Günter Ludyk

Quantum Mechanics in Matrix Form

Springer

Quantum Mechanics in Matrix Form

Günter Ludyk

Quantum Mechanics in Matrix Form

Springer

Günter Ludyk
Physics and Electrical Engineering
University of Bremen
Bremen
Germany

ISBN 978-3-319-26364-9 ISBN 978-3-319-26366-3 (eBook)
https://doi.org/10.1007/978-3-319-26366-3

Library of Congress Control Number: 2017954872

© Springer International Publishing AG 2018
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

In memory of my parents

Martha and Paul Ludyk

Preface and Introduction

Though this be madness, yet there is method in't.

Hamlet, William Shakespeare

Matrix mechanics was introduced in 1925 by the German physicist Werner Heisenberg[1] [13]. However, the American Nobel laureate Steven Weinberg[2] writes about this publication [26]:

> If the reader is mystified at what Heisenberg was doing, he or she is not alone. I have tried several times to read the paper that Heisenberg wrote on returning from Helgoland, and, although I think I understand quantum mechanics, I have never understood Heisenberg's motivations for the mathematical steps in his paper.

For years, Heisenberg and his colleagues had been struggling with a problem that had been raised in 1913 by Niels Bohr[3] in his atomic theory: why do electrons in atoms occupy only certain permitted orbits with certain well-defined energies? Heisenberg took a completely new approach to this question. Because the trajectory of an electron in an atom cannot be directly observed, he decided to deal only with measurable quantities (namely the allowed energies for the quantum states of all electrons in the atom, as well as the frequencies with which an atom spontaneously jumps from one of these quantum states to a different state while emitting a light particle, i.e., a photon). Heisenberg introduced a sort of "table" of these frequencies. He performed mathematical operations on it, which led to new tables for the various physical quantities such as position, velocity, or the square of the velocity of an electron.

To be more precise, the table entries were the so-called transition amplitudes, that is, quantities whose squares specify a transition probability. When returning

---

[1] Werner Heisenberg, 1901–1976, German physicist, Nobel Prize 1932.

[2] Steven Weinberg, * 1933, American physicist, Nobel Prize 1979.

[3] Niels Bohr, 1885–1962, Danish physicist, Nobel Prize 1922.

from Helgoland (where he first had this crucial idea) to Göttingen, Heisenberg found out that the operations he applied to these tables were well known to mathematicians. The tables were called *matrices,* and the operations that he used to get from the table representing the electron velocity to the table representing the square was named *matrix multiplication.* Starting from the known dependence between the energy of a particle and its velocity and position in a simple system, Heisenberg could calculate a table of the system's energies in its different quantum states, similar to Newton's[4] calculation of the energy of a planet based on its position and velocity.

At the time, Heisenberg was constantly in touch with some influential theoretical physicists, including the German researchers Max Born[5] and Pascual Jordan[6] and the English physicist Paul Dirac.[7] Until the end of the year 1925, they transformed Heisenberg's ideas into a comprehensive and systematic version of quantum mechanics, which today we refer to as *matrix mechanics.* With the help of the new matrix mechanics, the German physicist Wolfgang Pauli[8] managed in the following January to solve the paradigmatic problem in atomic physics, namely the calculation of the energy levels of the hydrogen atom. His calculations proved the earlier ad hoc results of Bohr.

H.S. Green, an employee of Max Born, writes in [12]:

> Most books on quantum theory emphasize the wave mechanical approach (of Schrödinger[9]), probably because it is supposed to be easier to understand for those who already have a solid knowledge on differential equations.

In this book, however, we restrict ourselves to the algebraic method using matrices and only briefly describe Schrödinger's wave mechanics in order to show the equivalence with Heisenberg's matrix mechanics. By implementing numerical algorithms in standard software such as MAPLE or MATHEMATICA, matrices and matrix equations are easy to handle these days [23].

I would like to thank Dr. Claus Ascheron from Springer for his kind assistance during the compilation of this manuscript. Last but not least, I would like to express my gratitude towards my wife Renate, without whom this book would have never been published.

Bremen, Germany Günter Ludyk

---

[4] Isaac Newton, 1642–1727.

[5] Max Born, 1882–1970, German physicist, Nobel Prize 1954.

[6] Pascual Jordan, 1902–1980, German physicist.

[7] Paul Dirac, 1902–1984, English physicist, Nobel Prize 1933.

[8] Wolfgang Pauli, 1900–1958, German physicist, Nobel Prize 1945.

[9] Erwin Schrödinger, 1887–1961, Austrian physicist, Nobel Prize 1933.

Notations

Important definitions, facts, and **theorems** are shown in boxes. Important **intermediate results** are double underlined.

Scalars are written in normal font:

$$a, b, c, \alpha, \beta, \gamma, \ \ldots$$

Vectors are written as lowercase letters in bold font:
$\boldsymbol{x}, \boldsymbol{p}, \boldsymbol{v}, \ \ldots$
Vectors in four-dimensional spacetime ($\mathbb{R}^4$) are written as bold lowercase letters with an arrow:

$$\vec{\boldsymbol{x}}, \vec{v}, \vec{\boldsymbol{u}}, \ldots$$

Matrices are written as uppercase letters in bold font:

$$\boldsymbol{X}, \boldsymbol{P}, \boldsymbol{R}, \boldsymbol{I}, \ldots$$

Matrix vectors are written as uppercase letters in bold Fraktur font:

$$\mathfrak{R}, \mathfrak{X}, \mathfrak{B}, \ \ldots$$

Matrix vectors are block matrices:

$$\mathfrak{R} \overset{\text{def}}{=} \begin{pmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \\ \boldsymbol{X}_3 \end{pmatrix}.$$

The identity matrix $\boldsymbol{I_n}$ of size n is an n-by-n matrix in which all the elements on the main diagonal are equal to 1 and all other elements are equal to 0:

$$\mathbf{I}_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

If the components are real numbers ($\mathbb{R}$), the **transpose** of a vector $\boldsymbol{x}$ or a matrix $\boldsymbol{A}$ is written as x^{T} and $\boldsymbol{A}^{\mathrm{T}}$, respectively. If the components are complex numbers ($\mathbb{C}$), the transpose is written as $\boldsymbol{x}^{\dagger}$ and $\boldsymbol{A}^{\dagger}$, respectively. In this case, the components of the transpose row vector are the complex conjugate. Note that this is necessary in order for the scalar product $\boldsymbol{x}^{\dagger}\boldsymbol{x}$ to be a real number. The same reasoning applies to the transpose matrix elements.

Contents

Chapter 1
Quantum Theory Prior to 1925

Abstract A description is given of the "Older Quantum Mechanics" as introduced by Nils Bohr and Arnold Sommerfeld in the years before 1925. Bohr's postulates are formulated and the atom size is derived from them.

1.1 Bohr–Sommerfeld Quantization Rule

Before the behavior of atoms could be described with the help of quantum mechanics, Bohr and Sommerfeld[1] explained the spectra of simple atoms based on Bohr's atom model ("Older Quantum Mechanics"). In this model, spectral lines appear as energy differences between two "discrete" electron orbits. The Bohr–Sommerfeld quantization rule requires that not only the equation of motion holds for the orbit of the electron around the nucleus, but that for each additional circulation

$$\int p\,\mathrm{d}x = nh \qquad (n = 1, 2, \dots) \tag{1.1}$$

applies, where p is the momentum, and the position x runs through one complete circle. As does the angular momentum, this integral has the dimension location times moment and is an action. The constant

$$h = 2\pi \cdot \hbar = 2\pi \cdot 1.054572 \cdot 10^{-34} Js = 6.6260755 \cdot 10^{-34} Js$$

is the *Planck quantum of action* or *Planck's constant*.[2] In other words, the action of each stationary electron orbit in the atom is quantized. The action can only occur as an integer multiple of Planck's constant. A detailed derivation of this principle by Sommerfeld [22] is presented in the following section.

---

[1] Arnold Sommerfeld, 1868–1951, German physicist.

[2] Max Planck, 1858–1947, German physicist, Nobel Prize 1919.

G. Ludyk, *Quantum Mechanics in Matrix Form*,
https://doi.org/10.1007/978-3-319-26366-3_1

1.2 Sommerfeld's Derivation

Let us consider an arbitrarily moving point mass with momentum

$$\boldsymbol{p} = m\boldsymbol{v}. \tag{1.2}$$

With $\boldsymbol{v} = \dot{\boldsymbol{q}}$ we get

$$\boldsymbol{p} = m\dot{\boldsymbol{q}}. \tag{1.3}$$

It is important to note that the dynamic triple of the momentum coordinates $\boldsymbol{p}$ is considered in addition to the geometric triple of the position coordinates $\boldsymbol{q}$. Newton supplies

$$\dot{\boldsymbol{p}} = \boldsymbol{f} = -\frac{\partial E_{pot}}{\partial \boldsymbol{q}}, \tag{1.4}$$

where it was assumed that the force $\boldsymbol{f}$ can be derived from a potential energy E_{pot} (which is a function of the position coordinates $\boldsymbol{q}$). Using (1.3), the kinetic energy is obtained as

$$E_{kin} = \frac{m}{2}\dot{\boldsymbol{q}}^{\mathsf{T}}\dot{\boldsymbol{q}} = \frac{1}{2m}\boldsymbol{p}^{\mathsf{T}}\boldsymbol{p}.$$

The total energy as a function of the q_k and p_k is also called the Hamilton function or Hamiltonian H. We hence have

$$H(\boldsymbol{q}, \boldsymbol{p}) = E_{kin} + E_{pot}, \quad \frac{\partial H}{\partial \boldsymbol{q}} = \frac{\partial E_{pot}}{\partial \boldsymbol{q}}, \quad \frac{\partial H}{\partial \boldsymbol{p}} = \frac{\partial E_{kin}}{\partial \boldsymbol{p}} = \frac{\boldsymbol{p}}{m}.$$

As a result, the basic (1.3) and (1.25) can be written as

$$\boxed{\frac{d\boldsymbol{q}}{dt} = \frac{\partial H}{\partial \boldsymbol{p}}, \quad \frac{d\boldsymbol{p}}{dt} = -\frac{\partial H}{\partial \boldsymbol{q}}}. \tag{1.5}$$

Assuming that the kinetic energy is expressed as a function of $\boldsymbol{q}$ and $\dot{\boldsymbol{q}}$, the equation

$$\boldsymbol{p} = \frac{\partial E_{kin}}{\partial \dot{\boldsymbol{q}}} \tag{1.6}$$

is identical to (1.3).

The values of the coordinates $\boldsymbol{q}$ and $\boldsymbol{p}$ determine the state of the system. To understand how the state of motion of the system depends on the location $\boldsymbol{q}$ and the speed/momentum $\boldsymbol{p}$, let us imagine a single mass point with three degrees of freedom. Its three location coordinates and its three momentum coordinates can be plotted as Cartesian coordinates in a state space of six dimensions, such that each point of this state space represents a potential state of our mass point. Accordingly, for a system with d degrees of freedom one gets a state space with $2d$ dimensions.

In a first step, we can restrict ourselves to systems with only one degree of freedom. In this case, the general state space is a simple (two-dimensional) state plane. In this state plane, we can introduce q and p as orthogonal coordinates. Afterwards, we can construct *state trajectories* in this state plane, that is, the sequence of points corresponding to the successive states of motion of the system. Taking each point as the initial condition, we could draw such state trajectories and cover the entire state plane with them, whereby the trajectories are arbitrarily close to each other. It is characteristic for quantum theory, however, that a discrete subset of state trajectories is singled out from the infinite amount of potential trajectories. For defining this subset, we first consider a surface area of the state plane that is bounded by any two state trajectories, and we associate this area with the corresponding state. Afterwards we construct our bundle such that the area of the state between two neighboring curves is always equal to Planck's constant h. In other words, h has the meaning of an elementary area of the state. Let us from now on consider this meaning the true definition of Planck's quantum h.

We illustrate these abstract ideas with the help of two important special cases, namely the oscillator and the rotator. The linear *oscillator* is a spring attached to a point mass m at rest position, whereby the mass can only move in one direction $x = q$ from its rest position. Due to the spring, the mass experiences a restoring force, but no damping. The oscillator is the simplest model of a vibration center, as it is used in optics to describe a "quasi-elastically bounded electron." The oscillator is specified as a *harmonic oscillator* if we want to emphasize that only a certain eigen-oscillation is allowed. The oscillation frequency of the oscillator (i.e., the number of free vibrations in one time unit) is ν. The oscillation process then looks like

$$x = q = a \sin 2\pi\nu t, \tag{1.7}$$

where a is the amplitude of the oscillation. In this case, the momentum p is simply $m\dot{q}$; that is,

$$p = 2\pi\,\nu\, m\, a \cos 2\pi\nu t. \tag{1.8}$$

Eliminating t from (1.7) and (1.8), we obtain an ellipse as the state trajectory in the p, q-plane:

$$\frac{q^2}{a^2} + \frac{p^2}{b^2} = 1, \tag{1.9}$$

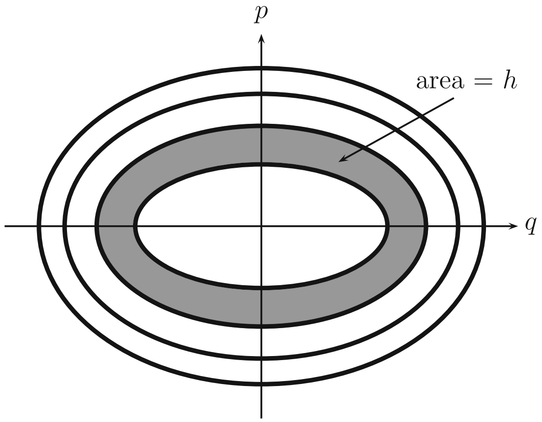
where the semi-minor axis b has the value

$$b = 2\pi\,\nu\, m\, a. \tag{1.10}$$

The area of the ellipse is

$$a\, b\, \pi = 2\pi^2 \nu\, m\, a^2.$$

Fig. 1.1 State plane of the linear oscillator



Note that this entity equals E/ν, where E is the constant vibrational energy. To see this, let us calculate E at the time $t = 0$, where the potential energy is equal to 0 and the kinetic energy is

$$\frac{m}{2}a^2(2\pi\nu)^2 = E. \tag{1.11}$$

Therefore in fact

$$a\,b\,\pi = \frac{E}{\nu}. \tag{1.12}$$

Figure 1.1 shows a family of similar ellipses in the state plane (p, q) that is obtained by changing E accordingly. After all, (1.10) ensures that the ratio b/a has the constant value $2\pi\,\nu\,m$. In the image, the ellipses follow one another in such a way that the resulting ellipse rings have the same area h. If we denote the difference of the energy constants for two consecutive ellipses with ΔE, we obtain from (1.12)

$$h = \frac{\Delta E}{\nu}, \quad \Delta E = h\,\nu. \tag{1.13}$$

By numbering the ellipses as $0, 1, \ldots, n$ and calling the corresponding energies $E_0, E_1, \ldots, E_n$, it follows from (1.13)

$$E_n = E_0 + h\,\nu\,n. \tag{1.14}$$

In the classical theory, all points of the state plane are equal and represent possible states of the oscillator. In quantum theory, however, those states are distinguished whose points lie on one of those ellipses. Such states are called the stationary states of the oscillator, because they can be continuously passed through without energy loss of the oscillator. In a way, they represent the charged mass point without radiation.

From time to time, however, the oscillator changes its energy. In particular, it emits energy once its state moves onto a smaller ellipse. Similarly, it absorbs energy once the state jumps to a larger ellipse. Note, however, that both emission and absorption occur in multiples of the energy quantum $h\nu$.

