\right)\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\right. \tag{10.392}$$
$$\left. -\,\frac{1}{r^2}\begin{pmatrix}(j-\frac{1}{2})(j+\frac{1}{2}) \;-\; Z^2e^4 & \pm iZe^2\\ \pm i\,Ze^2 & (j+\frac{1}{2})(j+\frac{3}{2}) \;-\; Z^2e^4\end{pmatrix}\right]\begin{pmatrix}h_{\pm}(r)\\ g_{\pm}(r)\end{pmatrix} \;=\; 0\,.$$

We seek to bring (10.392) into diagonal form. Any similarity transformation leaves the first term in (10.392), involving the 2×2 unit matrix, unaltered. However, such transformation can be chosen as to diagonalize the second term. Since, in the present treatment, we want to determine solely the spectrum, not the wave functions, we require only the eigenvalues of the matrices

$$B_{\pm} \;=\; \begin{pmatrix}(j-\frac{1}{2})(j+\frac{1}{2}) \;-\; Z^2e^4 & \pm iZe^2\\ \pm i\,Ze^2 & (j+\frac{1}{2})(j+\frac{3}{2}) \;-\; Z^2e^4\end{pmatrix}\,, \tag{10.393}$$

but do not explicitly consider further the wavefunctions. Obviously, the eigenvalues are independent of m. The two eigenvalues of both matrices are identical and can be written in the form

$$\lambda_1(j)\,[\lambda_1(j)\;+\;1] \qquad \text{and} \qquad \lambda_2(j)\,[\lambda_2(j)\;+\;1] \tag{10.394}$$

where

$$\lambda_1(j) \;=\; \sqrt{(j+\tfrac{1}{2})^2 \;-\; Z^2e^4} \tag{10.395}$$

$$\lambda_2(j) \;=\; \sqrt{(j+\tfrac{1}{2})^2 \;-\; Z^2e^4} \;-\; 1 \tag{10.396}$$

Equation (10.392) reads then in the diagonal representation

$$\left(\partial_r^2 \;-\; \frac{\lambda_{1,2}(j)[\lambda_{1,2}(j)+1)}{r^2} \;+\; \frac{2\epsilon Ze^2}{r} \;+\; \epsilon^2 \;-\; m^2\right) f_{1,2}(r) \;=\; 0 \tag{10.397}$$

This equation is identical to the Klein-Gordon equation for pionic atoms written in the form (10.200), except for the slight difference in the expression of $\lambda_{1,2}(j)$ as given by (10.395, 10.396) and (10.199), namely, the missing additive term $-\frac{1}{2}$, the values of the argument of $\lambda_{1,2}(j)$ being $j=\frac{1}{2},\,\frac{3}{2},\,\ldots$ rather than $\ell=0,\,1,\,\ldots$ as in the case of pionic atoms, and except for the fact that we have two sets of values for $\lambda_{1,2}(j)$, namely, $\lambda_1(j)$ and $\lambda_2(j)$.. We can, hence, conclude that the

spectrum of (10.397) is again given by eq. (10.203), albeit with some modifications. Using (10.395, 10.396) we obtain, accordingly,

$$\epsilon_1 \;=\; \frac{m}{\sqrt{1 \,+\, \frac{Z^2e^4}{(n'+1+\sqrt{(j+\frac{1}{2})^2 - Z^2e^4})^2}}} \quad ; \tag{10.398}$$

$$\epsilon_2 \;=\; \frac{m}{\sqrt{1 \,+\, \frac{Z^2e^4}{(n' \;+\; \sqrt{(j+\frac{1}{2})^2 - Z^2e^4})^2}}} \quad ; \tag{10.399}$$

$$n' = 0, 1, 2, \dots \;, \quad j = \tfrac{1}{2}, \tfrac{3}{2}, \dots \;, \quad m = -j, -j+1, \dots, j$$

where ϵ_1 corresponds to $\lambda_1(j)$ as given in (10.395) and ϵ_2 corresponds to $\lambda_2(j)$ as given in (10.396). For a given value of n' the energies ϵ_1 and ϵ_2 for identical j-values correspond to mixtures of states with orbital angular momentum $\ell = j - \frac{1}{2}$ and $\ell = j + \frac{1}{2}$. The magnitude of the relativistic effect is determined by Z^2e^4. Expanding the energies in terms of this parameter allows one to identify the relationship between the energies ϵ_1 and ϵ_2 and the non-relativistic spectrum. One obtains in case of (10.398, 10.399)

$$\epsilon_1 \;\approx\; m \,-\, \frac{mZ^2e^4}{2\,(n' \,+\, j \,+\, \frac{3}{2})^2} \,+\, O(Z^4e^8) \tag{10.400}$$

$$\epsilon_2 \;\approx\; m \,-\, \frac{mZ^2e^4}{2\,(n' \,+\, j \,+\, \frac{1}{2})^2} \,+\, O(Z^4e^8) \tag{10.401}$$

$$n' = 0, 1, 2, \dots \;, \quad j = \tfrac{1}{2}, \tfrac{3}{2}, \dots \;, \quad m = -j, -j+1, \dots, j \;.$$

These expressions can be equated with the non-relativistic spectrum. Obviously, the second term on the r.h.s. of these equations describe the binding energy. In case of non-relativistic hydrogen-type atoms, including spin-$\frac{1}{2}$, the stationary states have binding energies

$$E \;=\; -\frac{mZ^2e^4}{2n^2} \;, \quad \begin{matrix} n = 1, 2, \dots & \ell \,=\, 0, 1, \dots, n-1 \\ m \,=\, -\ell, -\ell+1, \dots, \ell & m_s \,=\, \pm\tfrac{1}{2} \end{matrix} \;. \tag{10.402}$$

In this expression n is the so-called main quantum number. It is given by $n = n' + \ell + 1$ where ℓ is the orbital angular momentum quantum number and $n' = 0, 1, \dots$ counts the nodes of the wave function. One can equate (10.402) with (10.400) and (10.401) if one attributes to the respective states the angular momentum quantum numbers $\ell = j + \frac{1}{2}$ and $\ell = j - \frac{1}{2}$. One may also state this in the following way: (10.400) corresponds to a non-relativistic state with quantum numbers n, ℓ and spin-orbital angular momentum $j = \ell - \frac{1}{2}$; (10.401) corresponds to a non-relativistic state with quantum numbers n, ℓ and spin-orbital angular momentum $j = \ell + \frac{1}{2}$. These considerations are summarized in the following equations

$$E_{\rm D}(n, \ell, j = \ell - \tfrac{1}{2}, m) \;=\; \frac{m}{\sqrt{1 \,+\, \frac{Z^2e^4}{(n - \ell + \sqrt{(\ell+1)^2 - Z^2e^4})^2}}} \quad ; \tag{10.403}$$

$$E_{\rm D}(n, \ell, j = \ell + \tfrac{1}{2}, m) \;=\; \frac{m}{\sqrt{1 \,+\, \frac{Z^2e^4}{(n - \ell - 1 + \sqrt{\ell^2 - Z^2e^4})^2}}} \quad ; \tag{10.404}$$

$$n \,=\, 1, 2, \dots \;; \quad \ell \,=\, 0, 1, \dots, n-1 \;; \quad m \,=\, -j, -j+1, \dots, j$$

spectr. notation	main quantum number n	orbital angular mom. ℓ	spin-orbital ang. mom. j	non-rel. binding energy / eV Eq. (10.402)	rel. binding energy / eV Eq. (10.405)
$1s_{\frac{1}{2}}$	1	0	$\frac{1}{2}$	-13.60583	-13.60601
$2s_{\frac{1}{2}}$	2	0	$\frac{1}{2}$	-3.40146	-3.40151
$2p_{\frac{1}{2}}$	2	1	$\frac{1}{2}$	$\Uparrow$	$\Uparrow$
$2p_{\frac{3}{2}}$	2	1	$\frac{3}{2}$	$\Uparrow$	- 3.40147
$3s_{\frac{1}{2}}$	3	0	$\frac{1}{2}$	-1.51176	-1.551178
$3p_{\frac{1}{2}}$	3	1	$\frac{1}{2}$	$\Uparrow$	$\Uparrow$
$3p_{\frac{3}{2}}$	3	1	$\frac{3}{2}$	$\Uparrow$	- 1.551177
$3d_{\frac{3}{2}}$	3	2	$\frac{3}{2}$	$\Uparrow$	$\Uparrow$
$3d_{\frac{5}{2}}$	3	2	$\frac{5}{2}$	$\Uparrow$	- 1.551176

Table 10.2:
Binding energies for the hydrogen ($Z = 1$) atom. Degeneracies are denoted by $\Uparrow$. The energies were evaluated with $m = 511.0041$ keV and $e^2 = 1/137.036$ by means of Eqs. (10.402, 10.405).

One can combine the expressions (10.403, 10.404) finally into the single formula

$$E_{\rm D}(n,\ell,j,m) = \frac{m}{\sqrt{1 + \frac{Z^2e^4}{(n-j-\frac{1}{2}+\sqrt{(j+\frac{1}{2})^2 - Z^2e^4})^2}}} \qquad \begin{array}{l} n = 1,2,\ldots \\ \ell = 0,1,\ldots,n-1 \\ j = \left\{ \begin{array}{ll} \frac{1}{2} & \text{for } \ell = 0 \\ \ell \pm \frac{1}{2} & \text{otherwise} \end{array} \right. \\ m = -j, -j+1, \ldots, j \end{array} \tag{10.405}$$

In order to demonstrate relativistic effects in the spectrum of the hydrogen atom we compare in Table 10.2 the non-relativistic [cf. (10.402)] and the relativistic [cf. (10.405)] spectrum of the hydrogen atom. The table entries demonstrate that the energies as given by the expression (10.405) in terms of the non-relativistic quantum numbers n, ℓ relate closely to the corresponding non-relativistic states, in fact, the non-relativistic and relativistic energies are hardly discernible. The reason is that the mean kinetic energy of the electron in the hydrogen atom, is in the range of 10 eV, i.e., much less than the rest mass of the electron (511 keV). However, in case of heavier nuclei the kinetic energy of bound electrons in the ground state scales with the nuclear charge Z like Z^2 such that in case $Z = 100$ the kinetic energy is of the order of the rest mass and relativistic effects become important. This is clearly demonstrated by the comparision of non-relativistic and relativistic spectra of a hydrogen-type atom with $Z = 100$ in Table 10.3.

Of particular interest is the effect of spin-orbit coupling which removes, for example, the non-relativistic degeneracy for the six $2p$ states of the hydrogen atom: in the present, i.e., relativistic, case these six states are split into energetically different $2p_{\frac{1}{2}}$ and $2p_{\frac{3}{2}}$ states. The $2p_{\frac{1}{2}}$ states with $j = \frac{1}{2}$ involve two degenerate states corresponding to $\mathcal{Y}_{\frac{1}{2}m}(1, \frac{1}{2}|\hat{r})$ for $m = \pm\frac{1}{2}$, the $2p_{\frac{3}{2}}$ states with $j = \frac{3}{2}$ involve four degenerate states corresponding to $\mathcal{Y}_{\frac{3}{2}m}(1, \frac{1}{2}|\hat{r})$ for $m = \pm\frac{1}{2}, \pm\frac{3}{2}$.

spectr. notation	main quantum number n	orbital angular mom. ℓ	spin-orbital ang. mom. j	non-rel. binding energy / keV Eq. (10.402)	rel. binding energy / keV Eq. (10.405)
$1s_{\frac{1}{2}}$	1	0	$\frac{1}{2}$	-136.1	-161.6
$2s_{\frac{1}{2}}$	2	0	$\frac{1}{2}$	-34.0	-42.1
$2p_{\frac{1}{2}}$	2	1	$\frac{1}{2}$	$\Uparrow$	$\Uparrow$
$2p_{\frac{3}{2}}$	2	1	$\frac{3}{2}$	$\Uparrow$	- 35.2
$3s_{\frac{1}{2}}$	3	0	$\frac{1}{2}$	-15.1	-17.9
$3p_{\frac{1}{2}}$	3	1	$\frac{1}{2}$	$\Uparrow$	$\Uparrow$
$3p_{\frac{3}{2}}$	3	1	$\frac{3}{2}$	$\Uparrow$	- 15.8
$3d_{\frac{3}{2}}$	3	2	$\frac{3}{2}$	$\Uparrow$	$\Uparrow$
$3d_{\frac{5}{2}}$	3	2	$\frac{5}{2}$	$\Uparrow$	- 15.3

Table 10.3:
Binding energies for the hydrogen-type ($Z = 100$) atom. Degeneracies are denoted by $\Uparrow$. The energies were evaluated with $m = 511.0041$ keV and $e^2 = 1/137.036$ by means of Eqs. (10.402, 10.405).

In order to investigate further the deviation between relativistic and non-relativistic spectra of hydrogen-type atoms we expand the expression (10.405) to order $O(Z^4e^8)$. Introducing $\alpha = Z^2e^4$ and $\beta = j + \frac{1}{2}$ (10.405) reads

$$\frac{1}{\sqrt{1 + \frac{\alpha}{(n - \beta + \sqrt{\beta^2 - \alpha})^2}}} \tag{10.406}$$

The expansion (10.206) provides in the present case

$$E_{\mathrm{D}}(n, \ell, j, m) \approx m - \frac{mZ^2e^4}{2n^2} - \frac{mZ^4e^8}{2n^3}\left[\frac{1}{j + \frac{1}{2}} - \frac{3}{4n}\right] + O(Z^6e^{12}) . \tag{10.407}$$

This expression allows one, for example, to estimate the difference between the energies of the states $2p_{\frac{3}{2}}$ and $2p_{\frac{1}{2}}$ (cf. Tables 10.2,10.3). It holds for $n = 2$ and $j = \frac{3}{2}, \frac{1}{2}$

$$E\,(2p_{\frac{3}{2}}) - E\,(2p_{\frac{1}{2}}) \approx -\frac{mZ^4e^8}{2 \cdot 2^3}\left[\frac{1}{2} - 1\right] = \frac{mZ^4e^8}{32} . \tag{10.408}$$

Radial Dirac Equation

We want to determine now the wave functions for the stationary states of a Dirac particle in a 4-vector potential

$$A_\mu = (V(r), 0, 0, 0) \tag{10.409}$$

where $V(r)$ is spherically symmetric. An example for such potential is the Coulomb potential $V(r) = -Ze^2/r$ considered further below. We assume for the wave function the stationary state

form

$$\Psi(x^\mu) \;=\; e^{-i\epsilon t} \left(\begin{array}{c} \Phi(\vec{r}) \\ \mathcal{X}(\vec{r}) \end{array} \right) , \tag{10.410}$$

where $\Phi(\vec{r})$, $\mathcal{X}(\vec{r}) \;\in\; \mathbb{C}^2$ describe the spatial and spin-$\frac{1}{2}$ degrees of freedom, but are time-independent. The Dirac equation reads then, according to (10.232, 10.345),

$$\vec{\sigma} \cdot \hat{\vec{p}}\, \mathcal{X} \;+\; m\, \Phi \;+\; V(r)\, \Phi \quad = \quad \epsilon\, \Phi \tag{10.411}$$

$$\vec{\sigma} \cdot \hat{\vec{p}}\, \Phi \;-\; m \mathcal{X} \;+\; V(r)\, \mathcal{X} \quad = \quad \epsilon\, \mathcal{X} \tag{10.412}$$

In this equation a coupling between the wave functions $\Phi(\vec{r})$ and $\mathcal{X}(\vec{r})$ arises due to the term $\vec{\sigma} \cdot \hat{\vec{p}}$. This term has been discussed in detail in Sect. 6.5 [see, in particular, pp. 168]: the term is a scalar (rank zero tensor) in the space of the spin-angular momentum states $\mathcal{Y}_{jm}(j \pm \frac{1}{2}, \frac{1}{2}|\hat{r})$ introduced in Sect. 6.5, i.e., the term is block-diagonal in the space spanned by the states $\mathcal{Y}_{jm}(j \pm \frac{1}{2}, \frac{1}{2}|\hat{r})$ and does not couple states with different j, m-values; $\vec{\sigma} \cdot \hat{\vec{p}}$ has odd parity and it holds [c.f. (6.197, 6.198)]

$$\vec{\sigma} \cdot \hat{\vec{p}}\, f(r)\, \mathcal{Y}_{jm}(j + \tfrac{1}{2}, \tfrac{1}{2}|\hat{r}) \quad = \quad i \left[\partial_r \;+\; \frac{j + \frac{3}{2}}{r} \right] f(r)\, \mathcal{Y}_{jm}(j - \tfrac{1}{2}, \tfrac{1}{2}|\hat{r}) \tag{10.413}$$

$$\vec{\sigma} \cdot \hat{\vec{p}}\, g(r)\, \mathcal{Y}_{jm}(j - \tfrac{1}{2}, \tfrac{1}{2}|\hat{r}) \quad = \quad i \left[\partial_r \;+\; \frac{\frac{1}{2} - j}{r} \right] g(r)\, \mathcal{Y}_{jm}(j + \tfrac{1}{2}, \tfrac{1}{2}|\hat{r}) \; . \tag{10.414}$$

These equations can be brought into a more symmetric form using

$$\partial_r \;+\; \frac{1}{r} \;=\; \frac{1}{r}\, \partial_r\, r$$

which allows one to write (10.413, 10.414)

$$\vec{\sigma} \cdot \hat{\vec{p}}\, r f(r)\, \mathcal{Y}_{jm}(j + \tfrac{1}{2}, \tfrac{1}{2}|\hat{r}) \quad = \quad i \left[\partial_r \;+\; \frac{j + \frac{1}{2}}{r} \right] r\, f(r)\, \mathcal{Y}_{jm}(j - \tfrac{1}{2}, \tfrac{1}{2}|\hat{r}) \tag{10.415}$$

$$\vec{\sigma} \cdot \hat{\vec{p}}\, r g(r)\, \mathcal{Y}_{jm}(j - \tfrac{1}{2}, \tfrac{1}{2}|\hat{r}) \quad = \quad i \left[\partial_r \;-\; \frac{j + \frac{1}{2}}{r} \right] r\, g(r)\, \mathcal{Y}_{jm}(j + \tfrac{1}{2}, \tfrac{1}{2}|\hat{r}) \; . \tag{10.416}$$

The differential equations (10.411, 10.412) are four-dimensional with $\vec{r}$-dependent wave functions. The arguments above allow one to eliminate the angular dependence by expanding $\Phi(\vec{r})$ and $\mathcal{X}(\vec{r})$ in terms of $\mathcal{Y}_{jm}(j + \frac{1}{2}, \frac{1}{2}|\hat{r})$ and $\mathcal{Y}_{jm}(j - \frac{1}{2}, \frac{1}{2}|\hat{r})$, i.e.,

$$\left(\begin{array}{c} \Phi(\vec{r}) \\ \mathcal{X}(\vec{r}) \end{array} \right) \;=\; \left(\begin{array}{c} \dfrac{a(r)}{r}\, \mathcal{Y}_{jm}(j + \frac{1}{2}, \frac{1}{2}|\hat{r}) \;+\; \dfrac{b(r)}{r}\, \mathcal{Y}_{jm}(j - \frac{1}{2}, \frac{1}{2}|\hat{r}) \\ \dfrac{c(r)}{r}\, \mathcal{Y}_{jm}(j + \frac{1}{2}, \frac{1}{2}|\hat{r}) \;+\; \dfrac{d(r)}{r}\, \mathcal{Y}_{jm}(j - \frac{1}{2}, \frac{1}{2}|\hat{r}) \end{array} \right) . \tag{10.417}$$

In general, such expansion must include states with all possible j, m values. Presently, we consider the case that only states for one specific j, m pair contribute. Inserting (10.417) into (10.411, 10.412), using (10.415, 10.416), the orthonormality property (6.157), and multiplying by r results in the following two independent pairs of coupled differential equations

$$\begin{aligned} i\left[\partial_r - \frac{j+\frac{1}{2}}{r}\right] d(r) + [\quad m + V(r) - \epsilon\,]\, a(r) &= 0 \\ i\left[\partial_r + \frac{j+\frac{1}{2}}{r}\right] a(r) + [\,-m + V(r) - \epsilon\,]\, d(r) &= 0 \end{aligned} \tag{10.418}$$

and

$$\begin{aligned} i\left[\partial_r + \frac{j+\frac{1}{2}}{r}\right] c(r) + [\quad m + V(r) - \epsilon\,]\, b(r) &= 0 \\ i\left[\partial_r - \frac{j+\frac{1}{2}}{r}\right] b(r) + [\,-m + V(r) - \epsilon\,]\, c(r) &= 0\,. \end{aligned} \tag{10.419}$$

Obviously, only $a(r), d(r)$ are coupled and $b(r), c(r)$ are coupled. Accordingly, there exist two independent solutions (10.417) of the form

$$\begin{pmatrix} \Phi(\vec{r}) \\ \mathcal{X}(\vec{r}) \end{pmatrix} = \begin{pmatrix} i\,\dfrac{f_1(r)}{r}\,\mathcal{Y}_{jm}(j+\frac{1}{2},\frac{1}{2}|\hat{r}) \\ -\dfrac{g_1(r)}{r}\,\mathcal{Y}_{jm}(j-\frac{1}{2},\frac{1}{2}|\hat{r}) \end{pmatrix} \tag{10.420}$$

$$\begin{pmatrix} \Phi(\vec{r}) \\ \mathcal{X}(\vec{r}) \end{pmatrix} = \begin{pmatrix} i\,\dfrac{f_2(r)}{r}\,\mathcal{Y}_{jm}(j-\frac{1}{2},\frac{1}{2}|\hat{r}) \\ -\dfrac{g_2(r)}{r}\,\mathcal{Y}_{jm}(j+\frac{1}{2},\frac{1}{2}|\hat{r}) \end{pmatrix} \tag{10.421}$$

where the factors i and -1 have been introduced for convenience. According to (10.418) holds for $f_1(r)$, $g_1(r)$

$$\begin{aligned} \left[\partial_r - \frac{j+\frac{1}{2}}{r}\right] g_1(r) + [\,\epsilon - m - V(r)\,]\, f_1(r) &= 0 \\ \left[\partial_r + \frac{j+\frac{1}{2}}{r}\right] f_1(r) - [\,\epsilon + m - V(r)\,]\, g_1(r) &= 0 \end{aligned} \tag{10.422}$$

and for $f_2(r)$, $g_2(r)$

$$\begin{aligned} \left[\partial_r + \frac{j+\frac{1}{2}}{r}\right] g_2(r) + [\,\epsilon - m - V(r)\,]\, f_2(r) &= 0 \\ \left[\partial_r - \frac{j+\frac{1}{2}}{r}\right] f_2(r) - [\,\epsilon + m - V(r)\,]\, g_2(r) &= 0 \end{aligned} \tag{10.423}$$

Equations (10.422) and (10.423) are identical, except for the opposite sign of the term $(j + \frac{1}{2})$; the equations determine, together with the appropriate boundary conditions at $r = 0$ and $r \to \infty$, the radial wave functions for Dirac particles in the potential (10.409).