Let us generalize this insight to any mechanical system with one degree of freedom. *The image point of the system in the state plane is restricted to certain quantum theoretically special "quantized" state trajectories, each of which includes an elementary region of size h with the following trajectory. The nth trajectory includes (when closed) the area*

$$J = \iint \mathrm{d}p\,\mathrm{d}q, \tag{1.15}$$

where the integral covers the inside of the nthtrajectory. Carrying out the integration with respect to p yields

$$J = \int p\,\mathrm{d}q, \tag{1.16}$$

where the integral is bound *by the nth trajectory* itself. We call (1.16) the *state integral* or *action.*

The final formulation of the quantum hypothesis is obtained by requiring that the difference of the state integrals for two consecutive trajectories must be equal to h:

$$\Delta J = h, \quad J = J_0 + n\,h. \tag{1.17}$$

This requirement selects a discrete (infinite) number of real, quantum theoretically possible trajectories out of the *continuous* manifold of all possible mechanical movements. Unlike this general version of quantum hypothesis, the original energy quantum hypothesis as formulated by Planck with respect to thermal radiation is only valid for the oscillator. The above evaluation of the state integral (1.15) was very easy because we could directly calculate the area of the ellipses by the formula $a\,b\,\pi$. In particular, (1.17) yields (assuming $J_0 = 0$) a formulation that is analogous to (1.14):

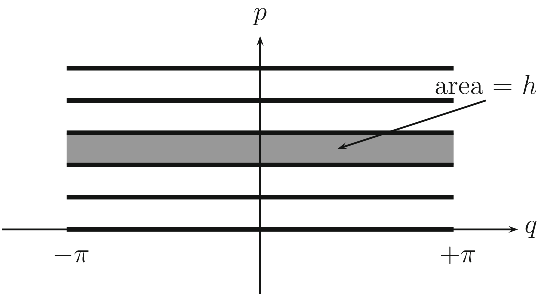
$$J = \int p\,\mathrm{d}q = n\,h. \tag{1.18}$$

Now let us move on to the *rotator*. A rotator is a point mass m that rotates uniformly on a circle of radius a around a fixed center. We can use the rotation angle $\varphi = q$ as a natural location coordinate. Starting from an arbitrary initial position $\varphi = 0$, the particle position on the circle is labeled by the angle φ. Because the velocity of the point mass m is equal to $a\dot{q}$, the kinetic energy is given by

$$E_{kin} = \frac{m}{2}\,a^2\dot{q}^2. \tag{1.19}$$

For a uniform rotation, the potential energy is independent of φ. Also, we can assume that it does not depend on a either, inasmuch as a is constant during the motion. We

Fig. 1.2 State plane of the rotator



can therefore write

$$E_{pot} = const.$$

According to (1.6) and (1.19), the momentum corresponding to $\dot{q}$ is

$$p = m\,a^2\dot{q}. \tag{1.20}$$

Because $\dot{q} = const$, this momentum p is also constant, in accordance with the equations of motion (1.5). Therefore, the state trajectory of the rotator in the state plane (p, q) is a line that is parallel to the q-axis; see Fig. 1.2.

Note that the state trajectory is not a closed curve in this case. Therefore we must first define what we understand as the area of the state trajectory. We remark that the state of the rotator (i.e., its location on the trajectory as well as the direction of its momentum) is repeated after each loop. The true state trajectory is therefore not an infinite, but a finite line that is repeatedly passed through from left to right. In the q-direction, the state area of the rotator has only the length 2π. Casually speaking, one can cut out the line $q = \pm\pi$ and glue it together to form a cylinder. The area of the cylinder between the nth and the $(n-1)$th state trajectory is then given (similar to a rectangle with the baseline 2π) by $2\pi(p_n - p_{n-1})$. This area must now be equal to h. For the area between the nth and the zero state trajectory (the latter being represented by the q-axis), we then get

$$2\,\pi\,p_n = n\,h. \tag{1.21}$$

This area plays the same role as the closed curves in the case of the oscillator.

As we can see, the rotator is *not quantized in energy quanta* (as the oscillator), but according to its angular momentum. More precisely, the angular momentum of a rotator is an integral multiple of $h/2\pi$. We can now calculate the (kinetic) energy of the rotator based on (1.19) and (1.20) as

$$E_{kin} = \frac{p\,\dot{q}}{2}.$$

From (1.21), with $\nu = \dot{q}/2\pi$ we further get:

$$E_{kin} = \frac{n\,h}{2}\frac{\dot{q}}{2\pi} = \frac{n\,h\,\nu}{2}. \tag{1.22}$$

Here, ν is the winding number of the rotator, that is, the number of full rotations in one time unit. This number is the analogue of the oscillation frequency of the oscillator.

1.3 Bohr's Postulates

Niels Bohr derived his model by adding three postulates to Rutherford's model, namely:

1. Electrons move on stable orbits around the nucleus. Contrary to the prediction of the theory of electrodynamics, the electron radiates no energy (in the form of electromagnetic radiation) when moving on the orbits.
2. The radius of the electron orbit does not change continuously, but abruptly. When the electron jumps from one orbit to another, electromagnetic radiation is emitted (or absorbed). The frequency of the radiation is given by Max Planck's relationship between energy and the frequency of light. According to this relationship, if E_n is the energy of the initial state and E_m is the energy of the target state, then a photon is emitted with frequency

$$\nu = \frac{(E_m - E_n)}{h}. \tag{1.23}$$

3. Electron orbits are only stable if the orbital angular momentum L of the electron is an integral multiple of the reduced Planck constant $\hbar = \frac{h}{2\pi}$: $L = n\hbar$. This postulate is often called the *selection condition*.

In addition, the classical equation of motion applies, that is, a centripetal Coulomb force acts.

1.4 Atom Sizes

Bohr's atom model allows a comparison of a series of numerical predictions with experimental results, especially with respect to the position of the lines in the hydrogen spectrum. The model treats the electron as a pointlike particle that is attracted by the opposite electric charge of the nucleus. This force shapes the path of the electron according to the laws of classical mechanics on circular orbits. That is why in Bohr's atomic model the distance of an electron to the nucleus is also referred to as the *classical atomic radius*. The angular momentum L of a particle with mass m and velocity v along a circular orbit of radius r is

$$L = mvr.$$

A centripetal force acts on the moving electron according to

$$F_{\text{centr}} = \frac{mv^2}{r}.$$

On the other hand, the electron with the elementary charge e experiences a force in the electric field of the proton according to Coulomb's law:

$$F_{\text{el}} = \frac{e^2}{4\pi\varepsilon_0 r^2}.$$

This Coulomb force keeps the electron on the circular orbit; that is, the two forces must be equal:

$$F_{\text{el}} = F_{\text{centr}} \quad \Leftrightarrow \quad \frac{e^2}{4\pi\varepsilon_0 r^2} = \frac{m_e v^2}{r}. \tag{1.24}$$

Hereby, the angular momentum must satisfy the above-postulated selection condition. A more sustainable formulation of this condition is that the length of the orbit ($2\pi r$) must be an integer multiple of the electron's elementary wavelength (or *de Broglie wavelength*), because otherwise destructive interference would occur. Along a permitted orbit, the electron forms a standing matter wave:

$$\lambda_{\text{dB}} = \frac{h}{p} = \frac{h}{m_e v} \quad \Rightarrow \quad 2\pi r = n\frac{h}{m_e v}.$$

Solving for v yields

$$v = \frac{n\hbar}{m_e r}.$$

This velocity v can be substituted in the above equation for the forces:

$$\frac{e^2}{4\pi\varepsilon_0 r^2} = \frac{m_e \left(\frac{n\hbar}{m_e r}\right)^2}{r} \quad \Leftrightarrow \quad \frac{e^2}{4\pi\varepsilon_0 r^2} = \frac{m_e n^2 \hbar^2}{m_e^2 r^3} \quad \Leftrightarrow$$

$$r_n = n^2 \frac{\varepsilon_0 \cdot h^2}{m_e e^2 \pi}. \tag{1.25}$$

The smallest radius ($n = 1$) is denoted as *Bohr's atomic radius*:

$$r_1 = \frac{\varepsilon_0 h^2}{m_e e^2 \pi} \approx 5.29 \cdot 10^{-11}\text{m}. \tag{1.26}$$

For each higher electron orbit, we get

$$r_n = n^2 \cdot r_1 \approx n^2 \cdot 5.29 \cdot 10^{-11}\text{m}. \tag{1.27}$$

The energy E_n of the electron on the nth orbit consists of kinetic as well as potential energy, whereby the potential energy is negative:

$$E_n = E_{kin} + E_{pot} = \frac{m_e \cdot v_n^2}{2} - \frac{1}{4\pi\varepsilon_0} \cdot \frac{e^2}{r_n}. \tag{1.28}$$

From (1.24), it follows that

$$v_n^2 = \frac{e^2}{4\pi\varepsilon_0 \cdot r_n \cdot m_e}.$$

We substitute this term in (1.28), which provides

$$E_n = -\frac{1}{2}\frac{1}{4\pi\varepsilon_0}\frac{e^2}{r_n}.$$

With (1.26) and (1.27), we get

$$\boxed{E_n = -\frac{1}{2}\frac{1}{4\pi\varepsilon_0}\frac{e^2}{n^2 r_1} = -\frac{m_e e^4}{8\varepsilon_0^2 h^2}\frac{1}{n^2} = -13.6\,eV \cdot \frac{1}{n^2}}. \tag{1.29}$$

In a hydrogen atom, an electron can only assume these discrete energy values, where the integer n denotes its current trajectory.

Chapter 2
Heisenberg's Year 1925

Abstract Starting from the known facts on spectral lines up to 1925, the crucial new ideas of Heisenberg are presented which led him to the introduction of his matrix quantum mechanics.

2.1 Spectral Lines

Nineteenth century research on spectral lines can be considered the starting point for quantum theory. These investigations are based on two measurable variables, namely the frequency and the brightness of the spectral lines. After all, the emitted light consists of very sharp frequencies. In the year 1850, Kirchhoff[1] and Bunsen[2] had discovered that chemical elements produce such characteristic lines in their spectra.

In the visible region of the *hydrogen spectrum*, four lines can be observed (the wavelengths get closer as they decrease). In 1885, the Swiss mathematician Balmer[3] (a teacher at a lyceum) discovered that the wavelength λ of these lines can be calculated with the simple formula

$$\lambda = A\left(\frac{n^2}{n^2-4}\right),$$

where $A = 364.56 \times 10^{-9}$ m and $n = 3, 4, 5$, or 6. This leads to the following wavelengths (in nanometers):

- 656.279 nm (red)
- 486.133 nm (blue-green)
- 434.047 nm (violet)
- 410.174 nm (violet)

---

[1] Gustav Robert Kirchhoff, 1824–1887, German physicist.

[2] Robert Wilhelm Eberhard Bunsen, 1811–1899, German chemist.

[3] Johann Jakob Balmer, 1825–1898.

G. Ludyk, *Quantum Mechanics in Matrix Form*,
https://doi.org/10.1007/978-3-319-26366-3_2

Table 2.1 Measured or calculated hydrogen spectral lines in the Balmer series

Transition $n \to m = 2$	$3 \to 2$	$4 \to 2$	$5 \to 2$	$6 \to 2$	$7 \to 2$	$8 \to 2$
Name	H_α	H_β	H_γ	H_δ	H_ϵ	H_ζ
Measured (nm)	656.2793	486.1327	434.0466	410.1738	397.0075	388.8052
Calculated	656.278	486.132	434.045	410.1735	397.0074	388.8057
Colour	Red	Blue-Green	Violet	Violet	Violet	Violet

For the wavelengths λ, we obtain the *Balmer formula*

$$\frac{1}{\lambda} = R\left(\frac{1}{4} - \frac{1}{n^2}\right),$$

where $R = 4/A$ is the so-called *Rydberg*[4] *constant*.[5] With explicit numbers, it reads

$$\frac{1}{\lambda} = 10973731\left(\frac{1}{4} - \frac{1}{n^2}\right)\ [m^{-1}].$$

Table 2.1 is a table of the currently known spectral lines in the Balmer series, measured or calculated with the Balmer formula.

Five years later, in 1890, Rydberg generalized the Balmer equation into the Rydberg formula

$$\frac{1}{\lambda} = R\left(\frac{1}{m^2} - \frac{1}{n^2}\right)$$

with $m = 1, 2, \ldots$ and $n = m + 1, m + 2, \ldots$. For $m = 1$, one obtains the Lyman series. For $m = 2$, we get again Balmer's series, and $m = 3$ yields the Paschen series. For $m = 4$, the Brackett series is obtained, and $m = 5$ leads to the Pfund series. Each series is named after its discoverer; see Table 2.2 and Fig. 2.1.

For the spectra of other elements, different Rydberg constants are obtained. The Swiss physicist Ritz[6] discovered in 1908 that one can derive new lines from the known spectral lines of an element without having to modify any constants. Here is why: from the above formulas it becomes obvious that the resulting frequencies ν depend on two integers, namely m and n, in other words,

$$\nu(m, n) = Rc\left(\frac{1}{m^2} - \frac{1}{n^2}\right). \tag{2.1}$$

---

[4] Johannes Robert Rydberg, 1854–1919, Swedish physicist.

[5] Today, R is an accurately known fundamental constant with $R_\infty = 10973731.568539(55)\, m^{-1}$. The index ∞ indicates that an infinitely large nucleus mass is assumed.

[6] Walter Ritz, 1878–1909, Swiss theoretical physicist.

Table 2.2 Hydrogen spectral lines series

Name	n_1	n_2	Formula	Spectrum range/Colour
LYMAN-Series	1	2, 3, 4, ...	$\widetilde{\nu} = R\left(1 - \frac{1}{n_2^2}\right)$	Vacuum-UV (121 nm → 91 nm)
BALMER-Series	2	3, 4, 5, ...	$\widetilde{\nu} = R\left(\frac{1}{2^2} - \frac{1}{n_2^2}\right)$	red, blue-green, 4× violet, then transition to near-UV → 365 nm
PASCHEN-Series	3	4, 5, 6, ...	$\widetilde{\nu} = R\left(\frac{1}{3^2} - \frac{1}{n_2^2}\right)$	IR-A (1875 nm → 820 nm)
BRACKETT-Series	4	5, 6, 7, ...	$\widetilde{\nu} = R\left(\frac{1}{4^2} - \frac{1}{n_2^2}\right)$	IR-B (4050 nm → 1460 nm)
PFUND-Series	5	6, 7, 8, ...	$\widetilde{\nu} = R\left(\frac{1}{5^2} - \frac{1}{n_2^2}\right)$	IR-B (7457 nm → 2280 nm)

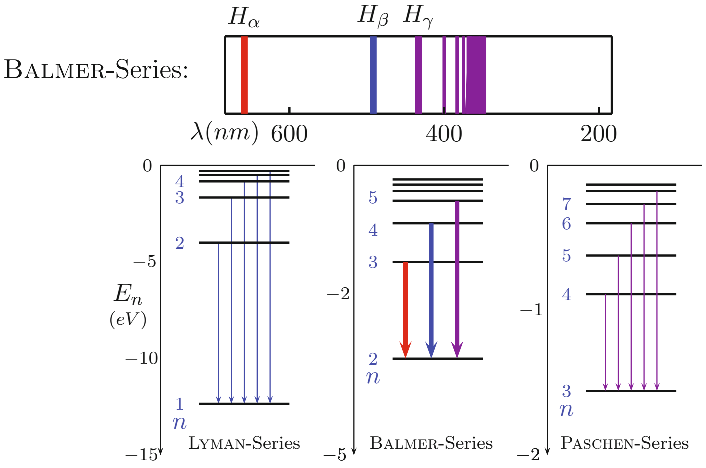


Fig. 2.1 The visible part of the hydrogen spectrum of the Balmer series (above). Energy levels and transitions (below)

Adding two different frequencies $\nu(m_1, n_1)$ and $\nu(m_2, n_2)$ of a spectrum, one obtains

$$\nu(m_1, n_1) + \nu(m_2, n_2) = Rc\left(\frac{1}{m_1^2} - \frac{1}{n_1^2} + \frac{1}{m_2^2} - \frac{1}{n_2^2}\right). \tag{2.2}$$

For $n_1 = m_2$ we obtain the new frequency according to the *Ritz combination principle*:

$$\underline{\underline{\nu(m_1, n_2) = \nu(m_1, n_1) + \nu(n_1, n_2)}} = R \cdot c\left(\frac{1}{m_1^2} - \frac{1}{n_2^2}\right). \tag{2.3}$$

And then the time of Niels Bohr came! It can be safely assumed that he was inspired by the Rydberg formula

$$\nu(m,n) = R \cdot c \left(\frac{1}{m^2} - \frac{1}{n^2} \right).$$

We can write this formula without brackets:

$$\nu(m,n) = R \cdot c \frac{1}{m^2} - R \cdot c \frac{1}{n^2}. \tag{2.4}$$

On the other hand, we have

$$E = h \cdot \nu,$$

and therefore

$$\nu = \frac{E}{h}.$$

It seems very plausible to rewrite for (2.4)

$$\nu(m,n) = \frac{E_m}{h} - \frac{E_n}{h}. \tag{2.5}$$

From this follows

$$E_k = \frac{h \cdot R \cdot c}{k^2}.$$

Note that the dimensions in the formula are consistent: the action quantum h has dimension Js, $R \cdot c$ has dimension s^{-1} and k is a dimensionless integer, so E_k has indeed the dimension J.

We need to emphasize once again that in quantum theory the mechanical behavior of an atom is characterized by two basic parameters, namely the energy E_n of the stationary state n and the probability per unit time $A(n,m)$ for the spontaneous transition from state n to the state m. In a spectroscopic study, we measure the radiation as emitted by the atom, that is, the line spectrum. Quantum theory then puts the mechanical properties (namely E_n and $A(n,m)$) into relation to the spectral characteristics (namely the frequency and intensity of the emitted light). Hereby, the transition energy $E_n - E_m$ determines the frequency of the light, and the transition probability $A(n,m)$ determines its intensity.

The emission of radiation by an atom is the result of electrons jumping between two discrete electron orbits. The transition probability determines the occurrence of a quantum jump. The radiation emitted during the transition $n \rightarrow m$ then has the frequency $\nu(n,m)$. Therefore, a mechanical energy loss $E_m - E_n$ occurs during the transition $n \to m$. Because[7] the photon energy is $h\,\nu(n,m) = \hbar\,\omega(n,m)$, energy conservation requires $E_m = E_n + \hbar\,\omega(n,m)$; that is,

$$E_m - E_n = \hbar\,\omega(n,m). \tag{2.6}$$

Note that a single photon only generates a flash of light on a spectral line. The complete spectral line is produced only if many atoms emit many photons. The number of photons with frequency $\nu(n,m)$ that arrives at a given area of the spectrometer per unit time determines the intensity of the line. In other words, the line intensity is a function of the number of jumping electrons, that is, the transition rate $A(n,m)$.

Let us consider a set of atoms, each in the state n. Then the light intensity $P(n,m)$ of the transition $n \to m$ defined by the amount of energy emitted per unit time Δt and per atom is

$$P(n,m) \overset{\text{def}}{=} \frac{1}{N_n}\frac{\Delta E(n,m)}{\Delta t}, \tag{2.7}$$

where N_n is the number of atoms in the state n, and $\Delta E(n,m)$ is the energy consumption of all atoms accomplishing the transition $n \to m$ in the time interval Δt. Again, the conservation of energy requires

$$\Delta E(n,m) = \Delta N(n,m)\hbar\,\omega(n,m),$$

where $\Delta N(n,m)$ is the number of atoms jumping from n to m in the time interval Δt. For large N_n, the portion of jumping atoms equals the probability for an atom to jump, namely

$$\Delta N(n,m)/N_n = A(n,m)\Delta t.$$

$A(n,m)$ is therefore the probability per unit time. From (2.7) we get

$$P(n,m) = \frac{\Delta N(n,m)\hbar\,\omega(n,m)}{N_n \Delta t} = A(n,m)\hbar\,\omega(n,m). \tag{2.8}$$

This result implies that intense lines are very probable transitions, and weak lines are improbable transitions.

2.2 Introduction of Matrices

Werner Heisenberg started off from the principle that concepts which do not correspond to *physically observable facts* should not be used in a theoretical description. Heisenberg therefore banished the idea of electron orbits with fixed radii and orbital

---

[7] $\hbar \overset{\text{def}}{=} \frac{h}{2\pi}$.

periods, because these quantities cannot be observed. Instead, he postulated that the theory was to be constructed with the help of abstract quadratic schemes. Rather than describing the motion by a time-dependent coordinate $x(t)$, he suggested determining a scheme of transition amplitudes x_{mn}. Heisenberg's theory is entirely based on measurable quantities, namely the frequencies and strengths of the spectral lines of atoms.

From Rydberg's formula it follows that each frequency $\nu(n, m)$ of the observed spectrum can be written as the difference of two energy terms E_n and E_m:

$$h\nu(n, m) = E_n - E_m. \tag{2.9}$$

This immediately leads to the Ritz combination principle

$$\nu(n, k) + \nu(k, m) = \frac{1}{h}\left((E_n - E_k) + (E_k - E_m)\right) = \frac{1}{h}(E_n - E_m) = \nu(n, m), \tag{2.10}$$

where the νs are the observable frequencies of the spectrum. From (2.9), we also find that

$$\nu(n, m) = \frac{1}{h}(E_n - E_m) = -\frac{1}{h}(E_m - E_n) = -\nu(m, n), \tag{2.11}$$

and that

$$\nu(n, n) = \frac{1}{h}(E_n - E_n) = 0. \tag{2.12}$$

For an objective observer, the frequencies and intensities of the spectral lines are the only available data of what is happening inside an atom. Because the occurring frequencies depend on *two* terms, it makes perfect sense to arrange the frequencies in a table. The first row contains all frequencies that are generated from E_0, that is,

$$\nu(0, 0), \nu(0, 1), \nu(0, 2), \ldots.$$

The second row contains frequencies that can be generated starting from E_1, namely

$$\nu(1, 0), \nu(1, 1), \nu(1, 2), \ldots.$$

The columns always contain all frequencies (in rising order) that can be generated upon reaching a certain level. Eventually, the table can be written in the form of a matrix Ω as

$$\Omega = 2\pi \begin{pmatrix} \nu(0,0) & \nu(0,1) & \nu(0,2) & \cdots \\ \nu(1,0) & \nu(1,1) & \nu(1,2) & \cdots \\ \nu(2,0) & \nu(2,1) & \nu(2,2) & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}. \tag{2.13}$$

For a harmonically bound electron, the position $x(n, t)$, where the stationary state is labeled by n, can be written as a Fourier series

$$x(n, t) = \sum_{\alpha=-\infty}^{\infty} a_\alpha e^{i\,\omega(n)t}. \tag{2.14}$$

Heisenberg observed that the αth component of the classical motion corresponds to the quantum mechanical transition from the state n to the state $n - \alpha$. So Heisenberg replaced the classical component

$$a_\alpha e^{i\omega(n)t}$$

by

$$a(n, n - \alpha)e^{i\omega(n,n-\alpha)t}.$$

In order to account for the transition from one stationary state n to another $(n - \alpha)$, he replaced (2.14) by

$$x \to a(n, n - \alpha)e^{i\omega(n,n-\alpha)t},$$

or

$$x_{nm} = a(n, m)e^{i\omega(n,m)t}, \tag{2.15}$$

and summed over transition components as in (2.14). Heisenberg represented the position by a set of transition components, x_{mn}, and he replaced $x_\alpha(n)$ by x_{mn} and $\alpha\omega(n)$ by $\omega(n, m)$.

Additionally Heisenberg modified the old Bohr–Sommerfeld quantization rule (1.18)

$$J = \int p\,\mathrm{d}q = \int m\dot{x}^2\mathrm{d}t = h\,n,$$

integrated over a full period of the motion. If one expressed this equation in terms of the Fourier series (2.14) for $x(n, t)$ one would obtain

$$h\,n = 2\pi m \sum_{\alpha=-\infty}^{\infty} |a_\alpha(n)|^2 \alpha^2 \omega_n. \tag{2.16}$$

The presence of the integer n in (2.16) seemed for Heisenberg to be an arbitrary condition, and he concluded that this equation must be replaced by a new condition and that the new condition must be about the transition between states. By differentiation (2.16) with respect to n Heisenberg found

$$h = 2\pi m \sum_{\alpha=-\infty}^{\infty} \alpha \frac{\mathrm{d}}{\mathrm{d}n}(\alpha\omega_n|a_\alpha|^2).$$

Heisenberg replaced the derivative by a difference:

$$h = 4\pi m \sum_{\alpha=0}^{\infty} \{|a(n, n+\alpha)|^2 \omega(n, n+\alpha) - |a(n, n-\alpha)|^2 \omega(n, n-\alpha)\}. \quad (2.17)$$

This is Heisenberg's quantum condition; it relates the amplitudes of different lines within an atomic spectrum.

How is the quantity $(x(t))^2$ to be represented in quantum mechanics, which appears, for example, by modeling an anharmonic oscillator? The answer in classical theory is obviously:

$$b_\beta(n) e^{i\omega(n)\beta t} = \sum_{\alpha=-\infty}^{\infty} a_\alpha a_{\beta-\alpha} e^{i\omega(n)(\alpha+\beta-\alpha)t}, \quad (2.18)$$

so that

$$(x(t))^2 = \sum_{\beta=-\infty}^{\infty} b_\beta(n) e^{i\omega(n)\beta t}. \quad (2.19)$$

To Heisenberg it seemed that in quantum mechanics the simplest and most natural assumption would be to replace (2.18) by:

$$b(n, n-\beta) e^{i\omega(n,n-\beta)t} = \sum_{\alpha=-\infty}^{\infty} a(n, n-\alpha) a(n-\alpha, n-\beta) e^{i\omega(n,n-\beta)t}. \quad (2.20)$$

Max Born studied Heisenberg's manuscript and discovered that Heisenberg's symbolic multiplication was nothing but matrix multiplication.

By collecting all possible transitions in a matrix (similar to the frequencies ν in Ω), we obtain a matrix of the form[8]

$$\boldsymbol{X} = \left(a(n, m)\, e^{i2\pi\nu(n,m)t} \right).$$

Due to $\nu(n, m) = -\nu(m, n)$ and $a(m, n) = a^*(n, m)$, we get $x(n, m) = x^*(m, n)$; that is, the matrix $\boldsymbol{X}$ is an Hermitian[9] matrix. When transposing an Hermitian matrix $\boldsymbol{X}$, each component becomes its complex conjugate value. If we introduce a matrix multiplication as $\boldsymbol{X}\boldsymbol{X}^\dagger$, the elements of the product matrix are

$$a(n, m) a(m, n) = a(n, m) a^*(n, m) = |a(n, m)|^2. \quad (2.21)$$

---

[8] The matrix components $x(n, m) = a(n, m)\, e^{2\pi i\nu(n,m)t}$ are not to be confused with the classical coefficients $a_\alpha e^{i\alpha\omega t}$ of a Fourier series (see Appendix C), where we sum up from $\alpha = -\infty$ to $\alpha = +\infty$ in order to obtain the periodic function $x(t)$.

[9] Charles Hermite, 1822–1901, French mathematician.

Born and Jordan postulated in [5] that (2.21) is the *likelihood* for the transitions $n \rightleftharpoons m$ from the atomic state n into the state m, and vice versa.

The description is based on the idea of state transitions, therefore the diagonal elements of the matrix $\boldsymbol{X}$ must vanish; that is,

$$\nu(n, n) = 0$$

for all n. After all, no transition takes place from n to n. Altogether we get the form

$$\boldsymbol{X} = \begin{pmatrix} 0 & a(0,1)\, e^{2\pi i\nu(0,1)t} & a(0,2)\, e^{2\pi i\nu(0,2)t} & \cdots \\ a(1,0)\, e^{2\pi i\nu(1,0)t} & 0 & a(1,2)\, e^{2\pi i\nu(1,2)t} & \cdots \\ a(2,0)\, e^{2\pi i\nu(2.0)t} & a(2,1)\, e^{2\pi i\nu(2,1)t} & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{2.22}$$

Note that it is not summed over any coefficients as in a Fourier series. Rather, all the transition components that are collected in the matrix $\boldsymbol{X}$ reproduce all the possibilities for transitions of the system from the perspective of quantum theory. The aim of quantum theory is to create a mathematical model such that:

1. It allows the calculation of measurable frequency spectra of atoms (namely frequency and intensity of the spectral lines).
2. The classical theory is contained in the limit $h \to 0$, where h is Planck's quantum of action.

For the time derivatives of the elements $x(n, m)$ of the matrix $\boldsymbol{X}$ one gets

$$\dot{x}(n, m) = 2\pi i\nu(n, m)a(n, m)\, e^{2\pi i\nu(n,m)t}. \tag{2.23}$$

We introduce the diagonal matrix $\boldsymbol{E}$ with the matrix elements

$$E(n, m) \overset{\text{def}}{=} \delta_{n,m} E_n, \tag{2.24}$$

that is,

$$\boldsymbol{E} = \begin{pmatrix} E_0 & 0 & 0 & \cdots \\ 0 & E_1 & 0 & \cdots \\ 0 & 0 & E_2 & \ddots \\ 0 & 0 & 0 & \ddots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

With (2.9), we then get for (2.23)

$$\dot{x}(n, m) = 2\pi i\nu(n, m)x(n, m) = \frac{2\pi i}{h}(E_n - E_m)x(n, m). \tag{2.25}$$

Because $\boldsymbol{E}$ is diagonal, we have for matrix elements

$$E_n x(n,m) = (\boldsymbol{E}\boldsymbol{X})(n,m)$$

and

$$E_m x(n,m) = (\boldsymbol{X}\boldsymbol{E})(n,m).$$

Therefore we get

$$\dot{x}(n,m) = \frac{2\pi i}{h}\left((\boldsymbol{E}\boldsymbol{X})(n,m) - (\boldsymbol{X}\boldsymbol{E})(n,m)\right),$$

and because this applies to every element of the matrix

$$\boxed{\dot{\boldsymbol{X}} = \frac{2\pi i}{h}(\boldsymbol{E}\boldsymbol{X} - \boldsymbol{X}\boldsymbol{E})}. \tag{2.26}$$

This is the simplest form of the so-called *quantum mechanical equation of motion*, or *Heisenberg's equation of motion*. Here for the first time a *commutator* enters in the form

$$[\boldsymbol{E}\boldsymbol{X}] \overset{\text{def}}{=} (\boldsymbol{E}\boldsymbol{X} - \boldsymbol{X}\boldsymbol{E}).$$

In the theory of quantum mechanics such commutators play an important role, as we show later. For a single matrix element, (2.26) means (2.25). Dividing by $2\pi i x(n,m)$ again yields the frequency condition $h\nu(n,m) = E_n - E_m$.

2.3 Problems

2.1 **Hermitian Matrices**: Under which conditions is the product of two Hermitian matrices again an Hermitian matrix?