Dirac Particle in Coulomb Field - Wave Functions

We want to determine now the wave functions of the stationary states of hydrogen-type atoms which correspond to the energy levels (10.405). We assume the 4-vector potential of pure Coulomb type (10.367) which is spherically symmetric such that equations (10.422, 10.423) apply for $V(r) = -Ze^2/r$. Equation (10.422) determines solutions of the form (10.420). In the non-relativistic limit, Φ in (10.420) is the large component and $\mathcal{X}$ is the small component. Hence, (10.422) corresponds to states

$$\Psi(x^\mu) \approx \begin{pmatrix} i\,\dfrac{f_1(r)}{r}\,\mathcal{Y}_{jm}(j+\frac{1}{2},\frac{1}{2}|\hat{r}) \\ 0 \end{pmatrix}, \tag{10.424}$$

i.e., to states with angular momentum $\ell = j + \frac{1}{2}$. According to the discussion of the spectrum (10.405) of the relativistic hydrogen atom the corresponding states have quantum numbers $n = 1, 2, \ldots, \ell = 0, 1, \ldots, n-1$. Hence, (10.422) describes the states $2p_{\frac{1}{2}}$, $3p_{\frac{1}{2}}$, $3d_{\frac{3}{2}}$, etc. Similarly, (10.423), determining wave functions of the type (10.421), i.e., in the non-relativistic limit wave functions

$$\Psi(x^\mu) \approx \begin{pmatrix} i\,\dfrac{f_2(r)}{r}\mathcal{Y}_{jm}(j-\frac{1}{2},\frac{1}{2}|\hat{r}) \\ 0 \end{pmatrix}, \tag{10.425}$$

covers states with angular momentum $\ell = j - \frac{1}{2}$ and, correspodingly the states $1s_{\frac{1}{2}}$, $2s_{\frac{1}{2}}$, $2p_{\frac{3}{2}}$, $3s_{\frac{1}{2}}$, $3p_{\frac{3}{2}}$, $3d_{\frac{5}{2}}$, etc.
We consider first the solution of (10.422). The solution of (10.422) follows in this case from the same procedure as that adopted for the radial wave function of the non-relativistic hydrogen-type atom. According to this procedure, one demonstrates first that the wave function at $r \to 0$ behaves as r^γ for some suitable γ, one demonstrates then that the wave functions for $r \to \infty$ behaves as $\exp(-\mu r)$ for some suitable μ, and obtains finally a polynomial function p(r) such that $r^\gamma \exp(-\mu r) p(r)$ solves (10.422); enforcing the polynomial to be of finite order leads to discrete eigenvalues ϵ, namely, the ones given in (10.405).

Behaviour at $r \to 0$

We consider first the behaviour of the solutions $f_1(r)$ and $g_1(r)$ of (10.422) near $r = 0$. We note that (10.422), for small r, can be written

$$\begin{aligned} \left[\partial_r - \frac{j+\frac{1}{2}}{r}\right] g_1(r) + \frac{Ze^2}{r} f_1(r) &= 0 \\ \left[\partial_r + \frac{j+\frac{1}{2}}{r}\right] f_1(r) - \frac{Ze^2}{r} g_1(r) &= 0\,. \end{aligned} \tag{10.426}$$

Setting

$$f_1(r) \underset{\to 0}{\sim} a\, r^{\gamma}\,, \quad g_1(r) \underset{\to 0}{\sim} b\, r^{\gamma} \tag{10.427}$$

yields

$$\begin{aligned} \gamma\, b\, r^{\gamma-1} \;-\; (j+\tfrac{1}{2})\, b\, r^{\gamma-1} \;+\; Ze^2\, a\, r^{\gamma-1} &= 0 \\ \gamma\, a\, r^{\gamma-1} \;+\; (j+\tfrac{1}{2})\, a\, r^{\gamma-1} \;-\; Ze^2\, b\, r^{\gamma-1} &= 0\,. \end{aligned} \tag{10.428}$$

or

$$\begin{pmatrix} \gamma + (j+\frac{1}{2}) & -Ze^2 \\ Ze^2 & \gamma - (j+\frac{1}{2}) \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = 0\,. \tag{10.429}$$

This equation poses an eigenvalue problem (eigenvalue $-\gamma$) for proper γ values. One obtains $\gamma = \pm\sqrt{(j+\frac{1}{2})^2 - Z^2e^4}$. The assumed r-dependence in (10.427) makes only the positive solution possible. We have, hence, determined that the solutions $f_1(r)$ and $g_1(r)$, for small r, assume the r-dependence in (10.427) with

$$\gamma = \sqrt{(j+\frac{1}{2})^2 - Z^2e^4}\,. \tag{10.430}$$

Note that the exponent in (10.427), in case $j+\frac{1}{2})^2 < Ze^2$, becomes imaginary. Such r-dependence would make the expectation value of the potential

$$\int r^2 dr\, \rho(\vec{r})\, \frac{1}{r} \tag{10.431}$$

infinite since, according to (10.266, 10.267, 10.420), for the particle density holds then

$$\rho(\vec{r}) \;\sim\; |r^{\gamma-1}|^2 \;=\; \frac{1}{r^2}\,. \tag{10.432}$$

Behaviour at $r \to \infty$

For very large r values (10.422) becomes

$$\begin{aligned} \partial_r\, g_1(r) &= -(\,\epsilon - m,)\, f_1(r) \\ \partial_r\, f_1(r) &= \quad (\,\epsilon + m\,)\, g_1(r) \end{aligned} \tag{10.433}$$

Iterating this equation once yields

$$\begin{aligned} \partial_r^2\, g_1(r) &= (\,m^2 - \epsilon^2\,)\, g_1(r) \\ \partial_r^2\, f_1(r) &= (\,m^2 - \epsilon^2\,)\, f_1(r) \end{aligned} \tag{10.434}$$

The solutions of these equations are $f_1, g_1 \sim \exp(\pm\sqrt{m^2-\epsilon^2}\, r)$. Only the exponentially decaying solution is admissable and, hence, we conclude

$$f_1(r) \underset{\to \infty}{\sim} e^{-\mu r}\,, \quad g_1(r) \underset{\to \infty}{\sim} e^{-\mu r}\,, \quad \mu = \sqrt{m^2 - \epsilon^2} \tag{10.435}$$

For bound states holds $\epsilon < m$ and, hence, μ is real. Let us consider then for the solution of (10.434)

$$f_1(r) \;=\; \sqrt{m+\epsilon}\, a\, e^{-\mu\, r}\;, \quad g_1(r) \;=\; -\sqrt{m-\epsilon}\, a\, e^{-\mu r}\;. \tag{10.436}$$

Insertion into (10.434) results in

$$\begin{aligned} (m-\epsilon)\sqrt{m+\epsilon}\, a \;-\; (m-\epsilon)\sqrt{m+\epsilon}\, a \quad &=\quad 0 \\ (m+\epsilon)\sqrt{m-\epsilon}\, a \;-\; (m+\epsilon)\sqrt{m-\epsilon}\, a \quad &=\quad 0 \end{aligned} \tag{10.437}$$

which is obviously correct.

Solution of the Radial Dirac Equation for a Coulomb Potential

To solve (10.422) for the Coulomb potential $V(r) = -Ze^2/r$ We assume a form for the solution which is adopted to the asymptotic solution (10.436). Accordingly, we set

$$f_1(r) \quad = \qquad \sqrt{m+\epsilon}\; e^{-\mu r}\; \tilde{f}_1(r) \tag{10.438}$$

$$g_1(r) \quad = \quad -\sqrt{m-\epsilon}\; e^{-\mu r}\; \tilde{g}_1(r) \tag{10.439}$$

where μ is given in (10.435). Equation (10.422) leads to

$$\begin{aligned} -\sqrt{m-\epsilon}\,[\partial_r \;-\; \frac{j+\frac{1}{2}}{r}]\,\tilde{g}_1 \;+\; \sqrt{m+\epsilon}\;\frac{Ze^2}{r}\;\tilde{f}_1 \;+& \\ (m-\epsilon)\,\sqrt{m+\epsilon}\;\tilde{g}_1 \;-\; (m-\epsilon)\,\sqrt{m+\epsilon}\;\tilde{f}_1 \;&=\; 0 \end{aligned} \tag{10.440}$$

$$\begin{aligned} \sqrt{m+\epsilon}\,[\partial_r \;+\; \frac{j+\frac{1}{2}}{r}]\,\tilde{f}_1 \;+\; \sqrt{m-\epsilon}\;\frac{Ze^2}{r}\;\tilde{g}_1 \;-& \\ (m+\epsilon)\sqrt{m-\epsilon}\tilde{f}_1 \;+\; (m+\epsilon)\sqrt{m-\epsilon}\tilde{g}_1 \;&=\; 0 \end{aligned} \tag{10.441}$$

The last two terms on the l.h.s. of both (10.440) and (10.441) correspond to (10.437) where they cancelled in case $\tilde{f}_1 = \tilde{g}_1 = a$. In the present case the functions $\tilde{f}_1$ and $\tilde{g}_1$ cannot be chosen identical due to the terms in the differential equations contributing for finite r. However, without loss of generality we can choose

$$\tilde{f}_1(r) \;=\; \phi_1(r) \,+\, \phi_2(r)\;, \quad \tilde{g}_1(r) \;=\; \phi_1(r) \,-\, \phi_2(r) \tag{10.442}$$

which leads to a partial cancellation of the asymptotically dominant terms. We also introduce the new variable

$$\rho \;=\; 2\mu\, r\;. \tag{10.443}$$

From this results after a little algebra

$$[\partial_\rho \;-\; \frac{j+\frac{1}{2}}{\rho}]\,(\phi_1 \;-\; \phi_2) \;-\; \sqrt{\frac{m+\epsilon}{m-\epsilon}}\;\frac{Ze^2}{\rho}\;(\phi_1 \,+\phi_2) \;+\; \phi_2 \;=\; 0 \tag{10.444}$$

$$[\partial_\rho \;+\; \frac{j+\frac{1}{2}}{\rho}]\,(\phi_1 \;+\; \phi_2) \;+\; \sqrt{\frac{m-\epsilon}{m+\epsilon}}\;\frac{Ze^2}{\rho}\;(\phi_1 \;-\; \phi_2) \;-\; \phi_2 \;=\; 0\;. \tag{10.445}$$

Addition and subtraction of these equations leads finally to the following two coupled differential equations for ϕ_1 and ϕ_2

$$\partial_\rho \phi_1 \, + \, \frac{j+\frac{1}{2}}{\rho}\,\phi_2 \, - \, \frac{\epsilon Z e^2}{\sqrt{m^2-\epsilon^2}\,\rho}\,\phi_1 \, - \, \frac{mZe^2}{\sqrt{m^2-\epsilon^2}\,\rho}\,\phi_2 \; = \; 0 \tag{10.446}$$

$$\partial_\rho \phi_2 \, + \, \frac{j+\frac{1}{2}}{\rho}\,\phi_1 \, + \, \frac{mZe^2}{\sqrt{m^2-\epsilon^2}\,\rho}\,\phi_1 \, + \, \frac{\epsilon Z e^2}{\sqrt{m^2-\epsilon^2}\,\rho}\,\phi_2 \, - \, \phi_2 \; = \; 0 \tag{10.447}$$

We seek solutions of (10.446 , 10.447) of the form

$$\phi_1(\rho) \quad = \quad \rho^\gamma \sum_{s=0}^{n'} \alpha_s \, \rho^2 \tag{10.448}$$

$$\phi_2(\rho) \quad = \quad \rho^\gamma \sum_{s=0}^{n'} \beta_s \, \rho^2 \tag{10.449}$$

for γ given in (10.430) which conform to the proper $r \, \to \, 0$ behaviour determined above [c.f. (10.426–10.430)]. Inserting (10.448, 10.449) into (10.446, 10.447) leads to

$$\sum_s \left[(s+\gamma)\,\alpha_s \, + \, (j+\tfrac{1}{2})\,\beta_s \, - \, \frac{\epsilon Z e^2}{\sqrt{m^2-\epsilon^2}}\,\alpha_s \right.$$
$$\left. - \, \frac{mZe^2}{\sqrt{m^2-\epsilon^2}}\,\beta_s \right] \, \rho^{s+\gamma-1} \; = 0 \tag{10.450}$$

$$\sum_s \left[(s+\gamma)\,\beta_2 \, + \, (j+\tfrac{1}{2})\,\alpha_s \, + \, \frac{mZe^2}{\sqrt{m^2-\epsilon^2}}\,\alpha_s \right.$$
$$\left. + \, \frac{\epsilon Z e^2}{\sqrt{m^2-\epsilon^2}}\,\beta_s \, - \, \beta_{s-1} \right] \; = 0 \tag{10.451}$$

From (10.450) follows

$$\frac{\alpha_s}{\beta_s} \; = \; \frac{\frac{mZe^2}{\sqrt{m^2-\epsilon^2}} \, - \, (j+\frac{1}{2})}{s \, + \, \gamma \, - \, \frac{\epsilon Z e^2}{\sqrt{m^2-\epsilon^2}}} \; . \tag{10.452}$$

From (10.451) follows

$$\begin{aligned} \beta_{s-1} \quad &= \quad \left(s \, + \, \gamma \, + \, \frac{\epsilon Z e^2}{\sqrt{m^2-\epsilon^2}} \right) \beta_s \, + \, \frac{\frac{m^2 Z^2 e^4}{m^2-\epsilon^2} \, - \, (j+\frac{1}{2})^2}{s \, + \, \gamma \, - \, \frac{\epsilon Z e^2}{\sqrt{m^2-\epsilon^2}}}\,\beta_s \\ &= \quad \frac{(s+\gamma)^2 \, + \, Z^2 e^4 \, - \, (j+\frac{1}{2})^2}{s \, + \, \gamma \, - \, \frac{\epsilon Z e^2}{\sqrt{m^2-\epsilon^2}}}\,\beta_s \; . \end{aligned} \tag{10.453}$$

Using (10.430) one can write this

$$\beta_s \; = \; \frac{s \, + \, \gamma \, - \, \frac{\epsilon Z e^2}{\sqrt{m^2-\epsilon^2}}}{s\,(s \, + \, 2\gamma)}\,\beta_{s-1} \; . \tag{10.454}$$

Defining

$$s_o \;=\; \frac{\epsilon Z e^2}{\sqrt{m^2-\epsilon^2}} \;-\; \gamma \tag{10.455}$$

one obtains

$$\begin{aligned}
\beta_s &= \frac{s-s_o}{s(s+2\gamma)}\,\beta_{s-1} \\
&= \frac{(s-1-s_o)(s-s_o)}{(s-1)s\,(s-1+2\gamma)(s+2\gamma)}\,\beta_{s-2} \\
&\;\;\vdots \\
&= \frac{(1-s_o)(2-s_o)\ldots(s-s_o)}{s!\,(2\gamma+1)(2\gamma+2)\cdots(2\gamma+s)}\,\beta_0
\end{aligned} \tag{10.456}$$

From (10.452) follows

$$\alpha_s \;=\; \frac{j+\frac{1}{2}\;-\;\frac{mZe^2}{\sqrt{m^2-\epsilon^2}}}{s_o}\;\frac{(-s_o)(1-s_o)(2-s_o)\ldots(s-s_o)}{s!\,(2\gamma+1)(2\gamma+2)\cdots(2\gamma+s)}\,\beta_0 \tag{10.457}$$

One can relate the polynomials $\phi_1(\rho)$ and $\phi_2(\rho)$ defined through (10.448, 10.449) and (10.456, 10.457) with the *confluent hypergeometric functions*

$$F(a,c;x) \;=\; 1 \;+\; \frac{a}{c}\,x \;+\; \frac{a(a+1)}{c(c+1)}\,\frac{x^2}{2!} \;+\; \ldots \tag{10.458}$$

or, equivalently, with the associated *Laguerre polynomials*

$$L_n^{(\alpha)} \;=\; F(-n,\alpha+1,x)\;. \tag{10.459}$$

It holds

$$\phi_1(\rho) \;=\; \beta_0\,\frac{j+\frac{1}{2}\;-\;\frac{mZe^2}{\sqrt{m^2-\epsilon^2}}}{s_o}\;\rho^\gamma\,F(-s_o,2\gamma+1;\rho) \tag{10.460}$$

$$\phi_2(\rho) \;=\; \beta_0\,\rho^\gamma\,F(1-s_o,2\gamma+1;\rho)\;. \tag{10.461}$$

In order that the wave functions remain normalizable the power series (10.448, 10.449) must be of finite order. This requires that all coefficients α_s and β_s must vanish for $s \geq n'$ for some $n' \in \mathbb{N}$. The expressions (10.456) and (10.457) for β_s and α_s imply that s_o must then be an integer, i.e., $s_o = n'$. According to the definitions (10.430, 10.455) this confinement of s_o implies discrete values for ϵ, namely,

$$\epsilon(n') \;=\; \frac{m}{\sqrt{1\;+\;\frac{Z^2e^4}{(n'+\sqrt{(j+\frac{1}{2})^2-Z^2e^4})^2}}}\;,\quad n' \;=\; 0,1,2,\ldots \tag{10.462}$$

This expression agrees with the spectrum of relativistic hydrogen-type atoms derived above and given by (10.405). Comparision with (10.405) allows one to identify $n' = n - j + \frac{1}{2}$ which, in fact, is an integer. For example, for the states $2p_{\frac{1}{2}}$, $3p_{\frac{1}{2}}$, $3d_{\frac{3}{2}}$ holds $n' = 1,2,1$. We can, hence, conclude that the polynomials in (10.461 for ϵ values given by (10.405) and the ensuing s_o values (10.455) are finite.

Altogether we have determined the stationary states of the type (10.421) with radial wave functions $f_1(r)$, $g_1(r)$ determined by (10.438, 10.439), (10.442), and (10.460, 10.461). The coefficients β_0 in (10.460, 10.461) are to be chosen to satisfy a normalization condition and to assign an overall phase. Due to the form (10.410) of the stationary state wave function the density $\rho(x^\mu)$ of the states under consideration, given by expression (10.267), is time-independent. The normalization integral is then

$$\int_0^\infty r^2 dr \int_0^\pi \sin\theta d\theta \int_0^{2\pi} d\phi \, (|\Phi(\vec{r})|^2 + |\mathcal{X}(\vec{r})|^2) = 1 \tag{10.463}$$

where Φ and $\mathcal{X}$, as in (10.421), are two-dimensional vectors determined through the explicit form of the spin-orbital angular momentum states $\mathcal{Y}_{jm}(j \pm \frac{1}{2}, \frac{1}{2}|\hat{r})$ in (6.147, 6.148). The orthonormality properties (6.157, 6.158) of the latter states absorb the angular integral in (10.463) and yield [note the $1/r$ factor in (10.421)]

$$\int_0^\infty dr \, (|f_1(r)|^2 + |g_1(r)|^2) = 1 \tag{10.464}$$

The evaluation of the integrals, which involve the *confluent hypergeometric functions* in (10.460, 10.461), can follow the procedure adopted for the wave functions of the non-relativistic hydrogen atom and will not be carried out here.

The wave functions (10.421) correspond to non-relativistic states with orbital angular momentum $\ell = j + \frac{1}{2}$. They are described through quantum numbers $n, j, \ell = j + \frac{1}{2}, m$. The complete wave function is given by the following set of formulas

$$\Psi(n, j, \ell = j + \tfrac{1}{2}, m|x^\mu) = e^{-i\epsilon t} \begin{pmatrix} iF_1(r)\, \mathcal{Y}_{j,m}(j + \frac{1}{2}, \frac{1}{2}|\hat{r}) \\ G_1(r)\, \mathcal{Y}_{j,m}(j - \frac{1}{2}, \frac{1}{2}|\hat{r}) \end{pmatrix} \tag{10.465}$$

$$F_1(r) = F_-(\kappa|r) \,, \quad G_1(r) = F_+(\kappa|r) \,, \quad \kappa = j + \tfrac{1}{2} \tag{10.466}$$

where[2]

$$F_\pm(\kappa|r) = \mp N (2\mu r)^{\gamma-1} e^{-\mu r} \left\{ \left[\frac{(n' + \gamma)m}{\epsilon} - \kappa \right] F(-n', 2\gamma + 1; 2\mu r) \right. \\ \left. \pm n' F(1 - n', 2\gamma + 1; 2\mu r) \right\} \tag{10.467}$$

$$N = \frac{(2\mu)^{\frac{3}{2}}}{\Gamma(2\gamma + 1)} \sqrt{\frac{m \mp \epsilon)\Gamma(2\gamma + n' + 1)}{4m \frac{(n'+\gamma)m}{\epsilon} \left(\frac{(n'+\gamma)m}{\epsilon} - \kappa \right) n'!}} \tag{10.468}$$

and

$$\begin{aligned} \mu &= \sqrt{(m - \epsilon)(m + \epsilon)} \\ \gamma &= \sqrt{(j + \tfrac{1}{2})^2 - Z^2 e^4} \\ n' &= n - j - \tfrac{1}{2} \\ \epsilon &= \frac{m}{\sqrt{1 + \frac{Z^2 e^4}{(n' + \gamma)^2}}} \,. \end{aligned} \tag{10.469}$$

[2]This formula has been adapted from "Relativistic Quantum Mechanics" by W. Greiner, (Springer, Berlin, 1990), Sect. 9.6.

We want to consider now the stationary states of the type (10.421) which, in the non-relativistic limit, become

$$\Psi(x^\mu) \;\approx\; e^{-1\epsilon t} \begin{pmatrix} i\,\dfrac{f_2(r)}{r}\mathcal{Y}_{jm}(j-\frac{1}{2},\frac{1}{2}|\hat{r}) \\ 0 \end{pmatrix} . \tag{10.470}$$

Obviously, this wavefunction has an orbital angular momentum quantum number $\ell = j - \frac{1}{2}$ and, accordingly, describes the complementary set of states $1s_{\frac{1}{2}}$, $2s_{\frac{1}{2}}$, $2p_{\frac{3}{2}}$, $3s_{\frac{1}{2}}$, $3p_{\frac{3}{2}}$, $3d_{\frac{5}{2}}$, etc. not not covered by the wave functions given by (10.465–10.469). The radial wave functions $f_2(r)$ and $g_2(r)$ in (10.421) are governed by the radial Dirac equation (10.423) which differs from the radial Dirac equation for $f_1(r)$ and $g_1(r)$ solely by the sign of the terms $(j+\frac{1}{2})/r$. One can verify, tracing all steps which lead from (10.422) to (10.469) that the following wave functions result

$$\Psi(n,j,\ell=j-\tfrac{1}{2},m|x^\mu) \;=\; e^{-i\epsilon t} \begin{pmatrix} iF_2(r)\,\mathcal{Y}_{j,m}(j-\frac{1}{2},\frac{1}{2}|\hat{r}) \\ G_2(r)\,\mathcal{Y}_{j,m}(j+\frac{1}{2},\frac{1}{2}|\hat{r}) \end{pmatrix} \tag{10.471}$$

$$F_2(r) \;=\; F_-(\kappa|r)\,, \quad G_2(r) \;=\; F_+(\kappa|r)\,, \quad \kappa \;=\; -\,j\,-\,\tfrac{1}{2} \tag{10.472}$$

where $F_\pm(\kappa|r)$ are as given in (10.467–10.469). We have, hence, obtained closed expressions for the wave functions of all the stationary bound states of relativistic hydrogen-type atoms.

Chapter 11

Spinor Formulation of Relativistic Quantum Mechanics

11.1 The Lorentz Transformation of the Dirac Bispinor

We will provide in the following a new formulation of the Dirac equation in the chiral representation defined through (10.225–10.229). Starting point is the Lorentz transformation $\tilde{\mathcal{S}}(\vec{w}, \vec{\vartheta})$ for the bispinor wave function $\tilde{\Psi}$ in the chiral representation as given by (10.262). This transformation can be written

$$\tilde{\mathcal{S}}(\vec{z}) = \begin{pmatrix} a(\vec{z}) & 0 \\ 0 & b(\vec{z}) \end{pmatrix} \tag{11.1}$$

$$a(\vec{z}) = \exp\left(\frac{1}{2}\vec{z}\cdot\vec{\sigma}\right) \tag{11.2}$$

$$b(\vec{z}) = \exp\left(-\frac{1}{2}\vec{z}^{*}\cdot\vec{\sigma}\right) \tag{11.3}$$

$$\vec{z} = \vec{w} - i\,\vec{\vartheta}\,. \tag{11.4}$$

We have altered here slightly our notation of $\tilde{\mathcal{S}}(\vec{w}, \vec{\vartheta})$, expressing its dependence on $\vec{w}$, $\vec{\vartheta}$ through a complex variable $\vec{z}, \vec{z} \in \mathbb{C}^3$.
Because of its block-diagonal form each of the diagonal components of $\tilde{\mathcal{S}}(\vec{z})$, i.e., $a(\vec{z})$ and $b(\vec{z})$, must be two-dimensional irreducible representations of the Lorentz group. This fact is remarkable since it implies that the representations provided through $a(\vec{z})$ and $b(\vec{z})$ are of lower dimension then the four-dimensional natural representation[1] $L(\vec{w}, \vec{\vartheta})^{\mu}{}_{\nu}$. The lower dimensionality of $a(\vec{z})$ and $b(\vec{z})$ implies, in a sense, that the corresponding representation of the Lorentz group is more basic than the natural representation and may serve as a building block for all representations, in particular, may be exploited to express the Lorentz-invariant equations of relativistic quantum mechanics. This is, indeed, what will be achieved in the following.
We will proceed by building as much as possible on the results obtained sofar in the chiral representation of the Dirac equation. We will characterize the space on which the transformations $a(\vec{z})$

[1] We will see below that the representations $a(\vec{z})$ and $b(\vec{z})$ are, in fact, isomorphic to the natural representation, i.e., different $L(\vec{w}, \vec{\vartheta})^{\mu}{}_{\nu}$ correspond to different $a(\vec{z})$ and $b(\vec{z})$.

and $b(\vec{z})$ act, the so-called spinor space, will establish the map between $L(\vec{w},\vec{\vartheta})^{\mu}{}_{\nu}$ and $a(\vec{z})$, $b(\vec{z})$, express 4-vectors A^{μ}, A_{ν}, the operator ∂_{μ} and the Pauli matrices $\vec{\sigma}$ in the new representation and, finally, formulate the Dirac equation, neutrino equation, and the Klein–Gordon equation in the spinor representation.

A First Characterization of the Bispinor States

We note that in case $\vec{w} = 0$ the Dirac transformations are pure rotations. In this case $a(\vec{z})$ and $b(\vec{z})$ are identical and read

$$a(i\,\vec{\vartheta}) \;=\; b(i\,\vec{\vartheta}) \;=\; \exp\left(-\frac{1}{2}\vec{\vartheta}\cdot\vec{\sigma}\right)\,, \qquad \theta \;\in\; \mathbb{R}^3\,. \tag{11.5}$$

The transformations in this case, i.e., for $\vec{z} = i\vec{\vartheta}, \vec{\vartheta} \in \mathbb{R}^3$, are elements of SU(2) and correspond, in fact, to the rotational transformations of spin-$\frac{1}{2}$ states as described by $D^{(\frac{1}{2})}_{m\,m'}(\vec{\vartheta})$, usually expressed as product of rotations and of functions of Euler angles α, β, γ (see Chapter 5). For $\vec{\vartheta} = (0,\beta,0)^T$ the transformations are

$$a(i\,\beta\,\hat{e}_2) \;=\; b(i\,\beta\,\hat{e}_2) \;=\; \left(\,d^{(\frac{1}{2})}_{mm'}(\beta)\,\right) \;=\; \begin{pmatrix} \cos\frac{\beta}{2} & -\sin\frac{\beta}{2} \\ \sin\frac{\beta}{2} & \cos\frac{\beta}{2} \end{pmatrix}. \tag{11.6}$$

as given by (5.243). This characterization allows one to draw conclusions regarding the state space in which $a(\vec{z})$ and $b(\vec{z})$ operate, namely, a space of vectors $\binom{\text{state1}}{\text{state2}}$ for which holds

$$\begin{pmatrix} \text{state 1} \;\sim\; |\frac{1}{2},+\frac{1}{2}\rangle \\ \text{state 2} \;\sim\; |\frac{1}{2},-\frac{1}{2}\rangle \end{pmatrix} \quad \vec{z} \;=\; i\vec{\vartheta},\; \vec{\vartheta} \in \mathbb{R}^3 \tag{11.7}$$

where "$\sim$" stands for "*transforms like*". Here $|\frac{1}{2},\pm\frac{1}{2}\rangle$ represents the familiar spin–$\frac{1}{2}$ states.
Since $a(\vec{z})$ acts on the first two components of the solution $\tilde{\Psi}$ of the Dirac equation, and since $b(\vec{z})$ acts on the third and fourth component of $\tilde{\Psi}$ one can characterize $\tilde{\Psi}$

$$\begin{pmatrix} \tilde{\Psi}_1(x^{\mu}) \;\sim\; |\frac{1}{2},+\frac{1}{2}\rangle \\ \tilde{\Psi}_2(x^{\mu}) \;\sim\; |\frac{1}{2},-\frac{1}{2}\rangle \\ \tilde{\Psi}_3(x^{\mu}) \;\sim\; |\frac{1}{2},+\frac{1}{2}\rangle \\ \tilde{\Psi}_4(x^{\mu}) \;\sim\; |\frac{1}{2},-\frac{1}{2}\rangle \end{pmatrix} \quad \vec{z} \;=\; i\vec{\vartheta},\; \vec{\vartheta} \in \mathbb{R}^3\,. \tag{11.8}$$

We like to stress that there exists, however, a distinct difference in the transformation behaviour of $\tilde{\Psi}_1(x^{\mu}), \tilde{\Psi}_2(x^{\mu})$ and $\tilde{\Psi}_3(x^{\mu}), \tilde{\Psi}_4(x^{\mu})$ in case $\vec{z} = \vec{w} + i\vec{\vartheta}$ for $\vec{w} \neq 0$. In this case holds $a(\vec{z}) \neq b(\vec{z})$ and $\tilde{\Psi}_1(x^{\mu}), \tilde{\Psi}_2(x^{\mu})$ transform according to $a(\vec{z})$ whereas $\tilde{\Psi}_3(x^{\mu}), \tilde{\Psi}_4(x^{\mu})$ transform according to $b(\vec{z})$.

Relationship Between $a(\vec{z})$ and $b(\vec{z})$

The transformation $b(\vec{z})$ can be related to the conjugate complex of the transformation $a(\vec{z})$, i.e., to

$$a^*(\vec{z}) \;=\; \exp\left(\frac{1}{2}\vec{z}^{\,*}\cdot\vec{\sigma}^{\,*}\right) \tag{11.9}$$

where $\vec{\sigma}^{\,*} = (\sigma_1^*, \sigma_2^*, \sigma_3^*)$. One can readily verify [c.f. (5.224)]

$$\sigma_1 = \sigma_1^* , \qquad \sigma_2 = -\sigma_2^* , \qquad \sigma_3 = \sigma_3^* . \tag{11.10}$$

From this one can derive

$$b(\vec{z}) = \epsilon\, a^*(\vec{z})\, \epsilon^{-1} \tag{11.11}$$

where

$$\epsilon = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} , \qquad \epsilon^{-1} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} . \tag{11.12}$$

To prove (11.11) one first demonstrates that for ϵ and ϵ^{-1} as given in (11.12) does, in fact, hold $\epsilon\,\epsilon^{-1} = \mathbb{1}$. One notices then, using $\epsilon f(a)\epsilon^{-1} = f(\epsilon a \epsilon^{-1})$,

$$\epsilon\, a^*(\vec{z})\epsilon^{-1} = \exp\left[\frac{1}{2}\, \vec{z}^{\,*}\big(\epsilon \vec{\sigma}^{\,*} \epsilon^{-1}\big) \right] . \tag{11.13}$$

Explicit matrix multiplication using (5.224, 11.10, 11.12) yields

$$\begin{aligned} \epsilon\sigma_1\epsilon^{-1} &= -\sigma_1 &&= -\sigma_1^* \\ \epsilon\sigma_2\epsilon^{-1} &= \sigma_2 &&= -\sigma_2^* \\ \epsilon\sigma_3\epsilon^{-1} &= -\sigma_3 &&= -\sigma_3^* , \end{aligned} \tag{11.14}$$

or in short

$$\epsilon\vec{\sigma}\epsilon^{-1} = -\vec{\sigma}^{\,*} . \tag{11.15}$$

Similarly, one can show

$$\epsilon^{-1}\vec{\sigma}\epsilon = -\vec{\sigma}^{\,*} , \tag{11.16}$$

a result to be used further below. Hence, one can express

$$\epsilon\, a^*(\vec{z})\, \epsilon^{-1} = \exp\left(-\frac{1}{2}\, \vec{z}^{\,*} \cdot \vec{\sigma} \right) = b(\vec{z}) . \tag{11.17}$$

We conclude, therefore, that the transformation (11.1) can be written in the form

$$\tilde{\mathcal{S}}(\vec{z}) = \begin{pmatrix} a(\vec{z}) & 0 \\ 0 & \epsilon\, a^*(\vec{z})\, \epsilon^{-1} \end{pmatrix} \tag{11.18}$$

with $a(\vec{z})$ given by (11.2, 11.4) and ϵ, ϵ^{-1} given by (11.12). This demonstrates that $a(\vec{z})$ is the transformation which characterizes both components of $\tilde{\mathcal{S}}(\vec{z})$.

Spatial Inversion

One may question from the form of $\tilde{\mathcal{S}}(\vec{z})$ why the Dirac equation needs to be four-dimensional, featuring the components $\tilde{\Psi}_1(x^\mu), \tilde{\Psi}_2(x^\mu)$ as well as $\tilde{\Psi}_3(x^\mu)$, $\tilde{\Psi}_4(x^\mu)$ even though these pairs of components transform independently of each other. The answer lies in the necessity that application of spatial inversion should transform a solution of the Dirac equation into another possible solution of the Dirac equation. The effect of inversion on Lorentz transformations is, however, that they alter $\vec{w}$ into $-\vec{w}$, but leave rotation angles $\vec{\vartheta}$ unaltered.

Let $\mathcal{P}$ denote the representation of spatial inversion in the space of the wave functions $\tilde{\Psi}$. Obviously, $\mathcal{P}^2 = \mathbb{1}$, i.e., $\mathcal{P}^{-1} = \mathcal{P}$. The transformation $\tilde{\mathcal{S}}(\vec{z})$ in the transformed space is then

$$\mathcal{P}\,\tilde{\mathcal{S}}(\vec{w} + i\vec{\vartheta})\,\mathcal{P} \;=\; \tilde{\mathcal{S}}(-\vec{w} + i\vec{\vartheta}) \;=\; \begin{pmatrix} b(\vec{z}) & 0 \\ 0 & a(\vec{z}) \end{pmatrix}\,, \tag{11.19}$$

i.e., the transformations $a(\vec{z})$ and $b(\vec{z})$ become interchanged. This implies

$$\mathcal{P}\begin{pmatrix} \tilde{\Psi}_1 \\ \tilde{\Psi}_2 \\ \tilde{\Psi}_3 \\ \tilde{\Psi}_4 \end{pmatrix} \;=\; \begin{pmatrix} \tilde{\Psi}_3 \\ \tilde{\Psi}_4 \\ \tilde{\Psi}_1 \\ \tilde{\Psi}_2 \end{pmatrix}\,. \tag{11.20}$$

Obviously, the space spanned by only two of the components of $\tilde{\Psi}$ is not invariant under spatial inversion and, hence, does not suffice for particles like the electron which obey inversion symmetry. However, for particles like the neutrinos which do not obey inversion symmetry two components of the wave function are sufficient. In fact, the Lorentz invariant equation for neutrinos is only 2–dimensional.

11.2 Relationship Between the Lie Groups SL(2,C) and SO(3,1)

We have pointed out that $a(i\vec{\vartheta}), \vec{\vartheta} \in \mathbb{R}^3$, which describes pure rotations, is an element of SU(2). However, $a(\vec{w} + i\vec{\vartheta})$ for $\vec{w} \neq 0$ is an element of

$$\mathrm{SL}(2,\mathbb{C}) \;=\; \{\, M,\, M \text{ is a complex } 2\times 2\text{–matrix}, \det(M) \,=\, 1 \,\}\,. \tag{11.21}$$

One can verify this by evaluating the determinant of $a(\vec{z})$

$$\det(\, a(\vec{z})\,) \;=\; \det\left(\, e^{\frac{1}{2}\vec{z}\cdot\vec{\sigma}}\,\right) \;=\; e^{\mathrm{tr}(\frac{1}{2}\vec{z}\cdot\vec{\sigma})} \;=\; 1 \tag{11.22}$$

which follows from the fact that for any complex, non-singular matrix M holds[2]

$$\det\left(\, e^{M}\,\right) \;=\; e^{\mathrm{tr}(M)} \tag{11.23}$$

and from [c.f. (5.224)]

$$\mathrm{tr}(\,\sigma_j\,) \;=\; 0\,, \qquad j \,=\, 1,2,3\,. \tag{11.24}$$

Exercise 11.2.1: Show that $SL(2,\mathbb{C})$ defined in (11.21) together with matrix multiplication as the binary operation forms a group.

---

[2]The proof of this important property is straightforward in case of hermitian M (see Chapter 5). For the general case the proof, based on the Jordan–Chevalley theorem, can be found in G.G.A. Bäuerle and E.A. de Kerf *Lie Algebras, Part* (Elsevier, Amsterdam, 1990), Exercise 1.10.3.

Mapping A^μ onto matrices $M(A^\mu)$

We want to establish now the relationship between $SL(2,\mathbb{C})$ and the group $\mathcal{L}_+^\uparrow$ of proper, orthochronous Lorentz transformations. Starting point is a bijective map between $\mathbb{R}^4$ and the set of two-dimensional hermitian matrices defined through

$$M(A^\mu) \;=\; \sigma_\mu\, A^\mu \tag{11.25}$$

where

$$\sigma_\mu \;=\; \left(\underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}_{\sigma_0}, \underbrace{\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}}_{\sigma_1}, \underbrace{\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}}_{\sigma_2}, \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}}_{\sigma_3} \right) . \tag{11.26}$$

The quantity σ_μ thus defined does not transform like a covariant 4–vector. In fact, one wishes that the definition (11.25) of the matrix $M(A^\mu)$ is independent of the frame of reference, i.e., in a transformed frame should hold

$$\sigma_\mu A^\mu \qquad \overset{L^\rho{}_\nu}{\longrightarrow} \qquad \sigma_\mu A'^\mu \; . \tag{11.27}$$

Straightforward transformation into another frame of reference would replace σ_μ by σ'_μ. Using $A^\mu = (L^{-1})^\mu{}_\nu A'^\nu$ one would expect in a transformed frame to hold

$$\sigma_\mu A^\mu \overset{L^\rho{}_\nu}{\longrightarrow} \sigma'_\mu\, (L^{-1})^\mu{}_\nu\, A'^\nu \; . \tag{11.28}$$

Consistency of (11.28) and (11.27) requires then

$$L_\nu{}^\mu \sigma'_\mu \;=\; \sigma_\nu \tag{11.29}$$

where we used (10.76). Noting that for covariant vectors according to (10.75) holds $a'_\nu = L_\nu{}^\mu a_\mu$ one realizes that σ_μ transforms *inversely* to covariant 4–vectors. We will prove below [cf. (11.135)] this transformation behaviour.
$M(A^\mu)$ according to (11.25) can also be written

$$M(A^\mu) \;=\; \begin{pmatrix} A^0 + A^3 & A^1 - iA^2 \\ A^1 + iA^2 & A^0 - A^3 \end{pmatrix} . \tag{11.30}$$

Since the components of A^μ are real, the matrix $M(A^\mu)$ is hermitian as can be seen from inspection of (11.26) or from the fact that the matrices $\sigma_0, \sigma_1, \sigma_2, \sigma_3$ are hermitian. The function $M(A^\mu)$ is bijective, in fact, one can provide a simple expression for the inverse of $M(A^\mu)$

$$M' \;=\; M(A^\mu) \qquad \leftrightarrow \qquad A^\mu \;=\; \frac{1}{2}\,\mathrm{tr}\left(M'\,\sigma_\mu \right) \; . \tag{11.31}$$

Exercise 11.2.2: Show that $\sigma_0, \sigma_1, \sigma_2, \sigma_3$ provide a linear–independent basis for the space of hermitian 2×2–matrices. Argue why $M(A^\mu) = \sigma_\mu A^\mu$ provides a bijective map. Demonstrate that (11.31) holds.