2.2 **Eigenvectors**: Show that eigenvectors belonging to different eigenvalues are linearly independent.

2.3 **Eigenvalues of an Hermitian Matrix**: Show that the eigenvalues of an Hermitian matrix are real.

2.4 **Eigenvalues of a Unitary Matrix**: What general property do the eigenvalues of a unitary matrix have?

2.5 **Eigenvectors**: In an N-dimensional space, N linear independent vectors $\boldsymbol{a}_j$ are given. Construct a set of N normalized orthogonal vectors $\boldsymbol{e}_j$.

2.6 **Normalized Eigenvectors**: The two linear independent vectors $\begin{pmatrix}1\\1\end{pmatrix}$ and $\begin{pmatrix}-1\\1\end{pmatrix}$ are given. Which transformation matrix $\boldsymbol{T}$ transforms these two vectors into $\begin{pmatrix}1\\0\end{pmatrix}$ and $\begin{pmatrix}0\\1\end{pmatrix}$? Which vectors would be obtained by using the method of Problem 2.5?

Chapter 3
Expansion of the Matrix Method

Abstract We describe the general ideas concerning the matrix method, which was further developed by Born, Heisenberg, and Jordan and extended to systems with several degrees of freedom.

3.1 Commutation Relation

In classical mechanics, the behavior of a system is described by canonical variables $\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{q}_1, \boldsymbol{q}_2, \ldots$, and its dynamical structure is contained in the Hamilton function $H(\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{q}_1, \boldsymbol{q}_2, \ldots)$. This scheme turned out to be very suitable for quantum theory as well. However, the canonical variables have to be replaced by more general mathematical objects, namely matrices. Because one wants to end up with the classical Hamilton theory, it makes sense to call the entries of the matrix $\boldsymbol{X}$ position coordinates and the correspondingly structured matrix $\boldsymbol{P} = m\dot{\boldsymbol{X}}$ a momentum matrix, where m is the particle mass.

Max Born and Pascual Jordan, later together with Werner Heisenberg, further expanded the matrix method. For this, we start with the two matrices $\boldsymbol{X}$ and $\boldsymbol{P}$. The matrices do not need to be commutative; that is, we do not require $\boldsymbol{X}\boldsymbol{P} = \boldsymbol{P}\boldsymbol{X}$. Therefore the difference

$$[\boldsymbol{P}, \boldsymbol{X}] \overset{\text{def}}{=} (\boldsymbol{P}\boldsymbol{X} - \boldsymbol{X}\boldsymbol{P}), \tag{3.1}$$

called the commutator of the two matrices, is normally not zero. Differentiation of the commutator with respect to the time yields

$$\frac{\mathrm{d}}{\mathrm{d}t}[\boldsymbol{P}, \boldsymbol{X}] = [\dot{\boldsymbol{P}}, \boldsymbol{X}] + [\boldsymbol{P}, \dot{\boldsymbol{X}}]. \tag{3.2}$$

Newton's law states $m\ddot{\boldsymbol{X}} = \dot{\boldsymbol{P}} = f(\boldsymbol{X})$. Substitution in (3.2) provides

$$\frac{\mathrm{d}}{\mathrm{d}t}[\boldsymbol{P}, \boldsymbol{X}] = \underbrace{[f(\boldsymbol{X}), \boldsymbol{X}]}_{f(\boldsymbol{X})\boldsymbol{X} - \boldsymbol{X}f(\boldsymbol{X})} + \underbrace{[\boldsymbol{P}, (1/m)\boldsymbol{P}]}_{(1/m)\boldsymbol{P}^2 - (1/m)\boldsymbol{P}^2} = \boldsymbol{0}. \tag{3.3}$$

G. Ludyk, *Quantum Mechanics in Matrix Form*,
https://doi.org/10.1007/978-3-319-26366-3_3

Note that $f(\boldsymbol{X})$ is composed of the matrix $\boldsymbol{X}$ and powers of the same. In this case, $f(\boldsymbol{X})$ commutes with $\boldsymbol{X}$; that is,

$$f(\boldsymbol{X})\boldsymbol{X} = \boldsymbol{X}f(\boldsymbol{X}).$$

The commutator $[\boldsymbol{P}, \boldsymbol{X}]$ *is thus a constant matrix!* What is the form of this constant matrix? Its elements consist of sums of products of the form

$$\dot{x}(\ell, j)x(j, k) = i\omega(\ell, j)x(\ell, j)x(j, k)$$

and

$$x(\ell, j)\dot{x}(j, k) = i\omega(j, k)x(\ell, j)x(j, k).$$

A general matrix element of the commutator of $\mathrm{d}\boldsymbol{X}/\mathrm{d}t$ with $\boldsymbol{X}$ is of the form

$$(\boldsymbol{X}\dot{\boldsymbol{X}} - \dot{\boldsymbol{X}}\boldsymbol{X})_{\ell k} = i\sum_j (\omega(\ell, j) - \omega(j, k))\, x(\ell, j)x(j, k). \tag{3.4}$$

The time derivative of this commutator is

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{X}\dot{\boldsymbol{X}} - \dot{\boldsymbol{X}}\boldsymbol{X})_{\ell k} &= i^2\sum_j (\omega(\ell, j) - \omega(j, k))\,(\omega(\ell, j) + \omega(j, k))\, x(\ell, j)x(j, k)\\ &= i^2\omega_{\ell k}\sum_j (\omega(\ell, j) - \omega(j, k))\, x(\ell, j)x(j, k) = 0,\end{aligned}$$

by the Ritz combination principle. It was already proved that this quantity vanishes. An off-diagonal component of this result is $\omega(\ell, k) \neq 0$; therefore the remaining sum must vanish. But this is just an off-diagonal element of (3.4). Therefore only diagonal elements of (3.4) are nonzero. It follows that $[\boldsymbol{P}, \boldsymbol{X}]$ is a *diagonal matrix*.

What are the diagonal elements of the matrix $[\boldsymbol{P}, \boldsymbol{X}]$? Let us use the energy matrix

$$\boldsymbol{H} = \frac{1}{2m}\boldsymbol{P}^2 + V(\boldsymbol{q}) \tag{3.5}$$

for the total energy of the system. The total energy is constant, that is, $\dot{\boldsymbol{H}} = \boldsymbol{0}$, thus the energy matrix $\boldsymbol{H}$ must also be constant. In fact, it must be a diagonal matrix like the commutator. Furthermore, the diagonal elements $H(i, i)$ must be related to the ith stationary state. It is therefore plausible to choose $H(i, i)$ to be exactly equal to the constant energy of the system state:

$$H(i, i) = E_i. \tag{3.6}$$

According to (2.26), we have

$$\dot{\boldsymbol{X}} = \frac{2\pi i}{h}(\boldsymbol{H}\boldsymbol{X} - \boldsymbol{X}\boldsymbol{H}), \tag{3.7}$$

and with $\boldsymbol{P} = m\dot{\boldsymbol{X}}$ we get

$$\boldsymbol{P} = \frac{2\pi i\, m}{h}[\boldsymbol{H}, \boldsymbol{X}]. \tag{3.8}$$

Because $\boldsymbol{H}$ represents the total energy, it must be of the form

$$\boldsymbol{H} = \frac{1}{2m}\boldsymbol{P}^2 + V(\boldsymbol{X}), \tag{3.9}$$

with the potential energy $V(\boldsymbol{X})$. The commutator of the potential energy $V(\boldsymbol{X})$ with $\boldsymbol{X}$ is equal to zero and inserting (3.9) into (3.8), one obtains

$$\boldsymbol{P} = \frac{\pi i}{h}[\boldsymbol{P}^2, \boldsymbol{X}]. \tag{3.10}$$

With

$$[\boldsymbol{P}^2, \boldsymbol{X}] = \boldsymbol{P}^2\boldsymbol{X} - \boldsymbol{X}\boldsymbol{P}^2 = \boldsymbol{P}^2\boldsymbol{X} - \boldsymbol{P}\boldsymbol{X}\boldsymbol{P} + \boldsymbol{P}\boldsymbol{X}\boldsymbol{P} - \boldsymbol{X}\boldsymbol{P}^2 = \boldsymbol{P}[\boldsymbol{P}, \boldsymbol{X}] + [\boldsymbol{P}, \boldsymbol{X}]\boldsymbol{P}$$

Equation (3.10) can be simplified to

$$\boldsymbol{P} = \frac{\pi i}{h}(\boldsymbol{P}[\boldsymbol{P}, \boldsymbol{X}] + [\boldsymbol{P}, \boldsymbol{X}]\boldsymbol{P}). \tag{3.11}$$

If we replace $[\boldsymbol{P}, \boldsymbol{X}]$ by the diagonal matrix $\boldsymbol{D}$ (as discussed above), we get

$$\boldsymbol{P} = \frac{\pi i}{h}(\boldsymbol{P}\boldsymbol{D} + \boldsymbol{D}\boldsymbol{P}). \tag{3.12}$$

Componentwise, this reads as

$$p(n, m)d_m + d_n p(n, m) = \frac{h}{\pi i}p(n, m). \tag{3.13}$$

Under the assumption that $p(n, m) \neq 0$ for $n \neq m$, (3.13) supplies

$$d_m + d_n = h/(\pi i).$$

Inasmuch as this condition applies to all m, n, it follows that $d_m = d_n = h/(2\pi i)$, or

$$\boldsymbol{D} = h/(2\pi i)\boldsymbol{I}. \tag{3.14}$$

Thus we have the final result for the commutation relation:

$$\boxed{\boldsymbol{P}\boldsymbol{X} - \boldsymbol{X}\boldsymbol{P} = \frac{h}{2\pi i}\boldsymbol{I}}. \tag{3.15}$$

Note that you can just as well ignore this "derivation" and simply *postulate* this commutation relation, as Born and Jordan did later. By the way, Dirac used a similar strategy with his "q-numbers".

From (3.15), the following commutation relations for powers of $\boldsymbol{X}$ and $\boldsymbol{P}$ can be derived:

$$\boldsymbol{X}^n\boldsymbol{P} - \boldsymbol{P}\boldsymbol{X}^n = n\frac{i\,h}{2\pi}\boldsymbol{X}^{n-1}, \tag{3.16}$$

and

$$\boldsymbol{P}^n\boldsymbol{X} - \boldsymbol{X}\boldsymbol{P}^n = -n\frac{i\,h}{2\pi}\boldsymbol{P}^{n-1}. \tag{3.17}$$

Proof of (3.16) (by induction): The case $n = 1$ corresponds to (3.15). Now suppose that (3.16) is true for n. Multiplication from the left with $\boldsymbol{X}$ yields

$$\boldsymbol{X}^{n+1}\boldsymbol{P} - \boldsymbol{X}\boldsymbol{P}\boldsymbol{X}^n = n\frac{i\,h}{2\pi}\boldsymbol{X}^n.$$

Calculate $\boldsymbol{X}\boldsymbol{P}$ from (3.15) and use it in the second term, which yields

$$\boldsymbol{X}^{n+1}\boldsymbol{P} - \left(\boldsymbol{P}\boldsymbol{X} + \frac{h}{2\pi i}\boldsymbol{I}\right)\boldsymbol{X}^n = n\frac{i\,h}{2\pi}\boldsymbol{X}^n.$$

We can now bring the second term in parentheses to the right side, which yields (3.16) for $n + 1$. **q.e.d.** Note that (3.17) can be proved similarly.

From the commutation relation, we can derive a further useful relation. Let $\boldsymbol{f}(\boldsymbol{P}, \boldsymbol{X})$ be any function of $\boldsymbol{P}$ and $\boldsymbol{X}$. Then

$$\boxed{\boldsymbol{f}\boldsymbol{X} - \boldsymbol{X}\boldsymbol{f} = \frac{h}{2\pi i}\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{P}}} \tag{3.18}$$

and

$$\boxed{\boldsymbol{P}\boldsymbol{f} - \boldsymbol{f}\boldsymbol{P} = \frac{h}{2\pi i}\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{X}}}. \tag{3.19}$$

Proof Suppose that (3.18) and (3.19) were correct for any two functions $\boldsymbol{f}_1$ and $\boldsymbol{f}_2$. Then they are also true for $\boldsymbol{f}_1 + \boldsymbol{f}_2$ and $\boldsymbol{f}_1 \cdot \boldsymbol{f}_2$. This fact is trivial for $\boldsymbol{f}_1 + \boldsymbol{f}_2$. For $\boldsymbol{f}_1 \cdot \boldsymbol{f}_2$, a short calculation shows

$$\boldsymbol{f_1 f_2 X} - \boldsymbol{X f_1 f_2} = \boldsymbol{f}_1(\boldsymbol{f_2 X} - \boldsymbol{X f_2}) + (\boldsymbol{f_1 X} - \boldsymbol{X f_1})\boldsymbol{f}_2 =$$

$$= \frac{h}{2\pi i}\left(\boldsymbol{f_1}\frac{\partial \boldsymbol{f_2}}{\partial \boldsymbol{P}} + \frac{\partial \boldsymbol{f}_1}{\partial \boldsymbol{P}}\boldsymbol{f}_2\right) = \frac{h}{2\pi i}\frac{\partial(\boldsymbol{f_1 f_2})}{\partial \boldsymbol{P}},$$

corresponding to $\boldsymbol{P f_1 f_2} - \boldsymbol{f_1 f_2 P}$. Now these relations (3.18) and (3.19) obviously apply for $\boldsymbol{P}$ and $\boldsymbol{X}$. Therefore, they also apply for each function $\boldsymbol{f}$ that can be expressed in powers of $\boldsymbol{P}$ and $\boldsymbol{X}$. **q.e.d.**

If we solve any $\boldsymbol{H}$ for $\boldsymbol{P} = \boldsymbol{p}(\boldsymbol{X}, \boldsymbol{H})$, and then choose $\boldsymbol{P} = \boldsymbol{f}$ in (3.18) we find

$$\boldsymbol{PX} - \boldsymbol{XP} = \frac{h}{2\pi i}\frac{\partial \boldsymbol{P}}{\partial \boldsymbol{P}} = \frac{h}{2\pi i}\boldsymbol{I}. \tag{3.20}$$

Thus we have shown that the commutation relation (3.15) is valid in general for any Hamiltonian $\boldsymbol{H}$.