The following important property holds for $M(A^\mu)$

$$\det\left(M(A^\mu) \right) \;=\; A^\mu A_\mu \tag{11.32}$$

which follows directly from (11.30).

Transforming the matrices $M(A^\mu)$

We define now a transformation of the matrix $M(A^\mu)$ in the space of hermitian 2×2–matrices

$$M \xrightarrow{a} M' = a\, M\, a^\dagger\,, \qquad a \in SL(2,\mathbb{C})\,. \tag{11.33}$$

This transformation conserves the hermitian property of M since

$$(M')^\dagger = (\, a\, M\, a^\dagger\,)^\dagger = (a^\dagger)^\dagger M^\dagger a^\dagger = a\, M\, a^\dagger = M' \tag{11.34}$$

where we used the properties $M^\dagger = M$ and $(a^\dagger)^\dagger = a$. Due to $\det(a) = 1$ the transformation (11.33) conserves the determinant of M. In fact, it holds for the matrix M' defined through (11.33)

$$\begin{aligned} \det(M') &= \det(\, a\, M\, a^\dagger\,) = \det(a)\det(a^\dagger)\det(M) \\ &= [\det(a)]^2 \det(M) = \det(M)\,. \end{aligned} \tag{11.35}$$

We now apply the transformation (11.33) to $M(A^\mu)$ describing the action of the transformation in terms of transformations of A^μ. In fact, for any $a \in \mathrm{SL}(2,\mathbb{C})$ and for any A^μ there exists an A'^μ such that

$$M(A'^\mu) = a\, M(A^\mu)\, a^\dagger\,. \tag{11.36}$$

The suitable A'^μ can readily be constructed using (11.31). Accordingly, any $a \in \mathrm{SL}(2,\mathbb{C})$ defines the transformation [c.f. (11.31)]

$$A^\mu \xrightarrow{a} A'^\mu = \frac{1}{2}\,\mathrm{tr}\left(\, a\, M(A^\mu)\, a^\dagger \sigma_\mu \right)\,. \tag{11.37}$$

Because of (11.32, 11.35) holds for this transformation

$$A'^\mu A'_\mu = A^\mu A_\mu \tag{11.38}$$

which implies that (11.37) defines actually a Lorentz transformation. The linear character of the transformation becomes apparent expressing A'^μ as given in (11.37) using (11.25)

$$A'^\mu = \frac{1}{2}\,\mathrm{tr}\left(\, a\, \sigma_\nu a^\dagger \sigma_\mu \right) A^\nu \tag{11.39}$$

which allows us to express finally

$$A'^\mu = L(a)^\mu{}_\nu\, A^\nu\,; \qquad L(a)^\mu{}_\nu = \frac{1}{2}\,\mathrm{tr}\left(\, a\, \sigma_\nu a^\dagger \sigma_\mu \right)\,. \tag{11.40}$$

Exercise 11.2.3: Show that $L(a)^\mu{}_\nu$ defined in (11.40) is an element of $\mathcal{L}_+^\uparrow$.

$L(a)^\mu{}_\nu$ provides a homomorphism

We want to demonstrate now that the map between SL(2, $\mathbb{C}$) and SO(3,1) defined through $L(a)^\mu{}_\nu$ [cf. (11.40)] respects the group property of SL(2, $\mathbb{C}$) and of SO(3,1), i.e.,

$$\bar{L}^\mu{}_\rho \;=\; \underbrace{L(a_1)^\mu{}_\nu\, L(a_2)^\mu{}_\rho}_{\substack{\text{product}\\ \text{in SO(3,1)}}} \;=\; L(\;\underbrace{a_1\, a_2}_{\substack{\text{product}\\ \text{in SL(2,}\mathbb{C}\text{)}}}\;)^\mu{}_\rho \tag{11.41}$$

For this purpose one writes using $\mathrm{tr}(AB) = \mathrm{tr}(BA)$

$$\begin{aligned} L(a_1)^\mu{}_\nu\, L(a_2)^\mu{}_\rho \;&=\; \sum_\nu \frac{1}{2}\mathrm{tr}\left(a_1\sigma_\nu a_1^\dagger\sigma_\mu\right)\frac{1}{2}\mathrm{tr}\left(a_2\sigma_\rho a_2^\dagger\sigma_\nu\right) \\ &=\; \sum_\nu \frac{1}{2}\mathrm{tr}\left(\sigma_\nu a_1^\dagger\sigma_\mu a_1\right)\frac{1}{2}\mathrm{tr}\left(a_2\sigma_\rho a_2^\dagger\sigma_\nu\right)\;. \end{aligned} \tag{11.42}$$

Defining

$$\Gamma \;=\; a_1^\dagger\sigma_\mu a_1\;, \qquad \Gamma' \;=\; a_2\sigma_\rho a_2^\dagger \tag{11.43}$$

and using the definition of $\bar{L}^\mu{}_\rho$ in (11.41) results in

$$\bar{L}^\mu{}_\rho \;=\; \frac{1}{4}\sum_{\substack{\nu,\alpha,\beta\\ \gamma,\delta}} (\sigma_\nu)_{\alpha\beta}\Gamma_{\beta\alpha}\,\Gamma_{\gamma\delta}(\sigma_\nu)_{\delta\gamma} \;=\; \frac{1}{4}\sum_{\substack{\alpha,\beta\\ \gamma,\delta}} A_{\alpha\beta\gamma\delta}\Gamma_{\beta\alpha}\,\Gamma_{\gamma\delta} \tag{11.44}$$

where

$$A_{\alpha\beta\gamma\delta} \;=\; \sum_\nu (\sigma_\nu)_{\alpha\beta}(\sigma_\nu)_{\delta\gamma}\;. \tag{11.45}$$

One can demonstrate through direct evaluation

$$A_{\alpha\beta\gamma\delta} \;=\; \begin{cases} 2 & \alpha=\beta=\gamma=\delta=1 \\ 2 & \alpha=\beta=\gamma=\delta=2 \\ 2 & \alpha=\gamma=1\,,\ \beta=\delta=2 \\ 2 & \alpha=\gamma=2\,,\ \beta=\delta=1 \\ 0 & \text{else} \end{cases} \tag{11.46}$$

which yields

$$\begin{aligned} \bar{L}^\mu{}_\rho \;&=\; \frac{1}{2}\left(\Gamma_{11}\Gamma'_{11} \;+\; \Gamma_{22}\Gamma'_{22} \;+\; \Gamma_{12}\Gamma'_{21} \;+\; \Gamma_{21}\Gamma'_{12}\right) \;=\; \frac{1}{2}\,\mathrm{tr}\left(\Gamma\Gamma'\right) \\ &=\; \frac{1}{2}\,\mathrm{tr}\left(a_1^\dagger\sigma_\mu a_1\, a_2\sigma_\rho a_2^\dagger\right) \;=\; \frac{1}{2}\,\mathrm{tr}\left(\sigma_\mu a_1 a_2\sigma_\rho a_2^\dagger a_1^\dagger\right) \\ &=\; \frac{1}{2}\,\mathrm{tr}\left(a_1 a_2\sigma_\rho(a_1 a_2)^\dagger\sigma_\mu\right) \;=\; L(a_1a_2)^\mu{}_\rho\;. \end{aligned} \tag{11.47}$$

This completes the proof of the homomorphic property of $L(a)^\mu{}_\nu$.

Generators for SL(2, $\mathbb{C}$) which correspond to $\vec{K}$, $\vec{J}$

The transformations $a \in \mathrm{SL}(2,\mathbb{C})$ as complex 2×2–matrices are defined through four complex or, correspondingly, eight real numbers. Because of the condition $\det(a) = 1$ only six independent real numbers actually suffice for the definition of a. One expects then that six generators G_j and six real coordinates f_j can be defined which allow one to represent a in the form

$$a \;=\; \exp\left(\sum_{j=1}^{6} f_j G_j\right) \;. \tag{11.48}$$

We want to determine now the generators of the transformation $a(\vec{z})$ as defined in (11.2) which correspond to the generators $K_1, K_2, K_3, J_1, J_2, J_3$ of the Lorentz transformations $L^\mu{}_\nu$ in the natural representations, i.e., correspond to the generators given by (10.47, 10.48). To this end we consider infinitesimal transformations and keep only terms of zero order and first order in the small variables. To obtain the generator of $a(\vec{z})$ corresponding to the generator K_1, denoted below as κ_1, we write (11.36)

$$M(L^\mu{}_\nu A^\nu) \;=\; a\, M(A^\mu)\, a^\dagger \tag{11.49}$$

assuming (note that $g^\mu{}_\nu$ is just the familiar Kronecker $\delta_{\mu\nu}$)

$$L^\mu{}_\nu \;=\; g^\mu{}_\nu \;+\; \epsilon\,(K_1)^\mu{}_\nu \tag{11.50}$$

$$a \;=\; \mathbb{1} \;+\; \epsilon\,\kappa_1 \;. \tag{11.51}$$

Insertion of $(K_1)^\mu{}_\nu$ as given in (10.48) yields for the l.h.s. of (11.49), noting the linearity of $M(A^\mu)$,

$$\begin{aligned} M(A^\mu \,+\, \epsilon\,(K_1)^\mu{}_\nu\, A^\nu) \;&=\; M(A^\mu) \;+\; \epsilon\, M(\,(-A^1, -A^0, 0, 0)\,) \\ &=\; M(A^\mu) \;-\; \epsilon\,\sigma_0 A^1 \;-\; \epsilon\,\sigma_1 A^0 \end{aligned} \tag{11.52}$$

where we employed (11.25) in the last step. For the r.h.s. of (11.49) we obtain using (11.51)

$$\begin{aligned} &(\,\mathbb{1} \,+\, \epsilon\,\kappa_1\,)\, M(A^\mu)\,(\,\mathbb{1} \,+\, \epsilon\,\kappa_1^\dagger\,) \\ &\quad=\; M(A^\mu) \;+\; \epsilon\,(\,\kappa_1 M(A^\mu) \;+\; M(A^\mu)\kappa_1^\dagger\,) \;+\; O(\epsilon^2) \\ &\quad=\; M(A^\mu) \;+\; \epsilon\,(\,\kappa_1\sigma_\mu \;+\; \sigma_\mu\kappa_1^\dagger\,)\, A^\mu \;+\; O(\epsilon^2)\;. \end{aligned} \tag{11.53}$$

Equating (11.52) and (11.53) results in the condition

$$\sigma_0 A^1 \;-\; \sigma_1 A^0 \;=\; (\,\kappa_1\sigma_\mu \;+\; \sigma_\mu\,\kappa_1^\dagger\,)\, A^\mu \;. \tag{11.54}$$

This reads for the four cases $A^\mu = (1,0,0,0)$, $A^\mu = (0,1,0,0)$, $A^\mu = (0,0,1,0)$, $A^\mu = (0,0,0,1)$

$$-\,\sigma_1 \;=\; \kappa_1\sigma_0 \;+\; \sigma_0\kappa_1^\dagger \;=\; \kappa_1 \;+\; \kappa_1^\dagger \tag{11.55}$$

$$\sigma_0 \;=\; \kappa_1\sigma_1 \;+\; \sigma_1\kappa_1^\dagger \tag{11.56}$$

$$0 \;=\; \kappa_1\sigma_2 \;+\; \sigma_2\kappa_1^\dagger \tag{11.57}$$

$$0 \;=\; \kappa_1\sigma_3 \;+\; \sigma_3\kappa_1^\dagger \;. \tag{11.58}$$

One can verify readily that

$$\kappa_1 \;=\; -\frac{1}{2}\,\sigma_1 \tag{11.59}$$

obeys these conditions. Similarly, one can show that the generators κ_2, κ_3 of $a(\vec{z})$ corresponding to K_2, K_3 and $\lambda_1, \lambda_2, \lambda_3$ corresponding to J_1, J_2, J_3 are given by

$$\kappa \;=\; -\frac{1}{2}\,\vec{\sigma}\,, \qquad \vec{\lambda} \;=\; \frac{i}{2}\,\vec{\sigma}\,. \tag{11.60}$$

We can, hence, state that the following two transformations are equivalent

$$\underbrace{L(\vec{w},\vec{\vartheta}) \;=\; e^{\vec{w}\cdot\vec{K}\,+\,\vec{\vartheta}\cdot\vec{J}}}_{\begin{array}{c}\in\,\mathrm{SO(3,1)}\\ \text{acts on 4–vectors } A^\mu\end{array}} \quad , \quad \underbrace{a(\vec{w}-i\vec{\vartheta}) \;=\; e^{-\frac{1}{2}(\vec{w}\,-\,i\vec{\vartheta})\cdot\vec{\sigma}}}_{\begin{array}{c}\in\,\mathrm{SL}(2,\mathbb{C})\\ \text{acts on } \textit{spinors}\ \phi^\alpha\in\mathbb{C}^2\\ \text{(characterized below)}\end{array}} \tag{11.61}$$

This identifies the transformations $a(\vec{z}=\vec{w}-i\vec{\vartheta})$ as representations of Lorentz transformations, $\vec{w}$ describing boosts and $\vec{\vartheta}$ describing rotations.

Exercise 11.2.4: Show that the generators (11.60) of $a \in \mathrm{SL}(2,\mathbb{C})$ correspond to the generators $\vec{K}$ and $\vec{J}$ of $L^\mu{}_\nu$.

11.3 Spinors

Definition of contravariant spinors

We will now further characterize the states on which the transformation $a(\vec{z})$ and its conjugate complex $a^*(\vec{z})$ act, the so-called *contravariant spinors*. We consider first the transformation $a(\vec{z})$ which acts on a 2-dimensional space of states denoted by

$$\phi^\alpha \;=\; \begin{pmatrix}\phi^1\\ \phi^2\end{pmatrix} \;\in\; \mathbb{C}^2\,. \tag{11.62}$$

According to our earlier discussion holds

$$\begin{array}{ll}\phi^1 & \text{transforms under rotations } (\vec{z}\,=\,i\vec{\vartheta}) \text{ like a spin–}\tfrac{1}{2}\text{ state } |\tfrac{1}{2},+\tfrac{1}{2}\rangle\\ \phi^2 & \text{transforms under rotations } (\vec{z}\,=\,i\vec{\vartheta}) \text{ like a spin–}\tfrac{1}{2}\text{ state } |\tfrac{1}{2},-\tfrac{1}{2}\rangle\,.\end{array}$$

We denote the general $(\vec{z}\,=\,\vec{w}\,+\,i\vec{\vartheta})$ transformation by

$$\phi'^\alpha \;=\; a^\alpha{}_\beta\,\phi^\beta \;\stackrel{\text{def}}{=}\; a^\alpha{}_1\phi^1 \,+\, a^\alpha{}_2\phi^2\,, \quad \alpha\,=\,1,2 \tag{11.63}$$

where we extended the summation convention of 4-vectors to spinors. Here

$$a(\vec{z}) \;=\; (\,a^\alpha{}_\beta\,) \;=\; \begin{pmatrix} a^1{}_1 & a^1{}_2\\ a^2{}_1 & a^2{}_2\end{pmatrix} \tag{11.64}$$

describes the matrix $a(\vec{z})$.

Definition of a scalar product

The question arises if for the states ϕ^α there exists a scalar product which is invariant under Lorentz transformations, i.e., invariant under transformations $a(\vec{z})$. Such a scalar product does, indeed, exist and it plays a role for spinors which is as central as the role of the scalar product $A^\mu A_\mu$ is for 4–vectors. To arrive at a suitable scalar product we consider first only rotational transformations $a(i\vec{\vartheta})$. In this case spinors ϕ^α transform like spin–$\frac{1}{2}$ states and an invariant, which can be constructed from products $\phi^1\chi^2$, etc., is the singlet state. In the notation developed in Chapter 5 holds for the singlet state

$$|\tfrac{1}{2},\tfrac{1}{2};0,0\rangle \;=\; \sum_{m=\pm\frac{1}{2}} (0,0|\tfrac{1}{2},m;\tfrac{1}{2},-m)\,|\tfrac{1}{2},m\rangle_1|\tfrac{1}{2},-m\rangle_2 \tag{11.65}$$

where $|\cdots\rangle_1$ describes the spin state of "particle 1" and $|\cdots\rangle_2$ describes the spin state of "particle 2" and $(0,0|\frac{1}{2},\pm\frac{1}{2};\frac{1}{2},\mp\frac{1}{2})$ stands for the Clebsch–Gordon coefficient. Using $(0,0|\frac{1}{2},\pm\frac{1}{2};\frac{1}{2},\mp\frac{1}{2}) = \pm 1/\sqrt{2}$ and equating the spin states of "particle 1" with the spinor ϕ^α, those of "particle 2" with the spinor χ^β one can state that the quantity

$$\Sigma \;=\; \frac{1}{\sqrt{2}}\,\left(\,\phi^1\,\chi^2 \;-\; \phi^2\,\chi^1\,\right) \tag{11.66}$$

should constitute a singlet spin state, i.e., should remain invariant under transformations $a(i\vec{\vartheta})$. In fact, as we will demonstrate below such states are invariant under general Lorentz transformations $a(\vec{z})$.

Definition of covariant spinors

Expression (11.66) is a bilinear form, invariant and as such has the necessary properties of a scalar[3] product. However, this scalar product is anti-symmetric, i.e., exchange of ϕ^α and χ^β alters the sign of the expression. The existence of a scalar product gives rise to the definition of a dual representation of the states ϕ^α denoted by ϕ_α. The corresponding states are defined through

$$\phi^1\,\chi^2 \;-\; \phi^2\,\chi^1 \;=\; \phi^1\chi_1 \;+\; \phi^2\chi_2 \tag{11.67}$$

It obviously holds

$$\chi_\alpha \;=\; \left(\begin{array}{c}\chi_1\\ \chi_2\end{array}\right) \;=\; \left(\begin{array}{c}\chi^2\\ -\chi^1\end{array}\right)\,. \tag{11.68}$$

We will refer to $\phi^\alpha,\chi^\beta,\ldots$ as *contravariant spinors* and to $\phi_\alpha,\chi_\beta,\ldots$ as *covariant spinors*. The relationship between the two can be expressed

$$\left(\begin{array}{c}\phi_1\\ \phi_2\end{array}\right) \;=\; \epsilon\left(\begin{array}{c}\phi^1\\ \phi^2\end{array}\right) \tag{11.69}$$

$$\epsilon \;=\; \left(\begin{array}{cc}0 & 1\\ -1 & 0\end{array}\right) \tag{11.70}$$

[3]'Scalar' implies invariance under rotations and is conventionally generalized to invariance under other symmetry trasnformations.

as can be verified from (11.68). The inverse of (11.69, 11.70) is

$$\begin{pmatrix} \phi^1 \\ \phi^2 \end{pmatrix} = \epsilon^{-1} \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \tag{11.71}$$

$$\epsilon^{-1} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} . \tag{11.72}$$

Exercise 11.3.1: Show that for any non-singular complex 2×2–matrix M holds

$$\epsilon\, M\, \epsilon^{-1} \;=\; \det(M)\, \left(M^{-1}\right)^T$$

The matrices ϵ, ϵ^{-1} connecting contravariant and covariant spinors play the role of the metric tensors $g_{\mu\nu}, g^{\mu\nu}$ of the Minkowski space [cf. (10.10, 10.74)]. Accordingly, we will express (11.69, 11.70) and (11.71, 11.72) in a notation analogous to that chosen for contravariant and covariant 4–vectors [cf. (10.72)]

$$\phi_\alpha = \epsilon_{\alpha\beta}\, \phi^\beta \tag{11.73}$$

$$\phi^\alpha = \epsilon^{\alpha\beta}\, \phi_\beta \tag{11.74}$$

$$\epsilon_{\alpha\beta} = \begin{pmatrix} \epsilon_{11} & \epsilon_{12} \\ \epsilon_{21} & \epsilon_{22} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{11.75}$$

$$\epsilon^{\alpha\beta} = \begin{pmatrix} \epsilon^{11} & \epsilon^{12} \\ \epsilon^{21} & \epsilon^{22} \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{11.76}$$

The scalar product (11.67) will be expressed as

$$\phi^\alpha \chi_\alpha \;=\; \phi^1 \chi_1 \;+\; \phi^2 \chi_2 \;=\; \phi^1\, \chi^2 \;-\; \phi^2\, \chi^1 \, . \tag{11.77}$$

For this scalar product holds

$$\phi^\alpha \chi_\alpha \;=\; -\chi^\alpha \phi_\alpha \, . \tag{11.78}$$

The transformation behaviour of ϕ_α according to (11.63, 11.73, 11.75) is given by

$$\phi'_\alpha \;=\; \epsilon_{\alpha\beta}\, a^\beta{}_\gamma\, \epsilon^{\gamma\delta}\, \phi_\delta \tag{11.79}$$

as can be readily verified.

Proof that $\phi^\alpha \chi_\alpha$ is Lorentz invariant

We want to verify now that the scalar product (11.77) is Lorentz invariant. In the transformed frame holds

$$\phi'^\alpha \chi'_\alpha \;=\; a^\alpha{}_\beta\, \epsilon_{\alpha\gamma} a^\gamma{}_\delta \epsilon^{\delta\kappa}\, \phi^\beta \chi_\kappa \, . \tag{11.80}$$

One can write in matrix notation

$$a^\alpha{}_\beta\, \epsilon_{\alpha\gamma} a^\gamma{}_\delta \epsilon^{\delta\kappa} \;=\; \left[\left(\epsilon\, a\, \epsilon^{-1}\right)^T \, a \right]_{\kappa\beta} \, . \tag{11.81}$$

Using (11.2) and (11.14) one can write

$$\epsilon\, a\, \epsilon^{-1} \;=\; \epsilon\, e^{\frac{1}{2}\vec{z}\cdot\vec{\sigma}}\epsilon^{-1} \;=\; e^{\frac{1}{2}\vec{z}\cdot\epsilon\vec{\sigma}\epsilon^{-1}} \;=\; e^{-\frac{1}{2}\vec{z}\cdot\vec{\sigma}^{\,*}} \tag{11.82}$$

and with $f(A)^T = f(A^T)$ for polynomial $f(A)$

$$\left(\epsilon\, a\, \epsilon^{-1}\right)^T \;=\; e^{-\frac{1}{2}\vec{z}\cdot(\vec{\sigma}^{\,*})^T} \;=\; e^{-\frac{1}{2}\vec{z}\cdot\vec{\sigma}} \;=\; a^{-1} \tag{11.83}$$

Here we have employed the hermitian property of $\vec{\sigma}$, i.e., $\vec{\sigma}^{\dagger} = (\vec{\sigma}^{\,*})^T = \vec{\sigma}$. Insertion of (11.83) into (11.81) yields

$$a^{\alpha}{}_{\beta}\,\epsilon_{\alpha\gamma}a^{\gamma}{}_{\delta}\epsilon^{\delta\kappa} \;=\; \left[\left(\epsilon\, a\, \epsilon^{-1}\right)^T\, a\right]_{\kappa\beta} \;=\; \left[a^{-1}a\right]_{\kappa\beta} \;=\; \delta_{\kappa\beta} \tag{11.84}$$

and, hence, from (11.80)

$$\phi'^{\alpha}\chi'_{\alpha} \;=\; \phi^{\beta}\chi_{\beta}\,. \tag{11.85}$$

The complex conjugate spinors

We consider now the *conjugate complex spinors*

$$(\phi^{\alpha})^* \;=\; \begin{pmatrix} (\phi^1)^* \\ (\phi^2)^* \end{pmatrix}\,. \tag{11.86}$$

A concise notation of the conjugate complex spinors is provided by

$$(\phi^{\alpha})^* \;=\; \phi^{\dot{\alpha}} \;=\; \begin{pmatrix} \phi^{\dot{1}} \\ \phi^{\dot{2}} \end{pmatrix} \tag{11.87}$$

which we will employ from now on. Obviously, it holds $\phi^{\dot{k}} = (\phi^k)^*, k = 1, 2$. The transformation behaviour of $\phi^{\dot{\alpha}}$ is

$$\phi'^{\dot{\alpha}} \;=\; (a^{\alpha}{}_{\beta})^*\,\phi^{\dot{\beta}} \tag{11.88}$$

which one verifies taking the conjugate complex of (11.63). As discussed above, $a^*(\vec{z})$ provides a representation of the Lorentz group which is distinct from that provided by $a(\vec{z})$. Hence, the conjugate complex spinors $\phi^{\dot{\alpha}}$ need to be considered separately from the spinors ϕ^{α}. We denote

$$(a^{\alpha}{}_{\beta})^* \;=\; a^{\dot{\alpha}}{}_{\dot{\beta}} \tag{11.89}$$

such that (11.88) reads

$$\phi'^{\dot{\alpha}} \;=\; a^{\dot{\alpha}}{}_{\dot{\beta}}\phi^{\dot{\beta}} \tag{11.90}$$

extending the summation convention to 'dotted' indices.
We also define covariant versions of $\phi^{\dot{\alpha}}$

$$\phi_{\dot{\alpha}} \;=\; (\phi_{\alpha})^*\,. \tag{11.91}$$

The relationship between contravariant and covariant conjugate complex spinors can be expressed in analogy to (11.73, 11.76)

$$\phi_{\dot\alpha} = \epsilon_{\dot\alpha\dot\beta}\,\phi^{\dot\beta} \tag{11.92}$$

$$\phi^{\dot\alpha} = \epsilon^{\dot\alpha\dot\beta}\,\phi_{\dot\beta} \tag{11.93}$$

$$\epsilon_{\dot\alpha\dot\beta} = \epsilon_{\alpha\beta} \tag{11.94}$$

$$\epsilon^{\dot\alpha\dot\beta} = \epsilon^{\alpha\beta} \tag{11.95}$$

where $\epsilon_{\alpha\beta}$ and $\epsilon^{\alpha\beta}$ are the real matrices defined in (11.75, 11.75). For the spinors $\phi^{\dot\alpha}$ and $\chi_{\dot\alpha}$ thus defined holds that the scalar product

$$\phi^{\dot\alpha}\chi_{\dot\alpha} \;=\; \phi^{\dot 1}\chi_{\dot 1} \;+\; \phi^{\dot 2}\chi_{\dot 2} \tag{11.96}$$

is Lorentz invariant, a property which is rather evident.
The transformation behaviour of $\phi_{\dot\alpha}$ is

$$\phi'_{\dot\alpha} \;=\; \epsilon_{\dot\alpha\dot\beta}a^{\dot\beta}{}_{\dot\gamma}\epsilon^{\dot\gamma\dot\delta}\phi_{\dot\delta}\;. \tag{11.97}$$

The transformation, in matrix notation, is governed by the operator $\epsilon\,a^*(\vec z)\,\epsilon^{-1}$ which arises in the Lorentz transformation (11.18) of the bispinor wave function $\tilde\Psi$, $\epsilon\,a^*(\vec z)\,\epsilon^{-1}$ accounting for the transformation behaviour of the third and fourth spinor component of $\tilde\Psi$. A comparision of (11.18) and (11.97) implies then that $\phi_{\dot\alpha}$ transforms like $\tilde\Psi_3, \tilde\Psi_4$, i.e., one can state

$\phi_{\dot 1}$ transforms under rotations ($\vec z = i\vec\vartheta$) like a spin–$\frac{1}{2}$ state $|\frac{1}{2}, +\frac{1}{2}\rangle$
$\phi_{\dot 2}$ transforms under rotations ($\vec z = i\vec\vartheta$) like a spin–$\frac{1}{2}$ state $|\frac{1}{2}, -\frac{1}{2}\rangle$.

The transformation behaviour of $\tilde\Psi$ (note that we do not include presently the space-time dependence of the wave function)

$$\tilde\Psi' \;=\; \tilde{\mathcal S}(\vec z)\,\tilde\Psi \;=\; \begin{pmatrix} a(\vec z)\begin{pmatrix}\tilde\Psi_1\\ \tilde\Psi_2\end{pmatrix} \\ \epsilon a^*(\vec z)\epsilon^{-1}\begin{pmatrix}\tilde\Psi_3\\ \tilde\Psi_4\end{pmatrix}\end{pmatrix} \tag{11.98}$$

obviously implies that the solution of the Dirac equation in the chiral representation can be written in spinor form

$$\begin{pmatrix}\tilde\Psi_1(x^\mu)\\ \tilde\Psi_2(x^\mu)\\ \tilde\Psi_3(x^\mu)\\ \tilde\Psi_4(x^\mu)\end{pmatrix} = \begin{pmatrix}\phi^1(x^\mu)\\ \phi^2(x^\mu)\\ \chi_{\dot 1}(x^\mu)\\ \chi_{\dot 2}(x^\mu)\end{pmatrix} = \begin{pmatrix}\phi^\alpha(x^\mu)\\ \chi_{\dot\beta}(x^\mu)\end{pmatrix}\;. \tag{11.99}$$

11.4 Spinor Tensors

We generalize now our definition of spinors ϕ^α to tensors. A tensor

$$t^{\alpha_1\alpha_2\cdots\alpha_k\dot\beta_1\dot\beta_2\cdots\dot\beta_\ell} \tag{11.100}$$

is a quantity which under Lorentz transformations behaves as

$$t'^{\alpha_1\alpha_2\cdots\alpha_k\dot\beta_1\dot\beta_2\cdots\dot\beta_\ell} \;=\; \prod_{m=1}^{k} a^{\alpha_m}{}_{\gamma_m} \prod_{n=1}^{\ell} a^{\dot\beta_n}{}_{\dot\delta_n}\, t^{\gamma_1\cdots\gamma_k\dot\delta_1\cdots\dot\delta_\ell} \;. \tag{11.101}$$

An example is the tensor $t^{\alpha\dot\beta}$ which will play an important role in the spinor presentation of the Dirac equation. This tensor transforms according to

$$t'^{\alpha\dot\beta} \;=\; a^{\alpha}{}_{\gamma} a^{\dot\beta}{}_{\dot\delta}\, t^{\gamma\dot\delta} \tag{11.102}$$

This reads in matrix notation, using conventional matrix indices j,k,ℓ,m,

$$t'_{jk} \;=\; \left(a\, t\, a^\dagger\right)_{jk} \;=\; \sum_{\ell,m} a_{j\ell} a^*_{km} t_{\ell m} \;. \tag{11.103}$$

Similarly, the transformation bevaviour of a tensor $t^{\alpha\beta}$ reads in spinor and matrix notation

$$t'^{\alpha\beta} \;=\; a^{\alpha}{}_{\gamma}\, a^{\beta}{}_{\delta}\, t^{\gamma\delta} \;, \qquad t'_{jk} \;=\; \left(a\, t\, a^T\right)_{jk} \;=\; \sum_{\ell,m} a_{j\ell} a_{km} t_{\ell m} \tag{11.104}$$

Indices on tensors can also be lowered employing the formula

$$t_{\alpha}{}^{\dot\beta} \;=\; \epsilon_{\alpha\gamma} t^{\gamma\dot\beta} \tag{11.105}$$

and generalizations thereof.
An example of a tensor is $\epsilon^{\alpha\beta}$ and $\epsilon_{\alpha\beta}$. This tensor is actually invariant under Lorentz transformations, i.e., it holds

$$\epsilon'^{\alpha\beta} \;=\; \epsilon^{\alpha\beta} \;, \qquad \epsilon'_{\alpha\beta} \;=\; \epsilon_{\alpha\beta} \tag{11.106}$$

Exercise 11.4.1: Prove equation (11.106).

The 4–vector A^μ in spinor form

We want to provide now the spinor form of the 4-vector A^μ, i.e., we want to express A^μ through a spinor tensor. This task implies that we seek a tensor, the elements of which are linear functions of A^μ. An obvious candidate is [cf. (11.25] $M(A^\mu) \;=\; \sigma_\mu A^\mu$. We had demonstrated that $M(A^\mu)$ transforms according to

$$M' \;=\; M(L^\mu{}_\nu A^\nu) \;=\; a\, M(A^\mu)\, a^\dagger \tag{11.107}$$

which reads in spinor notation [cf. (11.102, 11.103)

$$A'^{\alpha\dot\beta} \;=\; a^{\alpha}{}_{\gamma} a^{\dot\beta}{}_{\dot\delta} A^{\gamma\dot\delta} \;. \tag{11.108}$$

Obviously, this transformation behaviour is in harmony with the tensor notation adopted, i.e., with contravariant indices $\alpha\dot\beta$. According to (11.25) the tensor is explicitly

$$A^{\alpha\dot\beta} \;=\; \begin{pmatrix} A^{1\dot1} & A^{1\dot2} \\ A^{2\dot1} & A^{2\dot2} \end{pmatrix} \;=\; \begin{pmatrix} A^0 + A^3 & A^1 - iA^2 \\ A^1 + iA^2 & A^0 - A^3 \end{pmatrix} \;. \tag{11.109}$$

One can express $A^{\alpha\dot{\beta}}$ also through A_μ

$$A^{\alpha\dot{\beta}} \;=\; \begin{pmatrix} A_0 \,-\, A_3 & -A_1 \,+\, iA_2 \\ -A_1 \,-\, iA_2 & A_0 \,+\, A_3 \end{pmatrix} . \tag{11.110}$$

The 4-vectors A^μ, A_μ can also be associated with tensors

$$A_{\alpha\dot{\beta}} \;=\; \epsilon_{\alpha\gamma}\epsilon_{\dot{\beta}\dot{\delta}}A^{\gamma\dot{\delta}} . \tag{11.111}$$

This tensor reads in matrix notation

$$\begin{aligned} \begin{pmatrix} A_{1\dot{1}} & A_{1\dot{2}} \\ A_{2\dot{1}} & A_{2\dot{2}} \end{pmatrix} &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} A^{1\dot{1}} & A^{1\dot{2}} \\ A^{2\dot{1}} & A^{2\dot{2}} \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} A^{2\dot{2}} & -A^{2\dot{1}} \\ -A^{1\dot{2}} & A^{1\dot{1}} \end{pmatrix} . \end{aligned} \tag{11.112}$$

Hence, employing (11.109, 11.110) one obtains

$$A_{\alpha\dot{\beta}} \;=\; \begin{pmatrix} A^0 \,-\, A^3 & -A^1 \,-\, iA^2 \\ -A^1 \,+\, iA^2 & A^0 \,+\, A^3 \end{pmatrix} \tag{11.113}$$

$$A_{\alpha\dot{\beta}} \;=\; \begin{pmatrix} A_0 \,+\, A_3 & A_1 \,+\, iA_2 \\ A_1 \,-\, iA_2 & A_0 \,-\, A_3 \end{pmatrix} . \tag{11.114}$$

We finally note that the 4–vector scalar product $A^\mu B_\mu$ reads in spinor notation

$$A^\mu B_\mu \;=\; \frac{1}{2}A^{\alpha\dot{\beta}}B_{\alpha\dot{\beta}} . \tag{11.115}$$

Exercise 11.4.2: Prove that (11.115) is correct.

∂_μ in spinor notation

The relationship between 4–vectors A^μ, A_μ and tensors $t^{\alpha\dot{\beta}}$ can be applied to the partial differential operator ∂_μ. Using (11.110) one can state

$$\partial^{\alpha\dot{\beta}} \;=\; \begin{pmatrix} \partial_0 \,-\, \partial_3 & -\partial_1 \,+\, i\partial_2 \\ -\partial_1 \,-\, i\partial_2 & \partial_0 \,+\, \partial_3 \end{pmatrix} . \tag{11.116}$$

Similarly, (11.114) yields

$$\partial_{\alpha\dot{\beta}} \;=\; \begin{pmatrix} \partial_0 \,+\, \partial_3 & \partial_1 \,+\, i\partial_2 \\ \partial_1 \,-\, i\partial_2 & \partial_0 \,-\, \partial_3 \end{pmatrix} . \tag{11.117}$$

σ_μ in Tensor Notation

We want to develop now the tensor notation for σ_μ (11.26) and its contravariant analogue σ^μ

$$\begin{aligned}\sigma_\mu &= \left(\begin{pmatrix}1&0\\0&1\end{pmatrix},\begin{pmatrix}0&1\\1&0\end{pmatrix},\begin{pmatrix}0&-i\\i&0\end{pmatrix},\begin{pmatrix}1&0\\0&-1\end{pmatrix}\right)\\ \sigma^\mu &= \left(\begin{pmatrix}1&0\\0&1\end{pmatrix},\begin{pmatrix}0&-1\\-1&0\end{pmatrix},\begin{pmatrix}0&i\\-i&0\end{pmatrix},\begin{pmatrix}-1&0\\0&1\end{pmatrix}\right)\end{aligned} \tag{11.118}$$

For this purpose we consider first the transformation behaviour of σ^μ and σ_μ. We will obtain the transformation behaviour of σ^μ, σ_μ building on the known transformation behaviour of $\tilde{\gamma}^\mu$. This is possible since $\tilde{\gamma}^\mu$ can be expressed through σ^μ, σ_μ. Comparision of (10.229) and (11.118) yields

$$\tilde{\gamma}^\mu = \begin{pmatrix}0&\sigma^\mu\\ \sigma_\mu&0\end{pmatrix}. \tag{11.119}$$

Using (11.15) one can write

$$\sigma_\mu = \epsilon\,(\sigma^\mu)^*\,\epsilon^{-1} \tag{11.120}$$

and, hence,

$$\tilde{\gamma}^\mu = \begin{pmatrix}0&\sigma^\mu\\ \epsilon\,(\sigma^\mu)^*\,\epsilon^{-1}&0\end{pmatrix}. \tag{11.121}$$

One expects then that the transformation properties of σ^μ should follow from the transformation properties established already for γ^μ [c.f. (10.243)]. Note that (10.243) holds independently of the representation chosen, i.e., holds also in the chiral representation.
To obtain the transformation properties of σ^μ we employ then (10.243) in the chiral representation expressing $\mathcal{S}(L^\eta{}_\xi)$ by (11.18) and γ^μ by (11.121). Equation (10.243) reads then

$$\begin{pmatrix}a&0\\0&\epsilon a^*\epsilon^{-1}\end{pmatrix}\begin{pmatrix}0&\sigma^\mu\\ \epsilon(\sigma^\mu)^*\epsilon^{-1}&0\end{pmatrix}\begin{pmatrix}a^{-1}&0\\0&\epsilon a^*\epsilon^{-1}\end{pmatrix} L^\nu{}_\mu = \begin{pmatrix}0&\sigma^\nu\\ \epsilon(\sigma^\nu)^*\epsilon^{-1}&0\end{pmatrix}. \tag{11.122}$$

The l.h.s. of this equation is

$$\begin{pmatrix}0&a\sigma^\mu\epsilon^{-1}(a^*)^{-1}\epsilon\\ \epsilon a^*(\sigma^\mu)^*\epsilon^{-1}a^{-1}&0\end{pmatrix} L^\nu{}_\mu \tag{11.123}$$

and, hence, one can conclude

$$a\sigma^\mu\epsilon^{-1}(a^*)^{-1}\epsilon\, L^\nu{}_\mu = \sigma^\nu \tag{11.124}$$

$$\epsilon a^*(\sigma^\mu)^*\epsilon^{-1}a^{-1}\, L^\nu{}_\mu = \epsilon(\sigma^\nu)^*\epsilon^{-1}. \tag{11.125}$$

Equation (11.125) is equivalent to

$$a^*(\sigma^\mu)^*\epsilon^{-1}a^{-1}\epsilon\, L^\nu{}_\mu = (\sigma^\nu)^* \tag{11.126}$$

which is the complex conjugate of (11.124), i.e., (11.125) is equivalent to (11.124). Hence, (11.124) constitutes the essential transformation property of σ^μ and will be considered further.