Differentiation with Respect to a Matrix

When it comes to differentiating a matrix function with respect to a matrix, we use a very basic type of differentiation, namely if $\boldsymbol{F(X, Y, Z, \ldots)}$ is a function of the independent matrices $\boldsymbol{X, Y, Z, \ldots}$, then we define the *partial derivative* with respect to the matrix $\boldsymbol{X}$ as

$$\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{X}} \overset{\text{def}}{=} \lim_{\epsilon \to 0} \frac{\boldsymbol{F(X + \epsilon I, Y, Z, \ldots)} - \boldsymbol{F(X, Y, Z, \ldots)}}{\epsilon}.$$

With this definition, matrix differentiations give similar results to “normal” differentiation; for example,

$$\frac{\mathrm{d}\boldsymbol{X}}{\mathrm{d}\boldsymbol{X}} = \boldsymbol{I}$$

and

$$\frac{\mathrm{d}\boldsymbol{X}^2}{\mathrm{d}\boldsymbol{X}} = \lim_{\epsilon \to 0}\frac{1}{\epsilon}\sum_k [(X_{nk} + \epsilon\delta_{nk})(X_{km} + \epsilon\delta_{km}) - X_{nk}X_{km}] =$$

$$= (2X_{nm}) = 2\boldsymbol{X}.$$

For the differentiation of a product of two matrix functions, we get

$$\frac{\partial}{\partial \boldsymbol{X}}(\boldsymbol{FG}) = \frac{\partial \boldsymbol{F}}{\partial \boldsymbol{X}}\boldsymbol{G} + \boldsymbol{F}\frac{\partial \boldsymbol{G}}{\partial \boldsymbol{X}}. \tag{3.21}$$

This results from

$$\boldsymbol{F}(\boldsymbol{X}+\epsilon\boldsymbol{I},\boldsymbol{Y},\ldots)\boldsymbol{G}(\boldsymbol{X}+\epsilon\boldsymbol{I},\boldsymbol{Y},\ldots)-\boldsymbol{F}(\boldsymbol{X},\boldsymbol{Y},\ldots)\boldsymbol{G}(\boldsymbol{X},\boldsymbol{Y},\ldots)=$$

$$=\left[\boldsymbol{F}(\boldsymbol{X}+\epsilon\boldsymbol{I},\boldsymbol{Y},\ldots)-\boldsymbol{F}(\boldsymbol{X},\boldsymbol{Y},\ldots)\right]\boldsymbol{G}(\boldsymbol{X}+\epsilon\boldsymbol{I},\boldsymbol{Y},\ldots)$$

$$+\boldsymbol{F}(\boldsymbol{X},\boldsymbol{Y},\ldots)\left[\boldsymbol{G}(\boldsymbol{X}+\epsilon\boldsymbol{I},\boldsymbol{Y},\ldots)-\boldsymbol{G}(\boldsymbol{X},\boldsymbol{Y},\ldots)\right].$$

From (3.21), it follows directly that

$$\frac{\mathrm{d}\boldsymbol{X}^n}{\mathrm{d}\boldsymbol{X}}=n\boldsymbol{X}^{n-1}.$$

Therefore one can differentiate polynomials and power series of matrices similar to “normal” differentiation. For the exponential of a matrix

$$e^{\boldsymbol{X}}\overset{\text{def}}{=}\sum_{i=0}^{\infty}\frac{\boldsymbol{X}^i}{i!}$$

one obtains by termwise differentiation

$$\frac{\mathrm{d}\,e^{\boldsymbol{X}}}{\mathrm{d}\boldsymbol{X}}=e^{\boldsymbol{X}}.$$

3.2 Systems with Several Degrees of Freedom

Thus far, we know that $\boldsymbol{P}$ equals $m\dot{\boldsymbol{X}}$. Putting this formula in the commutation relation (3.15) provides a new form of the commutation relation:

$$\boxed{\boldsymbol{X}\dot{\boldsymbol{X}}-\dot{\boldsymbol{X}}\boldsymbol{X}=\frac{h}{2\pi m i}\boldsymbol{I}}. \tag{3.22}$$

From this form, it becomes obvious that the commutation relation refers to the entity $\boldsymbol{X}$ and its time derivative. However, this is not always the case, as we show later in the case of angular momentum. We can use this knowledge to generalize our findings to systems with several degrees of freedom:

$$\boldsymbol{X_k}\boldsymbol{P_k}-\boldsymbol{P_k}\boldsymbol{X_k}=\frac{-h}{2\pi i}\boldsymbol{I}. \tag{3.23}$$

Because the degree of freedom $\boldsymbol{X}_k$ does not depend on $\boldsymbol{X}_i$ for $i\neq k$, we get for $i\neq k$

$$\boldsymbol{X_k}\boldsymbol{P_i}-\boldsymbol{P_i}\boldsymbol{X_k}=\boldsymbol{0}, \tag{3.24}$$

$$\boldsymbol{X}_k \boldsymbol{X}_i - \boldsymbol{X}_i \boldsymbol{X}_k = \boldsymbol{0}, \tag{3.25}$$

and

$$\boldsymbol{P}_k \boldsymbol{P}_i - \boldsymbol{P}_i \boldsymbol{P}_k = \boldsymbol{0}. \tag{3.26}$$

In general, if $\boldsymbol{F}(\boldsymbol{P}, \boldsymbol{X})_k$ is a function of $\boldsymbol{P}_k$ and $\boldsymbol{X}_k$, then the following relations hold.

$$[\boldsymbol{X}_k, \boldsymbol{F}] = -\frac{h}{2\pi i} \frac{\partial \boldsymbol{F}}{\partial \boldsymbol{P}_k}, \tag{3.27}$$

$$[\boldsymbol{P}_k, \boldsymbol{F}] = \frac{h}{2\pi i} \frac{\partial \boldsymbol{F}}{\partial \boldsymbol{X}_k}. \tag{3.28}$$

In classical physics, a mechanical system is characterized by its energy as a function of the momenta and coordinates. In quantum mechanics, a system is similarly characterized by specifying the *energy function* (energy matrix) $\boldsymbol{H}(\boldsymbol{P}, \boldsymbol{X})$. In close analogy to the classical equations of motion[1]

$$\dot{x} = \frac{\partial H}{\partial p} = \{H, x\} \quad \text{and} \quad \dot{p} = -\frac{\partial H}{\partial x} = \{H, p\} \tag{3.29}$$

we use as quantum mechanical equations of motion these matrix relations, following from (3.7), (3.27), and (3.28):

$$\boxed{\dot{\boldsymbol{X}}_k = \frac{\partial \boldsymbol{H}}{\partial \boldsymbol{P}_k} = \frac{2\pi i}{h}[\boldsymbol{H}, \boldsymbol{X}_k]}, \tag{3.30}$$

$$\boxed{\dot{\boldsymbol{P}}_k = -\frac{\partial \boldsymbol{H}}{\partial \boldsymbol{X}_k} = \frac{2\pi i}{h}[\boldsymbol{H}, \boldsymbol{P}_k]}. \tag{3.31}$$

3.3 Transformations

Matrix differential equations of the form

$$\dot{\boldsymbol{X}} = \boldsymbol{A}\boldsymbol{X},$$

---

[1] Here, we use the Poisson–Jacobi bracket symbol:

$$\{F, G\} = -\{G, F\} = \sum_k \left(\frac{\partial F}{\partial p_k} \frac{\partial G}{\partial x_k} - \frac{\partial G}{\partial p_k} \frac{\partial F}{\partial x_k} \right).$$

.

where $\boldsymbol{A}$ is a constant matrix, have the solution

$$\boldsymbol{X}(t) = \boldsymbol{\Phi}(t)\boldsymbol{X}(0),$$

with the transition matrix [15]

$$\boldsymbol{\Phi}(t) \overset{\text{def}}{=} \sum_{k=0}^{\infty} \frac{1}{k!}(\boldsymbol{A}t)^k = \exp(\boldsymbol{A}t).$$

For the matrix equation

$$\dot{\boldsymbol{X}} = \frac{i}{\hbar}[\boldsymbol{H}, \boldsymbol{X}] \tag{3.32}$$

we can define a similar solution approach:

$$\boxed{\boldsymbol{\Phi}(t) = \exp\left(\frac{i}{\hbar}\boldsymbol{H}t\right)}, \tag{3.33}$$

$$\boxed{\boldsymbol{X}(t) = \boldsymbol{\Phi}(t)\boldsymbol{X}(0)\boldsymbol{\Phi}^{-1}(t)}, \tag{3.34}$$

where $\boldsymbol{X}(0)$ is the value of $\boldsymbol{X}$ at time $t = 0$. With $\dot{\boldsymbol{\Phi}}(t) = \frac{i}{\hbar}\boldsymbol{H}\boldsymbol{\Phi}(t)$, we then get indeed

$$\dot{\boldsymbol{X}} = \dot{\boldsymbol{\Phi}}(t)\boldsymbol{X}(0)\boldsymbol{\Phi}^{-1}(t) + \boldsymbol{\Phi}(t)\boldsymbol{X}(0)\dot{\boldsymbol{\Phi}}^{-1}(t) =$$

$$= \frac{i}{\hbar}\boldsymbol{H}\boldsymbol{X}(t) + \boldsymbol{\Phi}(t)\boldsymbol{X}(0)\left(-\frac{i}{\hbar}\boldsymbol{H}\boldsymbol{\Phi}^{-1}\right) = \frac{i}{\hbar}\left[\boldsymbol{H}\boldsymbol{X}(t) - \boldsymbol{X}(t)\boldsymbol{H}\right].$$

Note that we assumed the commutativity of $\boldsymbol{H}$ and $\boldsymbol{\Phi}^{-1}(t) = \boldsymbol{\Phi}(-t)$, which can be readily verified by expanding the series (3.33) of $\boldsymbol{\Phi}(t)$.

The following summary [14] turns out to be very useful. If the matrices $\boldsymbol{P}_k$, $\boldsymbol{X}_k$ satisfy the canonical commutation rules (3.24)–(3.26) and also have the property that their Hamilton function $\boldsymbol{H}(\boldsymbol{P}, \boldsymbol{X})$ is a diagonal matrix, the canonical equations of motion (3.30) and (3.31) are satisfied. In other words, if only (3.24)–(3.26) are satisfied and $\boldsymbol{H}$ is a diagonal matrix, you have already solved a given quantum mechanical problem!

By introducing the transformations

$$\boldsymbol{X} \Rightarrow \boldsymbol{X}' = \boldsymbol{T}\boldsymbol{X}\boldsymbol{T}^*$$

$$\boldsymbol{P} \Rightarrow \boldsymbol{P}' = \boldsymbol{T}\boldsymbol{P}\boldsymbol{T}^*$$

with

$$\boldsymbol{T} = \exp(-i\boldsymbol{H}t/\hbar) = \boldsymbol{\Phi}(-t),$$

the coordinates and momenta become independent of time. However, a formerly constant vector $\boldsymbol{v}$ now becomes a time-dependent vector $\boldsymbol{T}\boldsymbol{v} = \exp(-i\boldsymbol{H}t/\hbar)\boldsymbol{v}$. But $\exp(-i\boldsymbol{H}t/\hbar)\boldsymbol{v}$ is the solution of the differential equation

$$\dot{\boldsymbol{v}}(t) = -\frac{i}{\hbar}\boldsymbol{H}\boldsymbol{v}(t). \tag{3.35}$$

This is the Schrödinger equation for the state vector $\boldsymbol{v}$; see Chap. 11. In quantum mechanics in the Heisenberg picture the state vector $\boldsymbol{v}$ (see Chap. 4) does not change with time, whereas an observable $\boldsymbol{A}$ satisfies the Heisenberg equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A} = \frac{i}{h}[\boldsymbol{H}, \boldsymbol{A}].$$

3.4 Problems

3.1 Commutation Relations: What is

$$[\boldsymbol{A}, [\boldsymbol{B}, \boldsymbol{C}]] + [\boldsymbol{B}, [\boldsymbol{C}, \boldsymbol{A}]] + [\boldsymbol{C}, [\boldsymbol{A}, \boldsymbol{B}]]?$$

3.2 Powers in Commutation Relations: Show that

$$[\boldsymbol{X}, \boldsymbol{P}^n] = n\, i\hbar \boldsymbol{P}^{n-1}.$$

3.3 Anti-Commutation Relation for Hermitian Matrices: Show that for Hermitian matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, the sum $\boldsymbol{AB} + \boldsymbol{BA}$ is also an Hermitian matrix.

3.4 Coordinate Transformation: What does $\boldsymbol{X}$ look like when a coordinate transformation is applied according to (3.34) with a diagonal matrix $\boldsymbol{H}$?

3.5 Exponential of an Hermitian Matrix: Show that for an Hermitian matrix $\boldsymbol{H}$ the matrix $\exp(i\boldsymbol{H})$ is a unitary matrix.

3.6 Commutation Relation: Show that for Hermitian matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, the matrix $i\hbar\boldsymbol{C} = [\boldsymbol{AB} - \boldsymbol{BA}]$ is an Hermitian matrix.

3.7 Commutation Relation: Show that

$$a)[A, BC] = B[A, C] + [A, B]C$$

and

$$b)[AB, C] = A[B, C] + [A, C]B.$$

3.8 Exponential of a Nilpotent Matrix: What is the exponential matrix $\exp(t\boldsymbol{N})$ for the nilpotent matrix

$$\boldsymbol{N} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}?$$

Chapter 4
Observables and Uncertainty Relations

Abstract State vectors and other matrices are introduced. We also define projection matrices for the interpretation of experiments and density matrices for the description of mixed states. Also, Heisenberg's famous uncertainty relation is derived and interpreted.

4.1 State Vector

In classical physics, the dynamic behaviour of a system is completely described through the state vector

$$\boldsymbol{\xi} \overset{\text{def}}{=} \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{p} \end{pmatrix}.$$

The temporal behaviour of the system is determined by a system of first order differential equations, the Hamilton equations

$$\dot{\boldsymbol{\xi}} = \begin{pmatrix} \dot{\boldsymbol{x}} \\ \dot{\boldsymbol{p}} \end{pmatrix} = \begin{pmatrix} \frac{\partial H}{\partial \boldsymbol{p}} \\ \\ -\frac{\partial H}{\partial \boldsymbol{x}} \end{pmatrix}.$$

The state variables (e.g., in astronomy, the space coordinates x_i and the momentum components p_i of a celestial body) can be measured simultaneously at each given time. These physical entities are real numbers that can be measured with a finite precision, e.g. Δx_i. More specifically, the measured value is a multiple of this finite precision, i.e. $x_i = n_i \cdot \Delta x_i$, where n_i is an integer. Once the product $\Delta x_i \Delta p_i$ of the accuracies Δx_i and Δp_i reach the magnitude of the Planck constant h (which is obviously the case in quantum mechanics), x_i and p_i can no longer be measured simultaneously with arbitrary precision.

G. Ludyk, *Quantum Mechanics in Matrix Form*,
https://doi.org/10.1007/978-3-319-26366-3_4

4.2 The Stern-Gerlach Experiment

In their famous Stern-Gerlach experiment in 1922, the physicists Otto Stern[1] and Walther Gerlach[2] were the first to observe the directional quantization of angular momenta of atoms. The experiment is one of the most fundamental experiments in physics, and it is repeatedly referred to for explaining quantum mechanical effects that cannot be understood in the framework of classical physics.

In this experiment, a beam of (electrically neutral) silver atoms is sent through a vacuum between the poles of a magnet. One of the poles has the form of a cutting edge (parallel to the beam), the other pole looks like a groove in a flat plane. This setup implies that the magnetic field is strongly inhomogeneous in the direction transverse to the beam. Eventually, the silver atoms are displayed on a screen. It turns out that the silver atoms can be found in two separate patches. In other words, the magnetic field splits the beam into two separate sub-beams.

Here is a quantum mechanical explanation of the experiment. The silver atom has a magnetic moment $\vec{\mu}$, which points into the same direction as its angular momentum $\vec{S}$. The magnetic field can be represented as

$$\vec{B}_{\text{total}} = \vec{B}_{\text{homogeneous}} + \vec{B}_{\text{inhomogeneous}} = \left(B_{\text{homogeneous}} + B_{\text{inhomogeneous}}\right) \cdot \vec{e}_z .$$

The angular momentum has a quantum number $\frac{1}{2}$. Therefore, only the settings $\frac{-\hbar}{2}$ or $\frac{+\hbar}{2}$ are allowed in the z-direction (also called “spin down” and “spin up”). In contrast, a classical angular momentum could point in any direction with respect to this axis. In the inhomogeneous field, the force

$$\vec{F} = -\nabla(-\vec{\mu} \cdot \vec{B}_{\text{total}}) = (\vec{\mu} \cdot \nabla)\vec{B}_{\text{total}}) = \begin{pmatrix} 0 \\ 0 \\ \mu_z \cdot \frac{\partial B}{\partial z} \end{pmatrix}$$

acts on the magnetic moment of the atom. Since $\vec{\mu}$ is proportional to $\vec{S}$, the z-component of $\vec{\mu}$ can only have a positive value or a negative value of equal size. Depending on the orientation of the angular momentum, a force that is of the same value but opposite in direction therefore acts perpendicular to the flight direction of the atom. As a result, the beam splits into two sub-beams. From a classical point of view, the magnetic moment $\vec{\mu}$ can occupy any angle with respect to the z-axis. The deflection force would therefore also have a continuously distributed value, and the silver atoms would be visible in a continuous strip.