One can rewrite (11.124) using (11.16, 11.2)

$$\epsilon^{-1}(a^*)^{-1}\epsilon \;=\; \epsilon^{-1}e^{-\frac{1}{2}\vec{z}^*\cdot\vec{\sigma}^*}\epsilon \;=\; e^{-\frac{1}{2}\vec{z}^*\cdot\epsilon^{-1}\vec{\sigma}^*\epsilon} \;=\; e^{\frac{1}{2}\vec{z}^*\cdot\vec{\sigma}}\;. \tag{11.127}$$

Exploiting the hermitian property of $\vec{\sigma}$, e.g., $(\vec{\sigma}^*)^T = \vec{\sigma}$ yields using (11.2)

$$\epsilon^{-1}(a^*)^{-1}\epsilon \;=\; e^{\frac{1}{2}\vec{z}^*\cdot(\vec{\sigma}^*)^T} \;=\; \left[e^{\frac{1}{2}\vec{z}^*\cdot\vec{\sigma}^*}\right]^T \;=\; [a^*]^T\;. \tag{11.128}$$

One can express, therefore, equation (11.124)

$$a\,\sigma^\mu\;[a^*]^T\;L^\nu{}_\mu \;=\; \sigma^\nu\;. \tag{11.129}$$

We want to demonstrate now that the expression $a\sigma^\mu[a^*]^T$ is to be interpreted as the transform of σ^μ under Lorentz transformations. In fact, under rotations the Pauli matrices transform like $(j = 1, 2, 3)$

$$\sigma_j \;\longrightarrow\; a(i\vec{\vartheta})\,\sigma_j\left(a(i\vec{\vartheta})\right)^\dagger \;=\; a(i\vec{\vartheta})\,\sigma_j\left(a^*(i\vec{\vartheta})\right)^T\;. \tag{11.130}$$

We argue in analogy to the logic applied in going from (11.107) to (11.108) that the same transformation behaviour applies then for general Lorentz transformations, i.e., transformations (11.2, 11.4) with $\vec{w} \neq 0$. One can, hence, state that σ^μ in a new reference frame is

$$\sigma'^\mu \;=\; a\,\sigma^\mu\,a^\dagger \tag{11.131}$$

where a is given by (11.2, 11.4). This transformation behaviour, according to (11.102, 11.103) identifies σ^μ as a tensor of type $t^{\alpha\dot{\beta}}$. It holds according to (11.118)

$$\begin{pmatrix} (\sigma^\mu)^{1\dot{1}} & (\sigma^\mu)^{1\dot{2}} \\ (\sigma^\mu)^{2\dot{1}} & (\sigma^\mu)^{2\dot{2}} \end{pmatrix} \;=$$

$$\left(\underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}_{\mu=0}, \underbrace{\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}}_{\mu=1}, \underbrace{\begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}}_{\mu=2}, \underbrace{\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}}_{\mu=3},\right) \tag{11.132}$$

and

$$\begin{pmatrix} (\sigma_\mu)^{1\dot{1}} & (\sigma_\mu)^{1\dot{2}} \\ (\sigma^\mu)_{2\dot{1}} & (\sigma^\mu)_{2\dot{2}} \end{pmatrix} \;=$$

$$\left(\underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}_{\mu=0}, \underbrace{\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}}_{\mu=1}, \underbrace{\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}}_{\mu=2}, \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}}_{\mu=3},\right)\;. \tag{11.133}$$

Combining (11.129, 11.131) one can express the transformation behaviour of σ^μ in the succinct form

$$L^\nu{}_\mu\,\sigma'^\mu \;=\; \sigma^\nu\;. \tag{11.134}$$

Inverting contravariant and covariant indices one can also state

$$L_\nu{}^\mu \, \sigma'_\mu \; = \; \sigma_\nu \; . \tag{11.135}$$

This is the property surmised already above [cf. (11.29)]. We can summarize that σ^μ and σ_μ transform like a 4–vector, however, the transformation is *inverse* to that of ordinary 4–vectors. Each of the $4 \times 4 \; = 16$ matrix elements in (11.132) and (11.133) is characterized through a 4-vector index $\mu, \mu \; = \; 0, 1, 2, 3$ as well as through two spinor indices $\alpha\dot{\beta}$. We want to express now σ^μ and σ_μ also with respect to the 4–vector index μ in spinor form employing (11.114). This yields

$$\sigma_{\alpha\dot{\beta}} \; = \; \begin{pmatrix} \sigma_0 \, + \, \sigma_3 & \sigma_1 \, + \, i\sigma_2 \\ \sigma_1 \, - \, i\sigma_2 & \sigma_0 \, - \, \sigma_3 \end{pmatrix} \; = \; 2 \begin{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \\ \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \end{pmatrix} \tag{11.136}$$

where on the rhs. the submatrices correspond to $t^{\alpha\dot{\beta}}$ spinors. We can, in fact, state

$$\left(\sigma_{\alpha\dot{\beta}}\right)^{\gamma\dot{\delta}} \; = \; \begin{pmatrix} \begin{pmatrix} (\sigma_{1\dot{1}})^{1\dot{1}} & (\sigma_{1\dot{1}})^{1\dot{2}} \\ (\sigma_{1\dot{1}})^{2\dot{1}} & (\sigma_{1\dot{1}})^{2\dot{2}} \end{pmatrix} & \begin{pmatrix} (\sigma_{1\dot{2}})^{1\dot{1}} & (\sigma_{1\dot{2}})^{1\dot{2}} \\ (\sigma_{1\dot{2}})^{2\dot{1}} & (\sigma_{1\dot{2}})^{2\dot{2}} \end{pmatrix} \\ \begin{pmatrix} (\sigma_{2\dot{1}})^{1\dot{1}} & (\sigma_{2\dot{1}})^{1\dot{2}} \\ (\sigma_{2\dot{1}})^{2\dot{1}} & (\sigma_{2\dot{1}})^{2\dot{2}} \end{pmatrix} & \begin{pmatrix} (\sigma_{2\dot{2}})^{1\dot{1}} & (\sigma_{2\dot{2}})^{1\dot{2}} \\ (\sigma_{2\dot{2}})^{2\dot{1}} & (\sigma_{2\dot{2}})^{2\dot{2}} \end{pmatrix} \end{pmatrix} \; . \tag{11.137}$$

Equating this with the r.h.s. of (11.136) results in the succinct expression

$$\frac{1}{2} \left(\sigma_{\alpha\dot{\beta}}\right)^{\gamma\dot{\delta}} \; = \; \delta_{\alpha\gamma}\delta_{\dot{\beta}\dot{\delta}} \; . \tag{11.138}$$

Note that all elements of $\sigma_{\alpha\dot{\beta}}$ are real and that there are only four non-vanishing elements.
In (11.138) the 'inner' covariant spinor indices, i.e., $\alpha, \dot{\beta}$, account for the 4–vector index μ, whereas the 'outer' contravariant spinor indices. i.e., $\gamma, \dot{\delta}$, account for the elements of the individual Pauli matrices. We will now consider the representation of σ^μ, σ_μ in which the contravariant indices are moved 'inside', i.e., account for the 4-vector μ, and the covariant indices are moved outside. The desired change of representation$(\sigma_\mu)^{\alpha\dot{\beta}} \; \longrightarrow \; (\sigma_\mu)_{\alpha\dot{\beta}}$ corresponds to a transformation of the basis of spin states

$$\begin{pmatrix} f \\ g \end{pmatrix} \; \longrightarrow \; \begin{pmatrix} g \\ -f \end{pmatrix} \; = \; \epsilon \begin{pmatrix} f \\ g \end{pmatrix} \tag{11.139}$$

and, hence, corresponds to the transformation

$$\begin{pmatrix} (\sigma_\mu)_{1\dot{1}} & (\sigma_\mu)_{1\dot{2}} \\ (\sigma_\mu)_{2\dot{1}} & (\sigma_\mu)_{2\dot{2}} \end{pmatrix} \; = \; \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} (\sigma_\mu)^{1\dot{1}} & (\sigma_\mu)^{1\dot{2}} \\ (\sigma_\mu)^{2\dot{1}} & (\sigma_\mu)^{2\dot{2}} \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{11.140}$$

where we employed the expressions (11.12) for ϵ and ϵ^{-1}. Using (11.110) to express $\sigma^{\alpha\dot{\beta}}$ in terms of σ_μ yields together with (5.224)

$$\sigma^{\alpha\dot{\beta}} \; = \; \begin{pmatrix} \sigma_0 \, - \, \sigma_3 & \sigma_1 \, + \, i\sigma_2 \\ -\sigma_1 \, - \, i\sigma_2 & \sigma_0 \, + \, \sigma_3 \end{pmatrix} \; = \; 2 \begin{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} \\ \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{pmatrix} \tag{11.141}$$

and, employing then transformation (11.140) to transform each of the four submatrices, which in (11.137) are in a basis $(\cdots)^{\alpha\dot{\beta}}$ to a basis $(\cdots)_{\alpha\dot{\beta}}$ results in

$$\begin{aligned}\left(\sigma^{\alpha\dot{\beta}}\right)_{\gamma\dot{\delta}} &= \begin{pmatrix} \begin{pmatrix} (\sigma^{1\dot{1}})_{1\dot{1}} & (\sigma^{1\dot{1}})_{1\dot{2}} \\ (\sigma^{1\dot{1}})_{2\dot{1}} & (\sigma^{1\dot{1}})_{2\dot{2}} \end{pmatrix} & \begin{pmatrix} (\sigma^{1\dot{2}})_{1\dot{1}} & (\sigma^{1\dot{2}})_{1\dot{2}} \\ (\sigma^{1\dot{2}})_{2\dot{1}} & (\sigma^{1\dot{2}})_{2\dot{2}} \end{pmatrix} \\ \begin{pmatrix} (\sigma^{2\dot{1}})_{1\dot{1}} & (\sigma^{2\dot{1}})_{1\dot{2}} \\ (\sigma^{2\dot{1}})_{2\dot{1}} & (\sigma^{2\dot{1}})_{2\dot{2}} \end{pmatrix} & \begin{pmatrix} (\sigma^{2\dot{2}})_{1\dot{1}} & (\sigma^{2\dot{2}})_{1\dot{2}} \\ (\sigma^{2\dot{2}})_{2\dot{1}} & (\sigma^{2\dot{2}})_{2\dot{2}} \end{pmatrix} \end{pmatrix} \\ &= 2\begin{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \\ \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \end{pmatrix} . \end{aligned} \tag{11.142}$$

This can be expressed

$$\frac{1}{2}\left(\sigma^{\alpha\dot{\beta}}\right)_{\gamma\dot{\delta}} = \delta_{\alpha\gamma}\delta_{\dot{\beta}\dot{\delta}} . \tag{11.143}$$

Combined with (11.138) one can conclude that the following property holds

$$\frac{1}{2}\left(\sigma_{\alpha\dot{\beta}}\right)^{\gamma\dot{\delta}} = \frac{1}{2}\left(\sigma^{\alpha\dot{\beta}}\right)_{\gamma\dot{\delta}} = \delta_{\alpha\gamma}\delta_{\dot{\beta}\dot{\delta}} . \tag{11.144}$$

The Dirac Matrices γ^μ in spinor notation

We want to express now the Dirac matrices $\tilde{\gamma}^\mu$ in spinor form. For this purpose we start from the expression (11.121) of $\tilde{\gamma}^\mu$. This expression implies that the element of $\tilde{\gamma}^\mu$ given by $\epsilon\,\sigma^\mu\epsilon^{-1}$ is in the basis $|_{\alpha\dot{\beta}}$ whereas the element of $\tilde{\gamma}^\mu$ given by σ^μ is in the basis $|^{\alpha\dot{\beta}}$. Accordingly, we write

$$\tilde{\gamma}^\mu = \begin{pmatrix} 0 & (\sigma^\mu)^{\alpha\dot{\beta}} \\ ((\sigma^\mu)_{\alpha\dot{\beta}})^* & 0 \end{pmatrix} . \tag{11.145}$$

Let A_μ be a covariant 4–vector. One can write then the scalar product using (11.115)

$$\begin{aligned}\tilde{\gamma}^\mu A_\mu &= \begin{pmatrix} 0 & (\sigma^\mu)^{\alpha\dot{\beta}}A_\mu \\ ((\sigma^\mu)_{\alpha\dot{\beta}})^* A_\mu & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & \frac{1}{2}(\sigma_{\gamma\dot{\delta}})^{\alpha\dot{\beta}}A^{\gamma\dot{\delta}} \\ \frac{1}{2}((\sigma^{\gamma\dot{\delta}})_{\alpha\dot{\beta}})^* A_{\gamma\dot{\delta}} & 0 \end{pmatrix} . \end{aligned} \tag{11.146}$$

Exploiting the property (11.144) results in the simple relationship

$$\tilde{\gamma}^\mu A_\mu = \begin{pmatrix} 0 & A^{\alpha\dot{\beta}} \\ A_{\alpha\dot{\beta}} & 0 \end{pmatrix} . \tag{11.147}$$

11.5 Lorentz Invariant Field Equations in Spinor Form

Dirac Equation

(11.147) allows us to rewrite the Dirac equation in the chiral representation (10.226)

$$i\,\gamma^{\mu}\partial_{\mu}\tilde{\Psi}(x^{\mu}) \;=\; \begin{pmatrix} 0 & i\,\partial^{\alpha\dot{\beta}} \\ i\,\partial_{\alpha\dot{\beta}} & 0 \end{pmatrix} \tilde{\Psi}(x^{\mu}) \;=\; m\,\tilde{\Psi}(x^{\mu})\,. \tag{11.148}$$

Employing $\tilde{\Psi}(x^{\mu})$ in the form (11.99) yields the Dirac equation in spinor form

$$i\,\partial^{\alpha\dot{\beta}}\,\chi_{\dot{\beta}} \;=\; m\,\phi^{\alpha} \tag{11.149}$$

$$i\,\partial_{\alpha\dot{\beta}}\,\phi^{\alpha} \;=\; m\,\chi_{\dot{\beta}}\,. \tag{11.150}$$

The simplicity of this equation is striking.

Chapter 12

Symmetries in Physics: Isospin and the Eightfold Way

by *Melih Sener* and *Klaus Schulten*

Symmetries and their consequences are central to physics. In this chapter we will discuss a particular set of symmetries that have played a seminal role in the development of elementary particle and nuclear physics. These are the *isospin symmetry* of nuclear interactions and its natural extension, the so-called *Eightfold Way*.
The organization of this chapter is as follows: In the next section we will discuss the relation between symmetries of a quantum mechanical system and the degeneracies between its energy levels. We will particularly use the example of spherically symmetric potentials. In the following section we will introduce the concept of isospin as an approximate $SU(2)$ symmetry, which identifies the proton and the neutron as different states of the same particle. We will also introduce the quark model as a natural framework to represent the observed symmetries. We will apply these concepts to an analysis of nucleon-nucleon and nucleon-meson scattering. In the final section, we will discuss the $SU(3)$ symmetry of three quark flavors. The algebraic structure and the representations of $SU(3)$ will be discussed in parallel to $SU(2)$ and particle families will be identified in terms of representations of the underlying symmetry group.

12.1 Symmetry and Degeneracies

The degeneracies of energy levels of a quantum mechanical system are related to its symmetries. Let us assume a continuous symmetry obeyed by a quantum mechanical system. The action of the symmetry operations on quantum mechanical states are given by elements of a corresponding Lie group, i.e.,

$$\mathcal{O} = \exp\left(\sum_k \alpha_k S_k\right). \tag{12.1}$$

Then the generators, S_k, will commute with the Hamiltonian of the system,

$$[H, S_k] = 0. \tag{12.2}$$

The action of any symmetry generator, S_k, on an energy eigenstate, $\psi_{E,\lambda_1,\ldots,\lambda_n}$, leaves the energy of the state invariant

$$H \exp(i\alpha_k S_k)\, \psi_{E,\lambda_1,\ldots,\lambda_n} \;=\; E \exp(i\alpha_k S_k)\, \psi_{E,\lambda_1,\ldots,\lambda_n}. \tag{12.3}$$

If the newly obtained state is linearly independent of the original one, this implies a degeneracy in the spectrum. We will investigate this shortly in more detail in the case of systems with spherical symmetry, where the symmetry generators, S_k, can be identified with the angular momentum operators, J_k, as studied in chapter 7.

Lie groups play an essential role in the discussion of mass degeneracies in particle physics. In order to illustrate this, we first consider a particular example of the implications of symmetry, namely motion in a spherically symmetric potential described by the group $SO(3)$ (or its double covering $SU(2)$ as discussed in section 5.12).

In chapter 7 the dynamics of a particle moving in three dimensions under the influence of a spherically symmetric potential, $V(r)$, has been discussed. The spherical symmetry implies the commutation of the Hamiltonian with angular momentum operators (7.8)

$$[\hat{H},\, J_k] \;=\; 0\,, \quad k \;=\; 1,2,3\,. \tag{12.4}$$

The stationary Schrödinger equation, (7.18), can then be reduced to a one-dimensional radial equation, (7.24), which yields a set of eigenstates of the form

$$\psi_{E,\ell,m}(\vec{r}) \;=\; v_{E,\ell,m}(r)\, Y_{\ell m}(\theta,\phi), \tag{12.5}$$

with $m = -l, \cdots, l$ and the corresponding energy levels are independent of m. Therefore, each energy level is $(2l+1)$-fold degenerate. This degeneracy follows from the fact that any rotation, as represented by an element of $SO(3)$, (7.39), generates a state which has the same energy as the original one.

Presence of additional symmetries may further increase the degeneracy of the system. As an example of this we will consider the Coulomb problem with the Hamiltonian

$$H \;=\; \frac{\vec{p}^2}{2m} - \frac{k}{r}. \tag{12.6}$$

From elementary quantum mechanics we know the spectrum of the hydrogen atom. The energy levels are

$$E_n = -\frac{mk^2}{2\hbar^2 n^2}, \tag{12.7}$$

where the orbital angular momentum, l, is allowed to take values in $0, \ldots, n-1$. The energy levels are totally independent of l. For example, the states $3s$, $3p$ and $3d$ all have the same energy. We want to understand this extra degeneracy in terms of the extra symmetry of the hydrogen atom given by an additional set of symmetry generators introduced below.

Classically, the additional symmetry generators of the Coulomb problem are the three components of the so-called *eccentricity* vector discovered by Hamilton

$$\vec{\epsilon} = \frac{1}{m}\vec{p} \times \vec{J} - k\frac{\vec{r}}{r}. \tag{12.8}$$

The vector $\vec{\epsilon}$ points along the symmetry axis of the elliptical orbit and its length equals the eccentricity of the orbit. The vector in (12.8) is not a hermitian operator. The corresponding quantum mechanical hermitian operator can be defined by

$$\begin{aligned}
\epsilon_1 &= \frac{1}{2m}\left(p_2 J_3 - p_3 J_2 + J_3 p_2 - J_2 p_3\right) - k\frac{x_1}{r}, && (12.9)\\
\epsilon_1 &= \frac{1}{2m}\left(J_3 p_2 - J_2 p_3 + i\hbar p_1\right) - k\frac{x_1}{r}, && (12.10)\\
\epsilon_2 &= \frac{1}{2m}\left(J_1 p_3 - J_3 p_1 + i\hbar p_2\right) - k\frac{x_2}{r}, && (12.11)\\
\epsilon_3 &= \frac{1}{2m}\left(J_2 p_1 - J_1 p_2 + i\hbar p_3\right) - k\frac{x_3}{r}, && (12.12)\\
& && (12.13)
\end{aligned}$$

It can be verified explicitly that its components commute with the Hamiltonian.
In order to understand the aforementioned extra degeneracy, we will compute the hydrogen spectrum using the additional symmetry. For this purpose we first note that

$$\vec{J}\cdot\vec{\epsilon} = 0. \tag{12.14}$$

This follows from $\vec{a}\cdot(\vec{a}\times\vec{b}) = (\vec{a}\times\vec{b})\cdot\vec{a} = 0$, which is valid even when $\vec{a}$ and $\vec{b}$ do not commute.
We will also need the following identity [4]

$$\vec{\epsilon}^2 = \frac{2H}{m}\left(\vec{J}^2 + \hbar^2\right) + k^2, \tag{12.15}$$

which can be proved after very considerable algebra.
In the following we consider the bound states, which have a negative energy E. Therefore, in the subspace of the Hilbert space corresponding to a certain energy we can replace H by E. Now we scale the eccentricity vector as follows

$$\vec{K} = \sqrt{-\frac{m}{2E}}\,\vec{\epsilon}. \tag{12.16}$$

Through some algebra [4] the following commutation relations can be verified

$$\begin{aligned}
[K_i, J_j] &= i\hbar\epsilon_{ijk}K_k, && (12.17)\\
[K_i, K_j] &= i\hbar\epsilon_{ijk}J_k, && (12.18)
\end{aligned}$$

which complement the familiar angular momentum algebra of section 5.3.
We introduce the following new operators

$$\vec{A} = \frac{1}{2}(\vec{J} + \vec{K}), \tag{12.19}$$

$$\vec{B} = \frac{1}{2}(\vec{J} - \vec{K}), \tag{12.20}$$

which can be shown to satisfy

$$[A_i, A_j] = i\hbar\epsilon_{ijk}A_k, \tag{12.21}$$

$$[B_i, B_j] = i\hbar\epsilon_{ijk}B_k, \tag{12.22}$$
$$[A_i, B_j] = 0, \tag{12.23}$$
$$\left[\vec{A}, H\right] = 0, \tag{12.24}$$
$$\left[\vec{B}, H\right] = 0. \tag{12.25}$$

So far we have shown that the symmetry generators form an algebra, which is identical to the the direct sum of the Lie algebra of of two $SO(3)$ (or $SU(2)$) algebras. By comparing to the rotation algebra introduced in chapter 5, we can read off the eigenvalues of $\vec{A}^2$ and $\vec{B}^2$ from (12.21) and (12.22):

$$\vec{A}^2 = a(a+1)\hbar^2\ ,\ a = 0, \frac{1}{2}, 1, \dots\ , \tag{12.26}$$
$$\vec{B}^2 = b(b+1)\hbar^2\ ,\ b = 0, \frac{1}{2}, 1, \dots\ . \tag{12.27}$$

Following (12.14) we note that

$$\vec{A}^2 - \vec{B}^2 = \vec{J}\cdot\vec{\epsilon} = 0. \tag{12.28}$$

This implies that $a = b$. In order to arrive at the spectrum a final bit algebra is needed

$$\vec{A}^2 + \vec{B}^2 = \vec{J}^2 + \vec{K}^2 \tag{12.29}$$
$$= \vec{J}^2 - \frac{m}{2E}\vec{\epsilon}^2 \tag{12.30}$$
$$= -\frac{mk^2}{4E} - \frac{1}{2}\hbar^2, \tag{12.31}$$

where we have used (12.15). Using this equation the energy eigenvalues can be written in terms of the eigenvalues of $\vec{A}^2$ and $\vec{B}^2$ operators. Noticing that $\vec{A}^2$ and $\vec{B}^2$ have the same eigenvalues because of (12.28), the energy eigenvalues are found to be

$$E = -\frac{mk^2}{2\hbar^2(2a+1)^2},\ a = 0, \frac{1}{2}, 1, \dots\ . \tag{12.32}$$

A comparison with (12.7) tells us that $(2a+1) = n$. Furthermore the bound on the orbital angular momentum, l, can be seen to follow from the triangle inequality as applied to $\vec{J} = \vec{A} + \vec{B}$, namely that

$$\left|\vec{J}\right| > \left|\vec{A}\right| - \left|\vec{B}\right| = 0 \tag{12.33}$$
$$\left|\vec{J}\right| < \left|\vec{A}\right| + \left|\vec{B}\right| = 2\left|\vec{A}\right| \tag{12.34}$$

It follows that l has to have values in $\{0 = |a-b|, 1, \dots, a+b = n-1\}$. This illustrates the effect of additional symmetries to the degeneracy structure of a quantum mechanical system.

In contrast to the discussion above about extra symmetries, a lack of symmetry implies a lack of degeneracy in the energy levels of a quantum mechanical system. The most extreme case of this is the quantum analogue of a classically chaotic system. Chaos is described classically as exponential

sensitivity to initial conditions, in the sense that nearby trajectories in the phase space diverge from each other over time. However, another manifestation of chaos is the lack of independent operators commuting with the Hamiltonian.
A typical example of this so called quantum chaos is the quantum billiard problem, which is a particle in box problem in two dimensions with a boundary which can be chosen arbitrarily. If the chosen boundary is 'irregular' in a suitably defined sense, the classical trajectories will diverge from each other after successive bounces from the boundary. For a more detailed discussion of quantum chaos in billiard systems we refer the reader to [7] and the references therein.
In the case of billiards and other examples of quantum chaos one common observation is the almost nonexistence of degeneracies and the fact that the energy levels are more evenly spaced. This is known as *level repulsion*.

In the next section we will proceed with the discussion of a symmetry, which was discovered by observing degeneracies in the particle spectrum.

12.2 Isospin and the $SU(2)$ flavor symmetry

The concept of *isospin* goes back to Heisenberg, who, after the discovery of the neutron in 1932, suggested that the proton and the neutron can be regarded as two states of a single particle. This was motivated by the observation that their masses are approximately equal: $m_p = 938.28 MeV/c^2$, $m_p = 939.57 MeV/c^2$. Following the mass-energy equivalence of special relativity

$$E \;=\; mc^2, \tag{12.35}$$

this mass equivalence can be viewed as an energy degeneracy of the underlying interactions.
This (approximate) degeneracy led into the idea of the existence of an (approximate) symmetry obeyed by the underlying nuclear interactions, namely, that the proton and the neutron behave identically under the so-called *strong interactions* and that their difference is solely in their charge content. (Strong interactions bind the atomic nucleus together.)
If the proton and the neutron are to be viewed as two linearly independent states of the same particle, it is natural to represent them in terms of a two component vector, analogous to the spin-up and spin-down states of a spin-$\frac{1}{2}$ system

$$p = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \; n = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{12.36}$$

In analogy to the concept of spin regarding the rotations in 3-space as discussed in chapter 5, the isospin symmetry is also governed by an $SU(2)$ group 'rotating' components in (12.36) into each other in abstract isospin space. This enables us to utilize what we already know about the $SU(2)$ symmetry group from the study of angular momentum. For example, we will be able to use the familiar Clebsch-Gordan coefficients to combine the isospin of two particles the same way we added spin in chapter 6.
It is important to place the isospin concept in its proper historical context. Originally it was believed that isospin was an exact symmetry of strong interactions and that it was violated by electromagnetic and weak interactions. (Weak interactions are responsible, for example, for the

beta decay). The mass difference between the neutron and the proton could then be attributed to the charge content of the latter. If the mass difference (or the energy difference) were to be to be purely electrostatic in nature, the proton had to be heavier. However, the proton is the lighter of the two. If it were otherwise the proton would be unstable by decaying into the neutron, spelling disaster for the stability of matter.
Isospin symmetry is *not* an exact symmetry of strong interactions, albeit it is a good approximate one. Therefore it remains a useful concept. Further than that, as we shall see below, it can be seen as part of a larger (and more approximate) symmetry which is of great utility to classify observed particle families.

We can describe isospin multiplets the same way we have described the angular momentum and spin multiplets. Denoting the total isospin, I, and its third component, I_3, as good quantum numbers, we can re-write (12.36) as a multiplet with $I = \frac{1}{2}$

$$p = \left| I = \frac{1}{2}, I_3 = \frac{1}{2} \right\rangle , \; n = \left| I = \frac{1}{2}, I_3 = -\frac{1}{2} \right\rangle . \tag{12.37}$$

As an example of a multiplet with $I = 1$ we have the three pions or π-mesons

$$\pi^+ = |1, 1\rangle , \; \pi^0 = |1, 0\rangle , \; \pi^- = |1, -1\rangle , \tag{12.38}$$

which have all nearly identical masses. ($m_{\pi^\pm} = 139.6 MeV/c^2$, $m_{\pi^0} = 135.0 MeV/c^2$). Shortly we will see how to describe both the pion and nucleon states as composites of more fundamental $I = \frac{1}{2}$ states.
In the framework of the quark of model, the fundamental representation of the isospin symmetry corresponds to the doublet that contains the so-called *up* and *down* quarks

$$u = \left| \frac{1}{2}, \frac{1}{2} \right\rangle , \; d = \left| \frac{1}{2}, -\frac{1}{2} \right\rangle . \tag{12.39}$$

All other isomultiplets, including the proton and the neutron, are made up of these two quarks. They can be constructed with the same rules that have been used for angular momentum addition in chapter 6. For example, the three pions in (12.38) are $u\bar{d}$, $u\bar{u}$ and $d\bar{u}$ states, respectively. They form an isotriplet:

$$\pi^+ = |1, 1\rangle = \left| \frac{1}{2}, \frac{1}{2} \right\rangle_1 \left| \frac{1}{2}, \frac{1}{2} \right\rangle_2 , \tag{12.40}$$

$$\pi^0 = |1, 0\rangle = \frac{1}{\sqrt{2}} \left(\left| \frac{1}{2}, \frac{1}{2} \right\rangle_1 \left| \frac{1}{2}, -\frac{1}{2} \right\rangle_2 + \left| \frac{1}{2}, \frac{1}{2} \right\rangle_1 \left| \frac{1}{2}, -\frac{1}{2} \right\rangle_2 \right) , \tag{12.41}$$

$$\pi^- = |1, -1\rangle = \left| \frac{1}{2}, -\frac{1}{2} \right\rangle_1 \left| \frac{1}{2}, -\frac{1}{2} \right\rangle_2 . \tag{12.42}$$

$$\tag{12.43}$$

Similarly, the proton and the neutron can be written as totally symmetric *uud* and *udd* states. For a precise description of the two nucleons as composite states, including the spin and color quantum numbers of their constituent quarks, we refer the reader to [1], sec. 2.11.

The mass of the up and down quarks are *not* identical but they are both of the order of a few MeV/c^2's which is minuscule compared to the typical energy scale of hadrons (i.e. strongly interacting particles) which is about a GeV/c^2. This is why isospin is such a good symmetry and why isomultiplets have nearly identical masses.
As it later turned out, the up and down quarks are not the only quark 'species' - or *flavors* as they are commonly called. In the late 1940's and early 1950's, a number *strange* particles have been found which presumably contained a third quark species: the strange quark. It shall be noted here that the quark model was not invented until 1960's, but at the time the empirical concepts like isospin and *strangeness* quantum numbers were in use. The value of the strangeness quantum number is taken, by accidental convention, to be -1 for the strange quark. The up and down quarks have strangeness zero. All other composite states have their strangeness given by the sum of the strangeness content of their constituents.
Before proceeding further, we shall setup some terminology: *baryons* are qqq states, such as proton and the neutron, whereas *mesons* are $q\bar{q}$ states, the pions being examples thereof. By convention baryons have *baryon number* 1, and quarks have baryon number $\frac{1}{3}$. All antiparticles have their quantum numbers reversed. Naturally, mesons have baryon number 0. The names, baryon and meson, originally refer to the relative weight of particles, baryons generally are heavy, mesons have intermediate mass ranges, where *leptons* (electron, muon, the neutrinos etc.) are light. If taken literally, this remains only an inaccurate naming convention today, as some mesons discovered later are heavier than some baryons and so on.
The relation between electric charge and isospin are given by the *Gell-Mann–Nishijima relation* which was first derived empirically

$$Q = I_3 + \frac{1}{2}(B + S), \tag{12.44}$$

where B is the baryon number and S is the strangeness. In the next section we will be able to view the Gell-Mann–Nishijima relation in the light of the representation theory for the flavor $SU(3)$ symmetry.

Now let us consider another example of combining the isospins of two particles. The reader may know that the *deuteron*, a hydrogen isotope, consists of a proton and a neutron. Therefore it has to have isospin, $I_3 = 0$. We will now try to describe its wave function in terms of its constituent nucleons. Following (12.37) and in analogy to (12.43), this will be mathematically identical to adding two spins. The possibilities are that of an *isosinglet*

$$|0,0\rangle = \frac{1}{\sqrt{2}}(|p>|n> - |n>|p>) \tag{12.45}$$

and that of an *isotriplet*

$$|1,1\rangle = |p>|p>, \tag{12.46}$$

$$|1,0\rangle = \frac{1}{\sqrt{2}}(|p>|n> + |n>|p>), \tag{12.47}$$

$$|1,-1\rangle = |n>|n>. \tag{12.48}$$

Is the deuteron an isosinglet state or an isotriplet? If it were an isotriplet ($|1,0\rangle$) we should have seen nn and pp bound states of comparable energy in nature (because of isospin symmetry), but such states do not exist. Therefore the deuteron has to be an isosinglet state ($|0,0\rangle$).

As an exercise on the implications of isospin symmetry we will consider nucleon-nucleon scattering. We will eventually be able to compute ratios of scattering cross-sections between different processes. For example, consider

$$\begin{array}{l} \text{(I) } p + p \rightarrow d + \pi^+ \\ \text{(II) } p + n \rightarrow d + \pi^0 \end{array} \tag{12.49}$$

The only assumption that we put in will be that the interaction is of the form $V = \alpha\, \mathbf{I}^{(1)} \cdot \mathbf{I}^{(2)}$. The dot product here refers to the abstract isospin space. The cross-section, σ, is proportional to $|\mathcal{M}|^2$, where $\mathcal{M}$ is the scattering amplitude given by

$$\mathcal{M} = \langle \text{ final } |\ \alpha\, \mathbf{I}^{(1)} \cdot \mathbf{I}^{(2)}\ | \text{ initial } \rangle\,. \tag{12.50}$$

The initial and final states can be denoted in more detail as

$$| \text{ initial } \rangle \ = \ \left| \, I^{(i)}, I_3^{(i)}, \gamma^{(i)} \, \right\rangle, \tag{12.51}$$

$$| \text{ final } \rangle \ = \ \left| \, I^{(f)}, I_3^{(f)}, \gamma^{(f)} \, \right\rangle, \tag{12.52}$$

where $\gamma^{(i)}$ and $\gamma^{(f)}$ denote degrees of freedom other than isospin, such as the spatial dependence of the wave function and spin.

Exercise. Consider the generalization of tensor operators discussed in section 6.7 to the case of isospin. Show that $\mathbf{I}^{(1)} \cdot \mathbf{I}^{(2)}$ is an 'isotensor' of rank zero. Refer to exercise 6.7.5 for the spin-analogue of the same problem.

Exercise. Show that the expectation of $\mathbf{I}^{(1)} \cdot \mathbf{I}^{(2)}$ is $\frac{1}{4}$ in an isotriplet state and $-\frac{3}{4}$ in an isosinglet state.

Using (12.38) and the fact that the deuteron is an isosinglet we know that the isospins of the final states in (I) and (II) are $|1,1\rangle$ and $|1,0\rangle$, respectively. According to (12.45) and (12.48) the initial states in (I) and (II) have isospin values $|1,1\rangle$ and $(1/\sqrt{2})(|1,0\rangle + |0,0\rangle)$. We will now employ the (isospin analogue) of the Wigner-Eckart theorem (6.259) discussed in detail in sections 6.7 and 6.8 to compute the ratio of the scattering amplitudes $\mathcal{M}_{\rm I}$ and $\mathcal{M}_{\rm II}$. For completeness let us start by restating the Wigner-Eckart theorem (6.259) in the present context:

$$\langle I^{(f)} I_3^{(f)}, \gamma^{(f)} | T_{00} | I^{(i)} I_3^{(i)}, \gamma^{(i)} \rangle \ = \tag{12.53}$$

$$(I^{(f)} I_3^{(f)} | 00 I^{(i)} I_3^{(i)})\ (-1)^{I^{(f)} - I^{(i)}}\ \frac{1}{\sqrt{2I^{(i)}+1}}\ \langle I^{(f)}, \gamma^{(f)} || T_0 || I^{(i)}, \gamma^{(i)} \rangle. \tag{12.54}$$

Here $T_{00} \equiv V = \alpha\, \mathbf{I}^{(1)} \cdot \mathbf{I}^{(2)}$, which is an isoscalar, as discussed in the exercise above. $(I^{(f)} I_3^{(f)} | 00 I^{(i)} I_3^{(i)})$ is a Clebsch-Gordon coefficient and $\langle I^{(f)}, \gamma^{(f)} || T_0 || I^{(i)}, \gamma^{(i)} \rangle$ is a reduced matrix element defined in the same sense as in section 6.8.
Now let us re-write more carefully the scattering amplitudes for the two processes in the light of what we have just learned

$$\mathcal{M}_{\rm I} \ = \ \left\langle \, I^{(f)} = 1, I_3^{(f)} = 1, \gamma^{(f)} \, \right| T_{00} \left| \, I^{(i)} = 1, I_3^{(i)} = 1, \gamma^{(i)} \, \right\rangle \tag{12.55}$$

$$= (11|0011)\,\frac{1}{\sqrt{3}}\,\langle I^{(f)}=1,\gamma^{(f)}\,||T_{00}||\,I^{(i)}=1,\gamma^{(i)}\rangle \tag{12.56}$$

$$\mathcal{M}_{\rm II} = \frac{1}{\sqrt{2}}\left\langle\, I^{(f)}=1,I_3^{(f)}=0,\gamma^{(f)}\,\middle|\,T_{00}\,\middle|\,I^{(i)}=1,I_3^{(i)}=0,\gamma^{(i)}\,\right\rangle \tag{12.57}$$

$$+\frac{1}{\sqrt{2}}\left\langle\, I^{(f)}=1,I_3^{(f)}=0,\gamma^{(f)}\,\middle|\,T_{00}\,\middle|\,I^{(i)}=0,I_3^{(i)}=0,\gamma^{(i)}\,\right\rangle \tag{12.58}$$

$$= (10|0010)\,\frac{1}{\sqrt{3}}\,\langle I^{(f)}=1,\gamma^{(f)}\,||T_{00}||\,I^{(i)}=1,\gamma^{(i)}\rangle \tag{12.59}$$

$$+0. \tag{12.60}$$

$$\tag{12.61}$$

Note that the second term in $\mathcal{M}_{\rm II}$ vanishes due to the isospin conservation, which is also manifested by a vanishing Clebsch-Gordon prefactor. The relevant Clebsch-Gordon coefficients are easily evaluated:

$$(11|0011) \;=\; (10|0010) \;=\; 1. \tag{12.62}$$

We can now write the ratio of the scattering amplitudes:

$$\frac{\mathcal{M}_{\rm I}}{\mathcal{M}_{\rm II}} = \frac{1}{(1/\sqrt{2})}\,, \tag{12.63}$$

where common dynamical factors (which would not be as easy to compute) have dropped out thanks to the Wigner-Eckart theorem. It follows

$$\frac{\sigma_{\rm I}}{\sigma_{\rm II}} = 2, \tag{12.64}$$

which is in approximate agreement with the observed ratio. [2]

As a further example, we will consider pion-nucleon scattering. We want to compute the ratio of *total* cross-sections assuming a similar interaction as in the previous example

$$\frac{\sigma(\,\pi^+ + p \,\rightarrow\, \text{anything}\,)}{\sigma(\,\pi^- + p \,\rightarrow\, \text{anything}\,)}\,.$$

The possibilities are

$$\begin{array}{l}\text{(a)}\;\pi^+ + p \,\rightarrow\, \pi^+ + p\,,\\ \text{(b)}\;\pi^- + p \,\rightarrow\, \pi^- + p\,,\\ \text{(c)}\;\pi^- + p \,\rightarrow\, \pi^0 + n\,.\end{array} \tag{12.65}$$

There are more exotic possibilities, involving, for example, particles with strangeness, but these are not dominant at relatively low energies.
Once again we need the isospins for the initial and final states, which are obtained by a standard Clebsch-Gordan expansion

$$\begin{aligned}&\pi^+ + p \,:\, |1,1\rangle\,\big|\tfrac{1}{2},\tfrac{1}{2}\big\rangle = \big|\tfrac{3}{2},\tfrac{3}{2}\big\rangle\;,\\ &\pi^- + p \,:\, |1,-1\rangle\,\big|\tfrac{1}{2},\tfrac{1}{2}\big\rangle = \tfrac{1}{\sqrt{3}}\,\big|\tfrac{3}{2},-\tfrac{1}{2}\big\rangle - \sqrt{\tfrac{2}{3}}\,\big|\tfrac{1}{2},-\tfrac{1}{2}\big\rangle\;,\\ &\pi^0 + n \,:\, |1,0\rangle\,\big|\tfrac{1}{2},-\tfrac{1}{2}\big\rangle = \sqrt{\tfrac{2}{3}}\,\big|\tfrac{3}{2},-\tfrac{1}{2}\big\rangle + \tfrac{1}{\sqrt{3}}\,\big|\tfrac{1}{2},-\tfrac{1}{2}\big\rangle\;.\end{aligned} \tag{12.66}$$

As in the example of nucleon-meson scattering we define the relevant matrix elements

$$\begin{aligned}\mathcal{M}_{\frac{3}{2}} &= \left\langle \tfrac{3}{2}, m \;\middle|\; V \;\middle|\; \tfrac{3}{2}, m \right\rangle, \\ \mathcal{M}_{\frac{1}{2}} &= \left\langle \tfrac{1}{2}, m \;\middle|\; V \;\middle|\; \tfrac{1}{2}, m \right\rangle,\end{aligned} \tag{12.67}$$

which are independent of m. A computation similar to the previous example of nucleon-nucleon scattering yields (apart from common prefactors) the following amplitudes for the reactions in (12.65)

$$\begin{aligned}\mathcal{M}_{\mathrm{a}} &= \mathcal{M}_{\frac{3}{2}} \\ \mathcal{M}_{\mathrm{b}} &= \tfrac{1}{3}\mathcal{M}_{\frac{3}{2}} + \tfrac{2}{3}\mathcal{M}_{\frac{1}{2}} \\ \mathcal{M}_{\mathrm{c}} &= \tfrac{\sqrt{2}}{3}\mathcal{M}_{\frac{3}{2}} - \tfrac{\sqrt{2}}{3}\mathcal{M}_{\frac{1}{2}}\end{aligned} \tag{12.68}$$

Guided by empirical data we will further assume that $\mathcal{M}_{\frac{3}{2}} >> \mathcal{M}_{\frac{1}{2}}$, which leads to the following ratios for the cross-sections

$$\sigma_a : \sigma_b : \sigma_c = 9 : 1 : 2. \tag{12.69}$$

As the total cross-section is the sum of individual processes we obtain

$$\frac{\sigma(\,\pi^+ + p\,)}{\sigma(\,\pi^- + p\,)} = \frac{\sigma_a}{\sigma_b + \sigma_c} = 3 \tag{12.70}$$

again in approximate agreement with the observed value. [2]

12.3 The Eightfold Way and the flavor $SU(3)$ symmetry

The discovery of the concept of *strangeness*, mentioned in the previous section, was motivated by the existence of particles that are produced strongly but decay only weakly. For instance, K^+, which can be produced by $\pi^- + p \to K^+ + \Sigma^-$, has a lifetime which is comparable to that of π^+ albeit being more than three times heavier. Hence Gell-Mann and independently Nishijima postulated the existence of a separate quantum number, S, called strangeness, such that $S(K^+) = 1$, $S(\Sigma^-) = -1$ and $S(\pi) = S(N) = 0$, etc. It was assumed that strong interactions conserved S (at least approximately), while weak interactions did not. Hence the strangeness changing strong decays of K^+ (or Σ^-) were forbidden, giving it a higher than usual lifetime.
The classification of the newly found particles as members of some higher multiplet structure was less obvious then the case of isospin, however. Strange partners of the familiar nucleons, for example, are up to 40% heavier, making an identification of the underlying symmetry and the multiplet structure less straightforward.
In the light of the quark model, it appears an obvious generalization to add another component for an extra quark to the isospin vector space

$$u = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, d = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, s = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{12.71}$$

In this case the transformations that 'rotate' the components of (12.71) into each other, while preserving the norm, have to be elements of the group $SU(3)$ (which we will investigate closely very soon). However, history followed the reverse of this path. *First* particle multiplets were identified as representations of the $SU(3)$ group, the same way nucleons and pions were identified as representations of the isospin $SU(2)$ symmetry. *Then* came the question of what the fundamental representation, as in (12.71), should correspond to, giving rise to the quark model. As quarks were never directly observed, for a period they were considered as useful bookkeeping devices without physical content.
In this perspective the flavor $SU(3)$ symmetry may appear to be mainly of historical interest. However $SU(3)$ symmetry appears in another and much more fundamental context in strong interaction physics. The quarks posses another quantum number, called *color*, which again form representations of an $SU(3)$ group. This is believed to be an *exact* symmetry of strong interactions, in fact modern theory of strong interactions is a 'gauge theory' of this color group, called *quantum chromodynamics*. (The reader is referred to section 8.3 for a brief discussion of gauge transformations). Flavor $SU(N_f)$ symmetry on the other hand, where N_f is the number quark flavors, becomes increasingly inaccurate for $N_f > 3$. The reason is that the other known quarks, namely charm, bottom (beauty) and top (truth) are significantly heavier than the hadronic energy scale. (The 'bare' mass of the *charm* quark is already heavier than the two nucleons, which set the hadronic energy scale. The bottom and top are even heavier. [2])

Before discussing the significance and the physical implications of the quark model, we will establish some mathematical preliminaries about the group $SU(3)$. In many respects it will resemble the more familiar group $SU(2)$ discussed in some detail in chapter 5, but there are a number of subtle differences. The reader shall note that most of what is being said trivially generalizes to other unitary groups, $SU(N)$, but we will stick to $N = 3$ in the following. The reader is also invited to revisit section 5.1 whenever necessary, in reference to Lie groups, Lie algebras and related concepts.