---

[1] Otto Stern, 1888–1969, German physicist, Nobel Prize 1943.

[2] Walther Gerlach, 1889–1979, German physicist.

4.3 States and Postulates

In the Stern-Gerlach experiment, we can assign two states to the atoms:

$$\text{atoms with } \mu_z > 0 : \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

$$\text{atoms with } \mu_z < 0 : \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The total state is thus characterized by a vector of the form

$$\boldsymbol{\xi} = \frac{1}{\sqrt{\alpha^2 + \beta^2}} \left(\alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) = \frac{1}{\sqrt{\alpha^2 + \beta^2}} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

Quantum Mechanics is an axiomatic theory because it is well-grounded on few postulates. A postulate is a statement, also known as an axiom, which is taken to be true without proof. Postulates are the basic structure from which lemmas and theorems are derived.

Postulate 1 *The physical state of a quantum system at a time t_0 is completely described through a state vector $\boldsymbol{\xi}$.*

In quantum mechanics, all physically measurable properties are described by matrices, for example $\boldsymbol{X}_i$ and $\boldsymbol{P}_i$.

Postulate 2 *Each physically measurable property $\mathcal{A}$ is described by a Hermitian matrix $\mathbf{A}$ in the state space. $\mathbf{A}$ is called the observable.*

Since Hermitian matrices have only real eigenvalues, measurement results are always real, as expected.

Postulate 3 *The measurement of the physical property $\mathcal{A}$ always returns an eigenvalue of the matrix $\mathbf{A}$.*

A quantum state $\boldsymbol{a}$ is called an eigenstate of the matrix $\boldsymbol{A}$ if the action of the matrix on the state returns the same state multiplied by some eigenvalue λ, that is, $\boldsymbol{A}\boldsymbol{a} = \lambda\boldsymbol{a}$. If the quantum system is in the state $\boldsymbol{a}$, then a measurement of the observable $\boldsymbol{A}$ will give the result λ. Note that λ must be a real number (since anything that is physically measurable is a real number).

As an example, quantized results are observed in the Stern-Gerlach experiment. As we just saw, the distribution of possible magnetic moments μ_z is not continuous (as predicted by the classical theory), but limited to two values. These are exactly the eigenvalues ±1 of the matrix

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

It follows that a state variable can be fully described by a linear combination of eigenvectors. In order for the eigenvalues (which are now the measurements results) to be always real, the associated matrices must always be Hermitian. For example, the condition $\boldsymbol{X}_i = \bar{\boldsymbol{X}}_i^{\mathsf{T}} = \boldsymbol{X}_i^{\dagger}$ must hold (i.e., the transposed matrix with complex conjugate elements is equal to the original matrix).

It can be shown that Hermitian matrices (that correspond to physical entities) have a complete set[3] of eigenvectors. Such matrices are also called *observables*, and any possible measurement result of a physical entity corresponding to the observable $\boldsymbol{A}$ can only be an eigenvalue λ_i of $\boldsymbol{A}$. If the state vector $\boldsymbol{\xi}$ can be composed of eigenvectors $\boldsymbol{e}_i$ of $\boldsymbol{A}$, i.e.

$$\boldsymbol{\xi} = \sum_i c_i \, \boldsymbol{e}_i , \tag{4.1}$$

then the probability for the measurement of $\boldsymbol{A}$ to give the result λ_i is equal to $|c_i|^2$ (under the condition that the eigenvectors are orthonormal). For the case that the state vector is equal to an eigenvector $\boldsymbol{e}_i$, you will surely find the measurement result λ_i.

4.4 Projection Matrices

If the eigenvectors $\boldsymbol{e}_i$ are normalized (i.e., $\boldsymbol{e}_i^{\dagger}\boldsymbol{e}_j = 1$ for $i = j$ and $\boldsymbol{e}_i^{\dagger}\boldsymbol{e}_j = 0$ for $i \neq j$), one obtains the coefficients c_j by multiplying (4.1) from the left with the row vector $\boldsymbol{e}_j^{\dagger}$:

$$\boldsymbol{e}_j^{\dagger}\boldsymbol{\xi} = \sum_i c_i \, \boldsymbol{e}_j^{\dagger}\boldsymbol{e}_i = c_j . \tag{4.2}$$

---

[3] A complete set of eigenvectors is a set of eigenvectors so that every vector is a linear combination of the eigenvectors.

If we multiply this equation from the left with the column vector $\boldsymbol{e}_j$, we obtain

$$\boldsymbol{e}_j\boldsymbol{e}_j^\dagger\boldsymbol{\xi} = \boldsymbol{P}_j\,\boldsymbol{\xi} = \sum_i c_i\,\boldsymbol{e}_j\boldsymbol{e}_j^\dagger\boldsymbol{e}_i = \boldsymbol{e}_j\boldsymbol{e}_j^\dagger\sum_i c_i\,\boldsymbol{e}_i = \boldsymbol{P}_j\sum_i c_i\,\boldsymbol{e}_i = c_j\,\boldsymbol{e}_j. \tag{4.3}$$

The matrix

$$\boldsymbol{P}_j \overset{\text{def}}{=} \boldsymbol{e}_j\boldsymbol{e}_j^\dagger \tag{4.4}$$

is called *projection matrix*, since it projects the state vector $\boldsymbol{\xi}$ on the eigenvector $\boldsymbol{e}_j$. Projection matrices, also called *projection operators*, have the following properties:

- $$\boldsymbol{P}_j^2 = \boldsymbol{P}_j,$$

  since $\boldsymbol{P}_j^2 = (\boldsymbol{e}_j\boldsymbol{e}_j^\dagger)(\boldsymbol{e}_j\boldsymbol{e}_j^\dagger) = \boldsymbol{e}_j(\boldsymbol{e}_j^\dagger\boldsymbol{e}_j)\boldsymbol{e}_j^\dagger = \boldsymbol{e}_j\boldsymbol{e}_j^\dagger = \boldsymbol{P}_j$.
- $$\sum_i \boldsymbol{P}_i = \boldsymbol{I}.$$

  Therefore the choice

  $$\boldsymbol{P}_2 = \begin{pmatrix} 0 & 0 & 0\cdots \\ 0 & 1 & 0\cdots \\ 0 & 0 & 0\cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

  and so on define valid projection matrices.

To some extent, the projection matrix describes the preparation of an experiment, since they are *mathematical models* for the experiment!

If we add a shield behind the magnetic field of the Stern-Gerlach apparatus such that the lower radiation component is blocked, then only the upper portion can reach the detector, i.e. only "eigenvectors"

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

The projection matrix is thus composed as

$$\boldsymbol{P}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}(1, 0) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

and we have

$$\boldsymbol{P}_1\boldsymbol{\xi} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\left(\alpha\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \beta\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = \alpha\begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

4.5 Probabilistic Interpretation

We can specify the probability p_i that the result of a measurement is a_i. In order to determine this probability, a large number of measurements must be carried out under steady state conditions. The result of such measurement series is the *expectation value* $\langle A \rangle$ of the observable A. e.g., the spectral line wavelengths of an atom are the mean values based on many transitions from one atom state to another. In the state representation

$$\boldsymbol{\xi} = \sum_i c_i \, \boldsymbol{e}_i ,$$

the element $c_i \, \boldsymbol{e}_i$ describes the *possibility* that the measurement of the appropriate entity related to A gives the eigenvalue a_i. If $\boldsymbol{\xi}$ and $\boldsymbol{e}_i$ are normalized, then according to Max Born the number $c_i^* c_i = \left| c_i^2 \right|$ is the probability that the measurement gives the eigenvalue a_i. Of course, the probabilities are positive and add up to 1:

$$\sum_j c_j^* c_j = 1.$$

The measured mean value of A is then given by

$$\sum_j c_j^* c_j \, a_j = \boldsymbol{\xi}^\dagger \boldsymbol{A} \boldsymbol{\xi},$$

where it was agreed that if the coefficient c_i is complex, one uses the transposed vector $\boldsymbol{\xi}^\dagger$ as

$$\boldsymbol{\xi}^\dagger = \sum_i c_i^* \, \boldsymbol{e}_i^\mathsf{T} ,$$

The above formula can be easily verified, since the eigenvalue equation $\boldsymbol{A}\boldsymbol{e}_i = a_i \boldsymbol{e}_i$ yields

$$\boldsymbol{A}\boldsymbol{\xi} = \boldsymbol{A} \sum_i c_i \, \boldsymbol{e}_i = \sum_i c_i \, \boldsymbol{A}\boldsymbol{e}_i = \sum_i c_i \, a_i \boldsymbol{e}_i .$$

Multiplying with $\boldsymbol{\xi}^\dagger$ from the left, we finally get (with $\boldsymbol{e}_i^\dagger \boldsymbol{e}_j = \delta_{ij}$)

$$\boldsymbol{\xi}^\dagger \boldsymbol{A} \boldsymbol{\xi} = \left(\sum_j c_j^* \, \boldsymbol{e}_j^\dagger \right) \left(\sum_i c_i \, a_i \, \boldsymbol{e}_i \right) = \sum_j \left| c_j^2 \right| a_j = \langle \boldsymbol{A} \rangle . \tag{4.5}$$

Let us sum up our findings:

The probability to measure the eigenvalue a_j of $\boldsymbol{A}$ in the state $\boldsymbol{\xi}$ is given by $\left|c_j^2\right|$. $c_j \boldsymbol{e}_j$ is the projection of the normalized vector $\boldsymbol{\xi}$ on the normalized eigenvector $\boldsymbol{e}_j$. The average value of the observable $\boldsymbol{A}$ for a system being in the state $\boldsymbol{\xi}$ is $\langle \boldsymbol{A} \rangle = \boldsymbol{\xi}^\dagger \boldsymbol{A}\, \boldsymbol{\xi}$. The expectation values are extremely importent in quantum mechanics. If the measurement of the observable $\boldsymbol{A}$ is repeated a large number of times, the average of all the results will approach $\langle \boldsymbol{A} \rangle$. The same is true if we measure the observable $\boldsymbol{A}$ in many independent yet identical systems simultaneously.

4.6 Density Matrix

4.6.1 Definitions

Until now, we considered systems that consist of different particles in specific states. Let now an *ensemble* be a collection of many identical particles. They may be in different states. However, assume the statistical distribution of these states is known. Albert Einstein was an early advocate of the ensemble interpretation [28]:

"The attempt to conceive the quantum-theoretical description as the complete description of the individual systems leads to unnatural theoretical interpretations, which become immediately unnessary if one accepts the interpretation that the description refers to ensembles of systems and not to individual systems."

While the state of a single particle is described by a vector, the state of an ensemble of particles is best described by a *state matrix*, also called *density matrix*. The density matrix $\boldsymbol{D}$ is obtained as follows. In the definition (4.5) of the expectation value $\langle \boldsymbol{A} \rangle$, we can insert the identity matrix in the special form

$$\boldsymbol{I} = \sum_i \boldsymbol{I}_i = \sum_i \boldsymbol{e}_i \boldsymbol{e}_i^\intercal ,$$

where $\boldsymbol{I}_i$ is a diagonal matrix that is 1 only in the i-th diagonal element and zero otherwise:

$$\langle \boldsymbol{A} \rangle = \boldsymbol{\xi}^\dagger \boldsymbol{A} \boldsymbol{\xi} = \boldsymbol{\xi}^\dagger \boldsymbol{A} \boldsymbol{I} \boldsymbol{\xi} = \boldsymbol{\xi}^\dagger \boldsymbol{A} \left(\sum_i \boldsymbol{I}_i \right) \boldsymbol{\xi} =$$

$$= \boldsymbol{\xi}^\dagger \boldsymbol{A} \left(\sum_i \boldsymbol{e}_i \boldsymbol{e}_i^\intercal \right) \boldsymbol{\xi} = \sum_i \underbrace{\boldsymbol{\xi}^\dagger \boldsymbol{A} \boldsymbol{e}_i}\, \underbrace{\boldsymbol{e}_i^\intercal \boldsymbol{\xi}} = \sum_i \underbrace{\boldsymbol{e}_i^\intercal \boldsymbol{\xi}}\, \underbrace{\boldsymbol{\xi}^\dagger \boldsymbol{A} \boldsymbol{e}_i} . \tag{4.6}$$

Now we can introduce the density matrix $\boldsymbol{D}$ as

$$\boldsymbol{D} \overset{\text{def}}{=} \boldsymbol{\xi}\boldsymbol{\xi}^{\dagger} = \sum_{i} c_i\, \boldsymbol{e}_i \cdot \sum_{j} c_j^{*}\, \boldsymbol{e}_j^{\mathsf{T}} = \sum_{i,j} c_i\, \boldsymbol{e}_i \cdot c_j^{*}\, \boldsymbol{e}_j^{\mathsf{T}}. \tag{4.7}$$

The matrix elements of $\boldsymbol{D}$ in the $\boldsymbol{e}$-basis are

$$d(k,\ell) = \boldsymbol{e}_k^{\mathsf{T}} \boldsymbol{D} \boldsymbol{e}_\ell = \boldsymbol{e}_k^{\mathsf{T}} \sum_{i,j} c_i\, \boldsymbol{e}_i \cdot c_j^{*}\, \boldsymbol{e}_j^{\mathsf{T}} = c_k c_\ell^{*}.$$

Generally the density matrix is not diagonal. The *trace of a matrix* is defined as the sum of the elements on the main diagonal. It can be written by means of the normalized eigenvectors $\boldsymbol{e}_\nu$ as

$$\operatorname{trace}(\boldsymbol{X}) \overset{\text{def}}{=} \sum_{\nu} \boldsymbol{e}_\nu^{\mathsf{T}} \boldsymbol{X} \boldsymbol{e}_\nu. \tag{4.8}$$

With this definition, we finally get from (4.6) and (4.7)

$$\boxed{\langle \boldsymbol{A} \rangle = \operatorname{trace}(\boldsymbol{D}\boldsymbol{A}).}. \tag{4.9}$$

The expectation value of a dynamical variable A, represented by the matrix $\boldsymbol{A}$, in the state $\boldsymbol{D}$, is (4.9).

If $\boldsymbol{\xi}$ is a normalized state vector (i.e. length 1), it follows from (4.7)

$$\operatorname{trace}(\boldsymbol{D}) = \sum_{i} \left| c_i^2 \right| = 1. \tag{4.10}$$

Also, we get

$$\boldsymbol{D}^2 = \boldsymbol{D}, \tag{4.11}$$

since

$$\boldsymbol{D}^2 = \boldsymbol{\xi}\, \underbrace{\boldsymbol{\xi}^{\dagger}\boldsymbol{\xi}}_{1}\, \boldsymbol{\xi}^{\dagger} = \boldsymbol{\xi}\boldsymbol{\xi}^{\dagger} = \boldsymbol{D}.$$

4.6.2 Mixed States

If all particles of the considered ensemble are in a definite quantum state, it is called a *pure state*. In this case, the measurement result according to (4.9) is obviously the expectation value. If, however, the particles are in different quantum states, we call it a *mixed state*. Now, if N_ν out of N particles are in the state $\boldsymbol{\xi}_\nu$, then the probability[4] to pick such a particle from the entire ensemble equals

$$p_\nu = \frac{N_\nu}{N},$$

which certainly assumes

$$\sum_\nu p_\nu = 1.$$

For the expectation value, we obtain

$$\langle \boldsymbol{A} \rangle = \sum_\nu p_\nu \boldsymbol{\xi}_\nu^\dagger \boldsymbol{A} \boldsymbol{\xi}_\nu. \tag{4.12}$$

For mixed states, we can now modify the density matrix as

$$\boldsymbol{D}_M \overset{\text{def}}{=} \sum_\nu p_\nu \boldsymbol{\xi}_\nu \boldsymbol{\xi}_\nu^\dagger. \tag{4.13}$$

This provides a straightforward extension of the state concept, since we can now also describe systems whose state is not known in all detail (e.g., as a state vector in a Hilbert space). For this modified density matrix $\boldsymbol{D}_M$, we again have

$$\underline{\underline{\text{trace}(\boldsymbol{D}_M \boldsymbol{A})}} = \sum_\nu \sum_\mu \boldsymbol{e}_\nu^\intercal p_\mu \boldsymbol{\xi}_\mu \boldsymbol{\xi}_\mu^\dagger \boldsymbol{A} \boldsymbol{e}_\nu = \sum_\nu \sum_\mu p_\mu \boldsymbol{\xi}_\mu^\dagger \boldsymbol{A} \boldsymbol{e}_\nu \boldsymbol{e}_\nu^\intercal \boldsymbol{\xi}_\mu =$$

$$= \sum_\mu p_\mu \boldsymbol{\xi}_\mu^\dagger \boldsymbol{A} \boldsymbol{\xi}_\mu = \underline{\underline{\langle \boldsymbol{A} \rangle}}. \tag{4.14}$$

However, we now have

$$\boldsymbol{D}_M^2 \neq \boldsymbol{D}_M, \tag{4.15}$$

since

$$\boldsymbol{D}_M^2 = \sum_\nu p_\nu \boldsymbol{\xi}_\nu \boldsymbol{\xi}_\nu^\dagger \sum_\mu p_\mu \boldsymbol{\xi}_\mu \boldsymbol{\xi}_\mu^\dagger = \sum_\nu p_\nu^2 \boldsymbol{\xi}_\nu \boldsymbol{\xi}_\nu^\dagger. \tag{4.16}$$

If you compare (4.13) with (4.16), the two are equal only if one $p_\nu = 1$, and all others are equal to zero. This is only the case, however, if the state is a *pure state*. For a *mixed state*, equation (4.15) is always true.

---

[4] The same holds if we only have incomplete information on the system, for example if the particle number is very large and when we can only make probability statements.

For the trace of $\boldsymbol{D}$, we obtain

$$\operatorname{trace}(\boldsymbol{D}) = \sum_{\nu} \boldsymbol{e}_{\nu}^{\intercal} \boldsymbol{\xi}_{\nu} \boldsymbol{\xi}_{\nu}^{\dagger} \boldsymbol{e}_{\nu} = \sum_{\nu} \boldsymbol{\xi}_{\nu}^{\dagger} \boldsymbol{e}_{\nu} \boldsymbol{e}_{\nu}^{\intercal} \boldsymbol{\xi}_{\nu} = \sum_{\nu} \boldsymbol{\xi}_{\nu}^{\dagger} \boldsymbol{\xi}_{\nu} = 1, \tag{4.17}$$

and since $\boldsymbol{D}^2 = \boldsymbol{D}$, we also get

$$\operatorname{trace}(\boldsymbol{D}^2) = 1. \tag{4.18}$$

For the mixed state is obtained, however,

$$\begin{aligned}\operatorname{trace}(\boldsymbol{D}_M^2) &= \sum_{\nu} \boldsymbol{e}_{\nu}^{\intercal} \left(\sum_{\mu} p_{\mu} \boldsymbol{\xi}_{\mu} \boldsymbol{\xi}_{\mu}^{\dagger} \sum_{\kappa} p_{\kappa} \boldsymbol{\xi}_{\kappa} \boldsymbol{\xi}_{\kappa}^{\intercal} \right) \boldsymbol{e}_{\nu} = \\ &= \sum_{\nu} \boldsymbol{e}_{\nu}^{\intercal} \left(\sum_{\mu} p_{\mu}^2 \boldsymbol{\xi}_{\mu} \boldsymbol{\xi}_{\mu}^{\dagger} \right) \boldsymbol{e}_{\nu} = \sum_{\nu} p_{\nu}^2 < 1. \end{aligned} \tag{4.19}$$

Note that $\operatorname{trace}(\boldsymbol{D}_M) = 1$ holds as well. With the trace of the squared density matrix, we now have a tool to determine whether the system is in a pure or a mixed state!

4.6.3 Examples

1. For the state vector

   $$\boldsymbol{\xi} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

   one obtains

   $$\boldsymbol{D} = \boldsymbol{\xi}\boldsymbol{\xi}^{\dagger} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} (1\ 0) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

   In fact

   $$\boldsymbol{D}^2 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \boldsymbol{D}.$$

2. For

   $$\boldsymbol{\xi} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}$$

   one gets

   $$\boldsymbol{D} = \boldsymbol{\xi}\boldsymbol{\xi}^{\dagger} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} \frac{1}{\sqrt{2}} (1\ \ -i) = \begin{pmatrix} \frac{1}{2} & -\frac{i}{2} \\ \frac{i}{2} & \frac{1}{2} \end{pmatrix}.$$

   Again we have $\boldsymbol{D}^2 = \boldsymbol{D}$, and for the trace we get $\operatorname{trace}(\boldsymbol{D}) = 1$.

3. In the case of the Stern-Gerlach apparatus, we assigned the vectors $\begin{pmatrix}1\\0\end{pmatrix}$ and $\begin{pmatrix}0\\1\end{pmatrix}$ to the two possible states. If the two vectors occur with the probabilities p_1 and p_2, we obtain the density matrix

$$\boldsymbol{D}_M = p_1 \begin{pmatrix}1\\0\end{pmatrix} (1\ 0) + p_2 \begin{pmatrix}0\\1\end{pmatrix} (0\ 1) = \begin{pmatrix}p_1 & 0\\0 & p_2\end{pmatrix}.$$

In this case, we get

$$\boldsymbol{D}_M^2 = \begin{pmatrix}p_1^2 & 0\\0 & p_2^2\end{pmatrix},$$

i.e., $\boldsymbol{D}_M^2 \neq \boldsymbol{D}_M$ and $\text{trace}(\boldsymbol{D}_M) = 1$. However $\text{trace}(\boldsymbol{D}_M^2) < 1$, since $p_1 < 1$, $p_2 < 1$ and $p_1 + p_2 = 1$ yields $p_1^2 + p_2^2 = 1 - p_1 p_2 < 1$.

4.7 Time Evolution of the Expectation Value

Although the state matrix $\boldsymbol{D}$ is constant, the expectation value $\langle \boldsymbol{A} \rangle = \left\langle \boldsymbol{\xi}^\dagger \boldsymbol{A} \boldsymbol{\xi} \right\rangle$ =trace $(\boldsymbol{D}\boldsymbol{A})$ of an observable $\boldsymbol{A}$ varies over time. This can be readily seen with the help of the Heisenberg equation of motion (if $\boldsymbol{A}$ does not explicitly depend on time):

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A} = \frac{i}{\hbar}\Big[\boldsymbol{H}\boldsymbol{A}\Big] = \frac{i}{\hbar}\Big(\boldsymbol{H}\boldsymbol{A} - \boldsymbol{A}\boldsymbol{H}\Big), \tag{4.20}$$

and thus

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \boldsymbol{A} \rangle = \frac{i}{\hbar}\langle (\boldsymbol{H}\boldsymbol{A} - \boldsymbol{A}\boldsymbol{H}) \rangle ,$$

or, respectively,

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \boldsymbol{A} \rangle = \text{trace}(\boldsymbol{D}\dot{\boldsymbol{A}}).$$

Let us now sum up our findings. A single particle is described by a quantum state $\boldsymbol{\xi}$. The only thing that you can measure, however, is the expectation value $\langle \boldsymbol{A} \rangle = \left\langle \boldsymbol{\xi}^\dagger \boldsymbol{A} \boldsymbol{\xi} \right\rangle$ of that state. In contrast, the density matrix $\boldsymbol{D}$ describes everything that there is to know about the state of an ensemble of particles. If we consider a single particle of this ensemble, we know that it is in the state $\boldsymbol{D}$ and that the expectation value is $\langle \boldsymbol{A} \rangle = \text{trace}(\boldsymbol{D}\boldsymbol{A})$.

4.8 Heisenberg's Uncertainty Principle

In addition to the average value $\langle \boldsymbol{A} \rangle$ (also called the *expectation value*), the *variance* is a good indicator for how strongly the measured values fluctuate. The variance is defined as the average of the squared deviations from the mean value $\langle \boldsymbol{A} \rangle$:

$$(\Delta A)^2 \overset{\text{def}}{=} \left\langle (\boldsymbol{A} - \langle \boldsymbol{A} \rangle \boldsymbol{I})^2 \right\rangle = \left\langle \boldsymbol{A}^2 - 2\boldsymbol{A} \langle \boldsymbol{A} \rangle + \langle \boldsymbol{A} \rangle^2 \boldsymbol{I} \right\rangle = \left\langle \boldsymbol{A}^2 \right\rangle - \langle \boldsymbol{A} \rangle^2 . \tag{4.21}$$

Let us now assume that we would like to measure the two physical parameters a and b, represented by two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$. Let $\boldsymbol{\xi}$ be a normalized state vector. The expectation values of a and b are then

$$\langle \boldsymbol{A} \rangle = \boldsymbol{\xi}^\dagger \boldsymbol{A}\, \boldsymbol{\xi}$$

and

$$\langle \boldsymbol{B} \rangle = \boldsymbol{\xi}^\dagger \boldsymbol{B}\, \boldsymbol{\xi}.$$

The average fluctuation squares are given by

$$(\Delta A)^2 = \boldsymbol{\xi}^\dagger (\boldsymbol{A} - \langle \boldsymbol{A} \rangle \boldsymbol{I})^2 \boldsymbol{\xi}$$

and

$$(\Delta B)^2 = \boldsymbol{\xi}^\dagger (\boldsymbol{B} - \langle \boldsymbol{B} \rangle \boldsymbol{I})^2 \boldsymbol{\xi}.$$

We define a complex matrix

$$\boldsymbol{M} \overset{\text{def}}{=} (\boldsymbol{A} - \langle \boldsymbol{A} \rangle \boldsymbol{I}) + i\, \alpha (\boldsymbol{B} - \langle \boldsymbol{B} \rangle \boldsymbol{I}),$$

where $\alpha > 0$ and real. Due to $(\boldsymbol{M}\boldsymbol{\xi})^\dagger (\boldsymbol{M}\boldsymbol{\xi}) \geq 0$, we get

$$(\boldsymbol{M}\boldsymbol{\xi})^\dagger (\boldsymbol{M}\boldsymbol{\xi}) = \boldsymbol{\xi}^\dagger \boldsymbol{M}^\dagger \boldsymbol{M} \boldsymbol{\xi} =$$

$$= \boldsymbol{\xi}^\dagger \left[(\boldsymbol{A} - \langle \boldsymbol{A} \rangle \boldsymbol{I})^2 + \alpha^2 (\boldsymbol{B} - \langle \boldsymbol{B} \rangle \boldsymbol{I})^2 + i\, \alpha (\boldsymbol{A}\boldsymbol{B} - \boldsymbol{B}\boldsymbol{A}) \right] \boldsymbol{\xi} =$$

$$= (\Delta A)^2 + \alpha^2 (\Delta B)^2 + i\, \alpha\, \boldsymbol{\xi}^\dagger (\boldsymbol{A}\boldsymbol{B} - \boldsymbol{B}\boldsymbol{A}) \boldsymbol{\xi} \geq 0.$$

It follows[5]

$$\alpha^{-1} (\Delta A)^2 + \alpha (\Delta B)^2 \geq -i\, \boldsymbol{\xi}^\dagger [\boldsymbol{A}, \boldsymbol{B}]\, \boldsymbol{\xi} = -i \,\langle [\boldsymbol{A}, \boldsymbol{B}] \rangle . \tag{4.22}$$

Varying α for fixed ΔA and ΔB, we find that the left side of this inequality has its

---

[5] With $\langle [\boldsymbol{A}, \boldsymbol{B}] \rangle \overset{\text{def}}{=} \boldsymbol{\xi}^\dagger [\boldsymbol{A}, \boldsymbol{B}]\, \boldsymbol{\xi}$.

minimum when α satisfies the equation

$$-\alpha^{-2}(\Delta A)^2 + (\Delta B)^2 = 0,$$

i.e., $\alpha = \Delta A/\Delta B$. For this value of α, the inequality (4.22) reads

$$2\Delta A \Delta B \geq -i \ \langle [\boldsymbol{A}, \boldsymbol{B}] \rangle \,,$$

that is,

$$\boxed{\Delta A \Delta B \geq -i \, \frac{1}{2} \, \langle [\boldsymbol{A}, \boldsymbol{B}] \rangle \quad .} \tag{4.23}$$

For $\boldsymbol{A} = \boldsymbol{X}$ and $\boldsymbol{B} = \boldsymbol{P}$, we get in particular

$$\langle [\boldsymbol{X}, \boldsymbol{P}] \rangle = \langle i\hbar \boldsymbol{I} \rangle = i\hbar,$$

and therefore

$$\boxed{\Delta X \Delta P \geq \frac{\hbar}{2} \quad .} \tag{4.24}$$

This is Heisenberg's famous *uncertainty relation*! It shows that the uncertainty ΔP of the momentum must increase to the extent that the uncertainty ΔX of the coordinate decreases, and vice versa. Note, however, that the inequality only holds if $\boldsymbol{AB} \neq \boldsymbol{BA}$ (which is the case here).

The situation changes for systems with multiple degrees of freedom. The equations (3.23)–(3.26) state that $\boldsymbol{X}_k$ does not depend on $\boldsymbol{X}_i$ for $k \neq i$ and therefore $\boldsymbol{X_k P_i} = \boldsymbol{P_i X_k}$, $\boldsymbol{X_k X_i} = \boldsymbol{X_i X_k}$ and $\boldsymbol{P_k P_i} = \boldsymbol{P_i P_k}$. Therefore we get for these observables

$$\Delta X_k \Delta P_i \geq 0 (k \neq i),$$

$$\Delta X_k \Delta X_i \geq 0 (k \neq i)$$

and

$$\Delta P_k \Delta P_i \geq 0 (k \neq i).$$

In other words, we can simultaneously determine both observables of *different particles* with arbitrary precision!

4.9 Problems

4.1 **Uncertainty Principle**: The speed of an electron is 1000 $\frac{m}{s}$, and it is measured with an accuracy of 0.1%. With which accuracy can you measure the position of the electron?

4.2 **Projection Matrices**: Under which condition is the product matrix $\boldsymbol{P} = \boldsymbol{P}_1 \cdot \boldsymbol{P}_2$ of two projection matrices $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ also a projection matrix?

4.3 **Density Matrix**: What are the density matrices for the two pure spin 1/2 states, if the system is in the state a) $\boldsymbol{e}_{3,1} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and b) $\boldsymbol{e}_{2,1} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}$?

4.4 **Projection Matrix**: What are the eigenvalues and eigenvectors of a measurement device that is represented by the projection matrix $\boldsymbol{M} = \boldsymbol{e}_1 \boldsymbol{e}_1^{\top} + \boldsymbol{e}_2 \boldsymbol{e}_2^{\top}$, where e_1 and e_2 are orthogonal vectors?