Given a complex vector, a_k, of three dimensions, we want to find those transformations

$$a_k \to U_{kl} a_l \tag{12.72}$$

that preserve the norm, $\sum_k a_k^* a_k$, of a. It is seen that such a matrix U has to satisfy the following unitary relation

$$U^\dagger = U^{-1}. \tag{12.73}$$

To verify that all such matrices form a group, we observe that

$$(UV)^\dagger = V^\dagger U^\dagger = V^{-1} U^{-1} = (UV)^{-1}, \tag{12.74}$$

for any two unitary matrices U and V. This group of 3×3 unitary matrices is denoted by $U(3)$. The unitarity relation imposes 9 constraints on the total of 18 real degrees of freedom of a 3×3 complex matrix. Hence the group $U(3)$ has 9 dimensions. Multiplying U by a phase, $e^{i\phi}$, still leaves the norm invariant. Therefore $U(3)$ can be decomposed into a direct product $U(1) \times SU(3)$ where $SU(3)$ consists of 3×3 unitary matrices of unit determinant. Because of this additional constraint $SU(3)$ has 8 dimensions. Since arbitrary phase factors are of no physical interest, it is the group

$SU(3)$ and not $U(3)$ that is of main interest. The reader is invited to compare the structure of $SU(3)$ to that of $SU(2)$ discussed in section 5.7.

As discussed in section 5.1, any unitary matrix, U, can be written in the form

$$U = e^{iH} \tag{12.75}$$

where H is a hermitian matrix. Therefore we will express elements of $SU(3)$ as

$$U = e^{i\sum_k \alpha_k \lambda_k} \tag{12.76}$$

where λ_k are 8 linearly independent matrices forming the basis of the Lie algebra of $SU(3)$. (We shall at times refer to the Lie algebra with the name of the group, the meaning being apparent from the context.) The unit determinant condition requires that all λ_k are traceless, since $\det(e^A) = e^{\mathrm{tr}A}$. An explicit basis is constructed in analogy to the Pauli algebra of spin operators

$$\begin{aligned}
\lambda_1 &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\
\lambda_4 &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \\
\lambda_6 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \\
\lambda_8 &= \tfrac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.
\end{aligned} \tag{12.77}$$

The generators, λ_k, obey the following relation

$$\mathrm{tr}(\lambda_j \lambda_k) = 2\delta_{jk}, \tag{12.78}$$

which can be verified explicitly as matrix identities.
The Lie algebra structure is given by the commutators of λ_k

$$[\lambda_j, \lambda_k] = 2i f_{jkl} \lambda_l, \tag{12.79}$$

where f_{jkl} are the antisymmetric structure constants similar to the ϵ_{jkl} of $SU(2)$ given in (5.32). We can also introduce the constants, δ_{jkl}, via the anticommutator relation

$$[\lambda_j, \lambda_k]_+ = \frac{4}{3}\delta_{jk} + 2\delta_{jkl}\lambda_l, \tag{12.80}$$

This is the *fundamental* or defining representation of $SU(3)$. As in the case of $SU(2)$ higher dimensional representations obeying the same structure can be found. The fundamental relation to be preserved is (12.79), regardless of the dimension of the representation.

As it turns out the set of generators in (12.77) is not the most useful basis in the study of $SU(3)$. In the case of $SU(2)$ the identification of 'ladder' operators $\{J_+, J_-\}$ proved useful, which satisfied an 'eigenvalue equation'

$$[J_0, J_\pm] = \pm\hbar J_\pm. \tag{12.81}$$

In chapter 5, these relations have been used to construct the angular momentum spectrum as well as the function space representation of the rotation algebra, namely spherical harmonics. The generators of $SU(3)$ can be arranged into a very similar form to that of $SU(2)$. We first introduce the *F-spin* operators

$$F_i = \frac{1}{2}\lambda_i. \tag{12.82}$$

With another change of basis we arrive at the 'standard' form of the generators of the Lie algebra of $SU(3)$

$$\begin{aligned} T_\pm &= F_1 \pm iF_2, && (12.83)\\ T_0 &= F_3, && (12.84)\\ V_\pm &= F_4 \pm iF_5, && (12.85)\\ U_\pm &= F_6 \pm iF_7, && (12.86)\\ Y &= \frac{2}{\sqrt{3}}F_8. && (12.87) \end{aligned}$$

Exercise. Using the convention in (12.71) show that T_0 is the isospin operator, I_3.

Exercise. Derive the Gell-Mann–Nishijima relation (12.44), starting with the observation that $Y = B + S$. (Recall that the strange quark has $S = -1$ by convention). Y is called the *hypercharge*.

In the basis (12.87) the commutation relations between the generators can be expressed in a succinct manner. First, we have

$$[Y, T_0] = 0, \tag{12.88}$$

which defines (not uniquely) a maximal set of mutually commuting operators $\{Y, T_0\}$ and

$$\begin{aligned} [Y, T_\pm] &= 0, && (12.89)\\ [Y, U_\pm] &= \pm U_\pm, && (12.90)\\ [Y, V_\pm] &= \pm V_\pm, && (12.91) \end{aligned}$$

$$\begin{aligned} [T_0, T_\pm] &= \pm T_\pm, && (12.92)\\ [T_0, U_\pm] &= \mp\frac{1}{2}U_\pm, && (12.93)\\ [T_0, V_\pm] &= \pm\frac{1}{2}V_\pm, && (12.94) \end{aligned}$$

which relate the remaining generators $\{T_\pm, V_\pm, U_\pm\}$ to this maximal set by 'eigenvalue equations' and

$$[T_+, T_-] = 2T_0, \tag{12.95}$$
$$[U_+, U_-] = \frac{3}{2}Y - T_0 = 2U_0, \tag{12.96}$$
$$[V_+, V_-] = \frac{3}{2}Y + T_0 = 2V_0, \tag{12.97}$$

which relate commutators of generators with opposite eigenvalues to the maximal set $\{Y, T_0\}$. Note that, U_0 and V_0 are linear combinations of T_0 and Y. Finally, we have

$$[T_+, V_-] = -U_-, \tag{12.98}$$
$$[T_+, U_+] = V_+, \tag{12.99}$$
$$[U_+, V_-] = T_-, \tag{12.100}$$
$$[T_+, V_+] = 0, \tag{12.101}$$
$$[T_+, U_-] = 0, \tag{12.102}$$
$$[U_+, V_+] = 0. \tag{12.103}$$

Any remaining commutators follow from hermiticity.
The same way the angular momentum ladder operators have been used to construct the representations of $SU(2)$, we will use these commutation relations to construct representations of $SU(3)$. In the case of $SU(2)$ the representations lay on a line on the J_0 axis. However, since there are two mutually commuting generators in $SU(3)$ as given in (12.88), the representations will now lie in a $T_0 - Y$-plane. The maximum number of mutually commuting generators of a Lie algebra is called its *rank*. Thus, $SU(2)$ has rank 1, while $SU(3)$ has rank 2.
When the basis of a Lie algebra is expressed in such a way to satisfy the form of the eigenvalue relations as given above, it is said to be in *Cartan-Weyl form*. This form is essential for easy labeling of the representations of the group, as the relation between the states in a given representation can be conveniently expressed in terms of ladder operators. A formal definition and a detailed discussion of the Cartan-Weyl form is beyond the scope of this chapter. The interested reader is instead referred to a very readable account given in chapter 12 of [4].

Another important property of $SU(2)$ is the existence of an operator (namely the total angular momentum, J^2) which commutes with *all* of the generators. An operator which commutes with all generators of a Lie group is called a *Casimir operator*. As in the case of J^2 and $SU(2)$, Casimir operators can be used to label irreducible representations of the Lie algebra, similar to the way it was done in section 5.5. We can construct two such independent Casimir operators for the group $SU(3)$.

$$C_1 = \sum_k \lambda_k^2 = -\frac{2i}{3}\sum_{jkl} f_{jkl}\lambda_j\lambda_k\lambda_l, \tag{12.104}$$
$$C_2 = \sum_{jkl} d_{jkl}\lambda_j\lambda_k\lambda_l \tag{12.105}$$

In general the number of independent Casimir operators of a Lie group is equal to its rank.

The utility of Casimir operators arises from the fact that all states in a given representation assume the same value for a Casimir operator. This is because the states in a given representation are connected by the action of the generators of the Lie algebra and such generators commute with the Casimir operators. This property can be used to label representations in terms of the values of the Casimir operators. For example, it was shown in section 5.5 how to label the irreducible representations of the angular momentum algebra $SU(2)$ in terms of the value of the total angular momentum.

Now we will construct explicit representations of $SU(3)$. Because of (12.88) we can label states by the eigenvalues of T_0 and Y operators, $|t_3, y\rangle$:

$$T_0 |t_3, y\rangle = t_3 |t_3, y\rangle , \tag{12.106}$$
$$Y |t_3, y\rangle = y |t_3, y\rangle . \tag{12.107}$$

From the commutation relations we have presented above (the Cartan-Weyl form) we can write down the effect of various generators on the state $|t_3, y\rangle$. For example, we have

$$U_0 |t_3, y\rangle = (\frac{3}{4}y - \frac{1}{2}t_3) |t_3, y\rangle , \tag{12.108}$$
$$V_0 |t_3, y\rangle = (\frac{3}{4}y + \frac{1}{2}t_3) |t_3, y\rangle . \tag{12.109}$$

The same way that $J_\pm |m\rangle$ is proportional to $|m \pm 1\rangle$ in the case of the angular momentum algebra, we have

$$T_\pm |t_3, y\rangle = \alpha \left| t_3 \pm \frac{1}{2}, y \right\rangle , \tag{12.110}$$
$$U_\pm |t_3, y\rangle = \beta \left| t_3 \pm \frac{1}{2}, y \pm 1 \right\rangle , \tag{12.111}$$
$$V_\pm |t_3, y\rangle = \gamma \left| t_3 \mp \frac{1}{2}, y \pm 1 \right\rangle . \tag{12.112}$$

The effect of these operators to the states in the $y - t_3$ plane have been outlined in Fig. (12.1).
The representations of $SU(3)$ are constructed analogous to those of $SU(2)$ by identifying the 'boundary' states annihilated by raising (or lowering) operators. All other states of the representation are then constructed by successive application of ladder operators $T_\pm, U_\pm, V_\pm$. The representations for hexagons with sides of length p and q in the $T_0 - Y$-plane. Such a representation is labeled as $D(p, q)$ and it has a dimensionality of $\frac{1}{2}(p + 1)(q + 1)(p + q + 2)$. Figure (12.2) shows the representation $D(2, 1)$ as an example. The details of this procedure is beyond the scope of this chapter. The interested reader is referred to [4], especially chapters 7 and 8.
As another example for the representations of $SU(3)$, the pion family forms part of an *octet* corresponding to the $D(1, 1)$ representation. The representations $D(1, 0)$ and $D(0, 1)$ correspond to the triplets of quarks and antiquarks, respectively. (See Fig. (12.3).) All other representations can be constructed by combining these two conjugate representations. For example the pion octet (or any other meson octet) is therefore realized as states of a quark - antiquark pair. A notation suggestive of the dimensionality of the representation can be used to identify representations. For example,

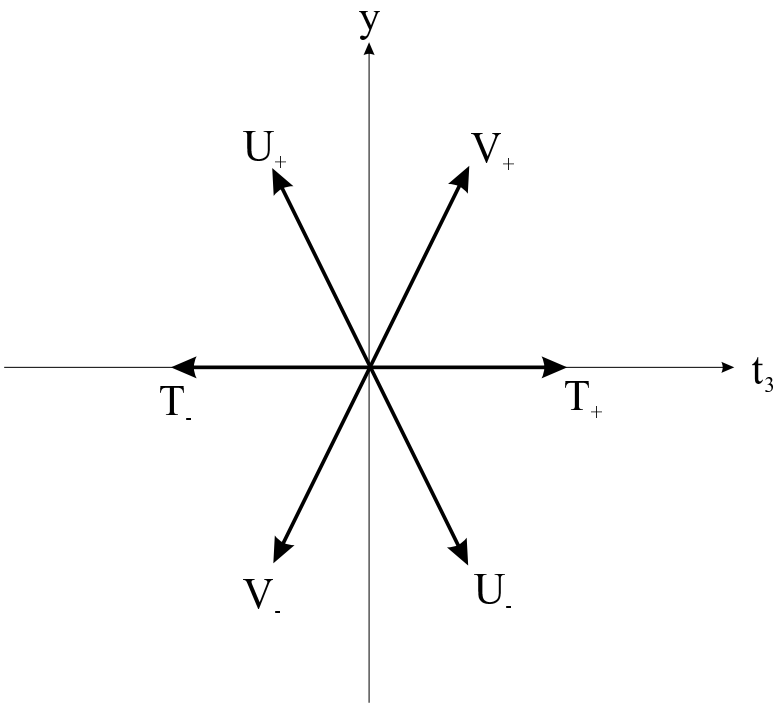


Figure 12.1: The effect of $SU(3)$ ladder operators on the $y - t_3$ plane.

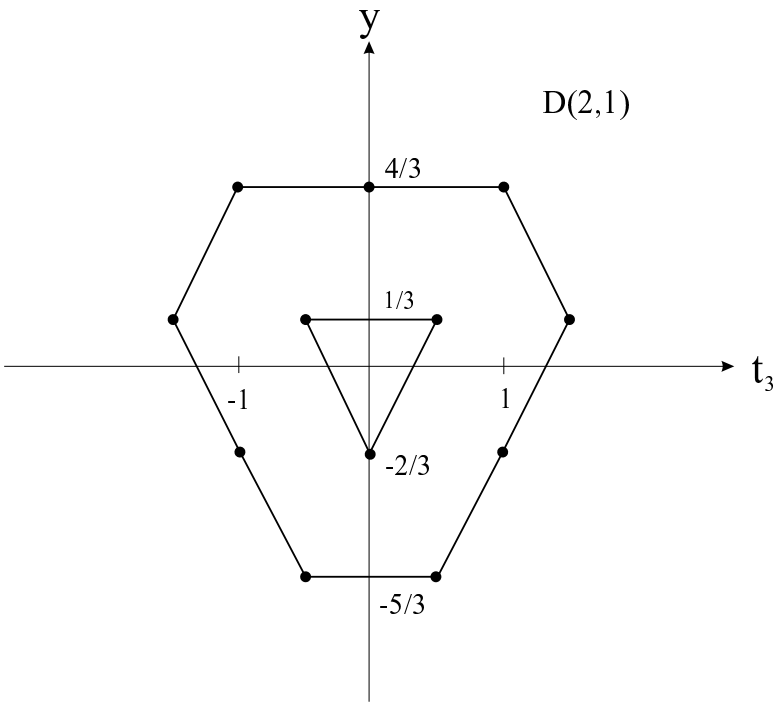


Figure 12.2: $D(2,1)$ representation of $SU(3)$. The states in the inner triangle are doubly degenerate.

$D(1,0)$ and $D(0,1)$ are denoted by [3] and $[\bar{3}]$, respectively. The octet $D(1,1)$ is written as [8] etc. This way the quark-antiquark states can be represented as follows

$$[3] \otimes [\bar{3}] = [8] \oplus [1] \tag{12.113}$$

The additional singlet state corresponds to the η' meson. This expansion can be compared, for example, to the case of adding two spin triplet states, in the case of $SU(2)$, where we would write

$$[3] \otimes [3] = [1] \oplus [3] \oplus [5]. \tag{12.114}$$

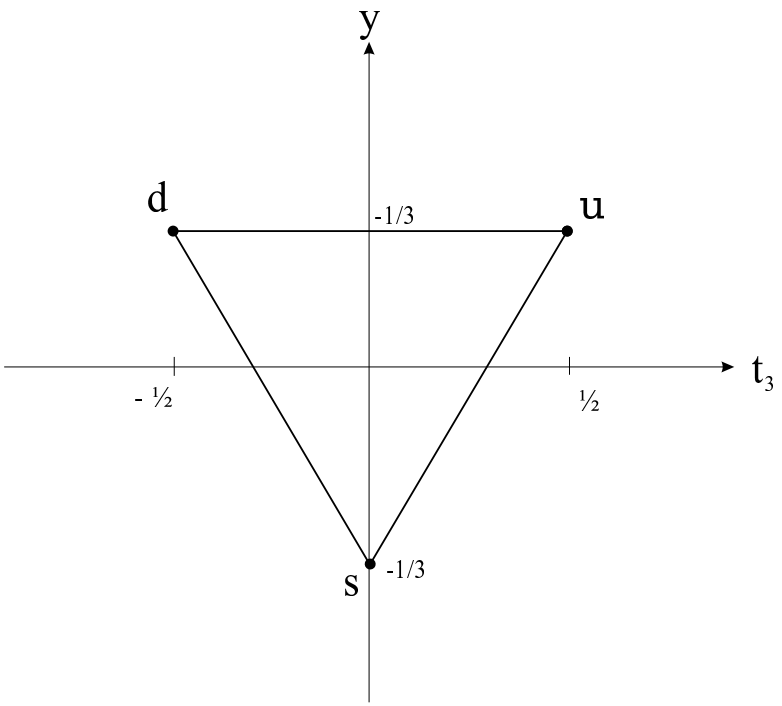


Figure 12.3: $D(1,0)$ or the fundamental representation of flavor $SU(3)$ symmetry.

The three quarks in the fundamental representation can now be written as

$$u = \left|\frac{1}{2}, \frac{1}{3}\right\rangle, \tag{12.115}$$

$$d = \left|-\frac{1}{2}, \frac{1}{3}\right\rangle, \tag{12.116}$$

$$s = \left|0, -\frac{2}{3}\right\rangle. \tag{12.117}$$

The Gell-Mann–Nishijima relation can then be succinctly expressed as

$$Q = \frac{1}{2}y + t_3, \tag{12.118}$$

from which the quark charges follow

$$Q_u = \frac{2}{3}, \tag{12.119}$$

$$Q_d = -\frac{1}{3}, \tag{12.120}$$

$$Q_s = -\frac{1}{3}. \tag{12.121}$$

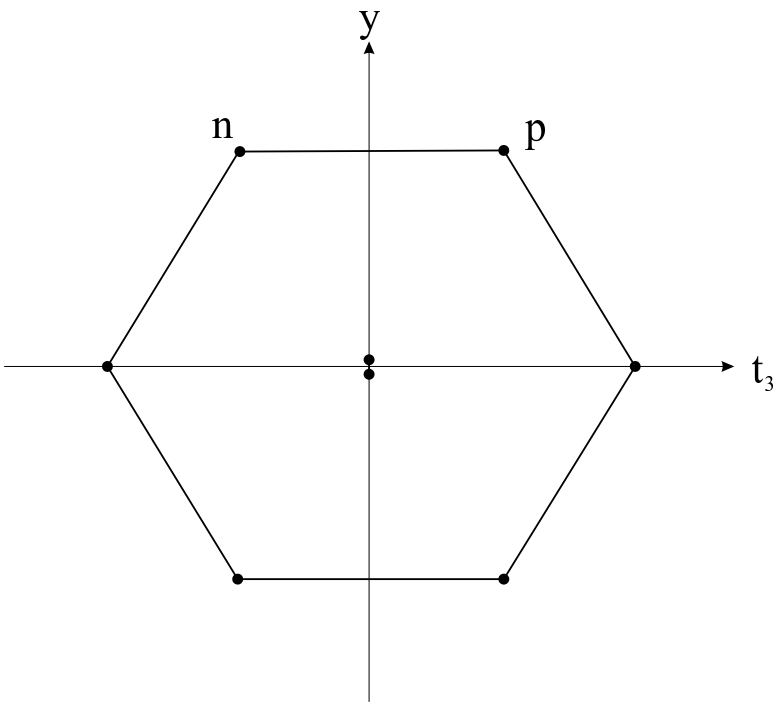


Figure 12.4: The baryon octet as a $D(1,1)$ representation of $SU(3)$.

References

[1] *Quarks and leptons: an introductory course in particle physics*, F. Halzen, A. D. Martin, 1984, Wiley.

[2] *Introduction to elementary particles*, D. Griffiths, 1987, Wiley.

[3] *Microscopic theory of the nucleus*, J. M. Eisenberg, W. Greiner, 1972, North holland.

[4] *Quantum mechanics: symmetries*, W. Greiner, B. Müller, 1989, Springer-Verlag

[5] *Principles of symmetry, dynamics, and spectroscopy*, W. G. Harter, 1993, Wiley.

[6] *Gauge theory of elementary particle physics*, T.-P. Cheng, L.-F. Li, 1984, Oxford.

[7] *Expansion method for stationary states of quantum billiards*, D. L. Kaufman, I. Kosztin, K. Schulten, *Am. J. Phys.* **67** (2), p. 133, (1999).