Chapter 5
The Harmonic Oscillator

Abstract As a first application of the matrix method, the quantum mechanical behavior of the harmonic oscillator is discussed in detail.

5.1 Physics of the Harmonic Oscillator

A typical harmonic oscillator consists of a mass attached to a spring. The oscillator has always been a fruitful dynamic model for an atom or molecule in quantum physics. For example, it is often assumed that the electrons in an atom are elastically coupled (one-dimensional elastic oscillators). Such a one-dimensional system can be mathematically described by

$$m\ddot{x} = -kx,$$

where m is the mass, x is the displacement from the equilibrium position, and k is the spring constant.[1] This can also be written as

$$\ddot{x} + \omega_0^2 x = 0, \tag{5.1}$$

with the angular frequency

$$\omega_0 \overset{\text{def}}{=} \sqrt{k/m}.$$

The Hamiltonian function for this simple system is

$$H = \frac{m}{2}\dot{x}^2 + \frac{m}{2}\omega_0^2 x^2,$$

where the first term is the kinetic energy and the second term is the potential energy. If we introduce the momentum $p = m\dot{x}$, it becomes

---

[1] We assume that the restoring force f is proportional to the deflection; that is, $f = -kx$.

G. Ludyk, *Quantum Mechanics in Matrix Form*,
https://doi.org/10.1007/978-3-319-26366-3_5

$$H = \frac{1}{2m} p^2 + \frac{m}{2} \omega_0^2 x^2. \tag{5.2}$$

Heisenberg [13], and later also Born and Jordan in a more sustainable mathematical framework [5], used this one-dimensional harmonic oscillator as a first case study for quantum mechanics. They tried to determine for this model on which energy levels the electrons can move. For the corresponding quantum system, we use the expression

$$\boldsymbol{H} = \frac{1}{2m} \boldsymbol{P}^2 + \frac{m}{2} \omega_0^2 \boldsymbol{X}^2 \tag{5.3}$$

as the matrix for the total energy. This equation can be slightly modified into

$$\boldsymbol{H} = \hbar\omega_0 \left(\underbrace{\frac{1}{2m\omega_0\hbar} \boldsymbol{P}^2}_{\widetilde{\boldsymbol{P}}^2} + \underbrace{\frac{m\omega_0}{2\hbar} \boldsymbol{X}^2}_{\widetilde{\boldsymbol{X}}^2} \right). \tag{5.4}$$

Therefore it is better to use scaled location and impulse operators:

$$\widetilde{\boldsymbol{X}} \overset{\text{def}}{=} \sqrt{\frac{m\omega_0}{2\hbar}}\, \boldsymbol{X}, \quad \text{respectively} \quad \boldsymbol{X} = \sqrt{\frac{2\hbar}{m\omega_0}}\, \widetilde{\boldsymbol{X}},$$

and

$$\widetilde{\boldsymbol{P}} \overset{\text{def}}{=} \sqrt{\frac{1}{2m\omega_0\hbar}}\, \boldsymbol{P}, \quad \text{respectively} \quad \boldsymbol{P} = \sqrt{2m\omega_0\hbar}\, \widetilde{\boldsymbol{P}}.$$

From basic algebra, we know that $(a + ib)(a - ib) = a^2 + b^2$. For noncommuting matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, however, we obtain

$$(\boldsymbol{A} + i\boldsymbol{B})(\boldsymbol{A} - i\boldsymbol{B}) = \boldsymbol{A}^2 + \boldsymbol{B}^2 + i(\boldsymbol{B}\boldsymbol{A} - \boldsymbol{A}\boldsymbol{B}).$$

Therefore

$$(\widetilde{\boldsymbol{X}} + i\,\widetilde{\boldsymbol{P}})(\widetilde{\boldsymbol{X}} - i\,\widetilde{\boldsymbol{P}}) = \widetilde{\boldsymbol{X}}^2 + \widetilde{\boldsymbol{P}}^2 - i(\widetilde{\boldsymbol{X}}\widetilde{\boldsymbol{P}} - \widetilde{\boldsymbol{P}}\widetilde{\boldsymbol{X}}). \tag{5.5}$$

Let us now introduce new terms for the matrices in parentheses on the left-hand side of the equation:

$$\boldsymbol{A} \overset{\text{def}}{=} \widetilde{\boldsymbol{X}} + i\widetilde{\boldsymbol{P}}. \tag{5.6}$$

Because $\boldsymbol{X}$ and $\boldsymbol{P}$ are Hermitian matrices (and therefore also $\widetilde{\boldsymbol{X}}$ and $\widetilde{\boldsymbol{P}}$), we get

$$\boldsymbol{A}^\dagger = \widetilde{\boldsymbol{X}} - i\widetilde{\boldsymbol{P}}. \tag{5.7}$$

$\boldsymbol{A}$ and $\boldsymbol{A}^\dagger$ do not commute, inasmuch as

$$\underline{\underline{[\boldsymbol{A}, \boldsymbol{A}^\dagger]}} = \boldsymbol{A}\boldsymbol{A}^\dagger - \boldsymbol{A}^\dagger\boldsymbol{A} = (\widetilde{\boldsymbol{X}} + i\widetilde{\boldsymbol{P}})(\widetilde{\boldsymbol{X}} - i\widetilde{\boldsymbol{P}}) - (\widetilde{\boldsymbol{X}} - i\widetilde{\boldsymbol{P}})(\widetilde{\boldsymbol{X}} + i\widetilde{\boldsymbol{P}}) =$$

$$= -\frac{i}{2\hbar}\underbrace{(\boldsymbol{X}\boldsymbol{P} - \boldsymbol{P}\boldsymbol{X})}_{i\hbar\boldsymbol{I}} + \frac{i}{2\hbar}\underbrace{(\boldsymbol{P}\boldsymbol{X} - \boldsymbol{X}\boldsymbol{P})}_{-i\hbar\boldsymbol{I}} = \underline{\underline{\boldsymbol{I}}}. \tag{5.8}$$

Note that we used the fact that the matrices $\boldsymbol{X}$ and $\boldsymbol{P}$ obey the commutation relation

$$\boldsymbol{X}\boldsymbol{P} - \boldsymbol{P}\boldsymbol{X} = i\hbar\boldsymbol{I}. \tag{5.9}$$

Also, we can calculate

$$\boldsymbol{A}^\dagger\boldsymbol{A} = \widetilde{\boldsymbol{X}}^2 + \widetilde{\boldsymbol{P}}^2 - \frac{1}{2}\boldsymbol{I},$$

and therefore

$$\widetilde{\boldsymbol{X}}^2 + \widetilde{\boldsymbol{P}}^2 = \boldsymbol{A}^\dagger\boldsymbol{A} + \frac{1}{2}\boldsymbol{I}.$$

With these findings, we can modify (5.4) into

$$\boldsymbol{H} = \hbar\omega_0\left(\boldsymbol{A}^\dagger\boldsymbol{A} + \frac{1}{2}\boldsymbol{I}\right). \tag{5.10}$$

The following commutation relations hold:

$$[\boldsymbol{H}, \boldsymbol{A}] = [\hbar\omega_0\boldsymbol{A}^\dagger\boldsymbol{A}, \boldsymbol{A}] = \hbar\omega_0[\boldsymbol{A}^\dagger, \boldsymbol{A}]\boldsymbol{A} = -\hbar\omega_0\boldsymbol{A}, \tag{5.11}$$

and

$$[\boldsymbol{H}, \boldsymbol{A}^\dagger] = [\hbar\omega_0\boldsymbol{A}^\dagger\boldsymbol{A}, \boldsymbol{A}^\dagger] = \hbar\omega_0\boldsymbol{A}^\dagger[\boldsymbol{A}, \boldsymbol{A}^\dagger] = \hbar\omega_0\boldsymbol{A}^\dagger. \tag{5.12}$$

In the following, we are looking for the eigenvalues of the energy matrix $\boldsymbol{H}$, because they represent (in a simple atomic model) the energy levels for the electrons and define the radiation that they absorb or release when jumping from one level to another. Let us suppose that we found an eigenvector $\boldsymbol{e}$ of $\boldsymbol{H}$. Then

$$\boldsymbol{H}\boldsymbol{e} = \lambda\boldsymbol{e}. \tag{5.13}$$

Now multiplying (5.11) by $\boldsymbol{e}$ from the right yields

$$\boldsymbol{H}\boldsymbol{A}\boldsymbol{e} - \boldsymbol{A}\boldsymbol{H}\boldsymbol{e} = -\hbar\omega_0\boldsymbol{A}\boldsymbol{e}.$$

With (5.13), we have

$$\underline{\underline{\boldsymbol{H}\boldsymbol{A}\boldsymbol{e} = (\lambda - \hbar\omega_0)\boldsymbol{A}\boldsymbol{e}.}} \tag{5.14}$$

In other words: if $\boldsymbol{e}$ is an eigenvector of $\boldsymbol{H}$ with eigenvalue λ, $\boldsymbol{A}\boldsymbol{e}$ is also an eigenvector of $\boldsymbol{H}$, but with the eigenvalue $\lambda - \hbar\omega_0$. The eigenvalue has thus been reduced by the energy $\hbar\omega_0 = h\nu_0$.

Let us now multiply (5.11) by $\boldsymbol{A}\boldsymbol{e}$ from the right. Using (5.14), we get

$$\boldsymbol{H}\boldsymbol{A}^2\boldsymbol{e} - \underbrace{\boldsymbol{A}\boldsymbol{H}\boldsymbol{A}\boldsymbol{e}}_{(\lambda-\hbar\omega_0)\boldsymbol{A}^2\boldsymbol{e}} = -\hbar\omega_0\boldsymbol{A}^2\boldsymbol{e},$$

which can be rewritten as

$$\underline{\underline{\boldsymbol{H}\boldsymbol{A}^2\boldsymbol{e} = (\lambda - 2\hbar\omega_0)\boldsymbol{A}^2\boldsymbol{e}.}} \tag{5.15}$$

In other words, the eigenvalue was once again reduced by the energy $\hbar\omega_0$. This process can be continued for arbitrary powers of $\boldsymbol{A}$. However, the eigenvalue of the energy $\boldsymbol{H}$ can never become negative! After all, if we multiply $\boldsymbol{H}$ in the form (5.10) from left and right with an eigenvector $\boldsymbol{e}$, we obtain

$$\boldsymbol{e}^\top\boldsymbol{H}\boldsymbol{e} = \lambda\,\boldsymbol{e}^\top\boldsymbol{e} = \hbar\omega_0\left(\boldsymbol{e}^\top\boldsymbol{A}^\dagger\boldsymbol{A}\boldsymbol{e} + \frac{1}{2}\boldsymbol{e}^\top\boldsymbol{e}\right) = \underbrace{(\boldsymbol{A}\boldsymbol{e}\cdot\boldsymbol{A}\boldsymbol{e})}_{\geq 0} + \frac{1}{2}\underbrace{\boldsymbol{e}^\top\boldsymbol{e}}_{\geq 0} \geq 0;$$

that is, the eigenvalues λ of $\boldsymbol{H}$ are all positive. For this reason, the above-explained reduction of energy has to come to an end at some point. Let this point be the basic eigenvector $\boldsymbol{e}_0$.Then we must have $\boldsymbol{A}\boldsymbol{e}_0 = \boldsymbol{0}$, because we could otherwise continue with another reduction loop. For the eigenvalue equation of the Hamiltonian, we obtain

$$\underline{\underline{\boldsymbol{H}\boldsymbol{e}_0}} = \left(\hbar\omega_0\boldsymbol{A}^\dagger\boldsymbol{A} + \frac{1}{2}\hbar\omega_0\right)\boldsymbol{e}_0 = \underline{\underline{\frac{1}{2}\hbar\,\omega_0\,\boldsymbol{e}_0.}} \tag{5.16}$$

Multiplying (5.12) with $\boldsymbol{e}_0$ from the right, we obtain

$$\boldsymbol{H}\boldsymbol{A}^\dagger\boldsymbol{e}_0 - \boldsymbol{A}^\dagger\boldsymbol{H}\boldsymbol{e}_0 = \hbar\omega_0\boldsymbol{A}^\dagger\boldsymbol{e}_0,$$

and therefore

$$\underline{\underline{\boldsymbol{H}\boldsymbol{A}^\dagger\boldsymbol{e}_0 = \hbar\omega_0\left(1 + \frac{1}{2}\right)\boldsymbol{A}^\dagger\boldsymbol{e}_0.}} \tag{5.17}$$

We see that the eigenvalue is now *increased* by the energy $\hbar\omega_0$. For that reason, the matrix $\boldsymbol{A}^\dagger$ is also called the *creation operator*. If we multiply (5.12) once again with the new eigenvector $\boldsymbol{A}^\dagger\boldsymbol{e}_0$ from the right, we get

$$\boldsymbol{H}(\boldsymbol{A}^\dagger)^2\boldsymbol{e}_0 - \boldsymbol{A}^\dagger\boldsymbol{H}\boldsymbol{A}^\dagger\boldsymbol{e}_0 = \hbar\omega_0(\boldsymbol{A}^\dagger)^2\boldsymbol{e}_0,$$

and therefore

$$\underline{\underline{\boldsymbol{H}(\boldsymbol{A}^\dagger)^2\boldsymbol{e}_0 = \hbar\omega_0\left(2+\frac{1}{2}\right)(\boldsymbol{A}^\dagger)^2\boldsymbol{e}_0.}} \tag{5.18}$$

As a general result, we find

$$\boldsymbol{e}_n \overset{\text{def}}{=} (\boldsymbol{A}^\dagger)^n\boldsymbol{e}_0,$$

where n can be an arbitrary large positive integer. We then get

$$\boldsymbol{H}\,\boldsymbol{e}_j = \underbrace{\hbar\omega_0\left(j+\frac{1}{2}\right)}_{\lambda_j=E_j}\boldsymbol{e}_j, \quad j=0,1,2,\ldots, \tag{5.19}$$

which implies

$$\boxed{\lambda_j = E_j = \hbar\omega_0\left(j+\frac{1}{2}\right), \quad j=0,1,2,\ldots} \tag{5.20}$$

In other words, you can pump as much energy into the system as you like. However, the lowest energy level is $E_0 = \frac{\hbar\omega_0}{2}$ (and not 0 as in classical theory). We can collect all energy levels in the diagonal matrix

$$\boxed{\boldsymbol{E} = \hbar\omega_0\begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 & \cdots \\ 0 & \frac{3}{2} & 0 & 0 & \cdots \\ 0 & 0 & \frac{5}{2} & 0 & \cdots \\ 0 & 0 & 0 & \frac{7}{2} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}}. \tag{5.21}$$

The energy spectrum consists of equidistant energy values with the distance $\hbar\omega_0$. When jumping from one level to the next, the harmonic oscillator can therefore absorb or emit only integral multiples of $\hbar\omega_0$ (Fig. 5.1).

What is the shape of the matrices $\boldsymbol{X}$ and $\boldsymbol{P}$ for the harmonic oscillator? If we adopt Heisenberg's assumption that only transitions between neighboring states are possible (i.e., only $x(k,k+1) \neq 0$ and $x(k+1,k) \neq 0$), one obtains the matrix

$$\boldsymbol{X} = \begin{pmatrix} 0 & x(0,1) & 0 & 0 & 0 & \cdots \\ x(1,0) & 0 & x(1,2) & 0 & 0 & \cdots \\ 0 & x(2,1) & 0 & x(2,3) & 0 & \cdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}. \tag{5.22}$$

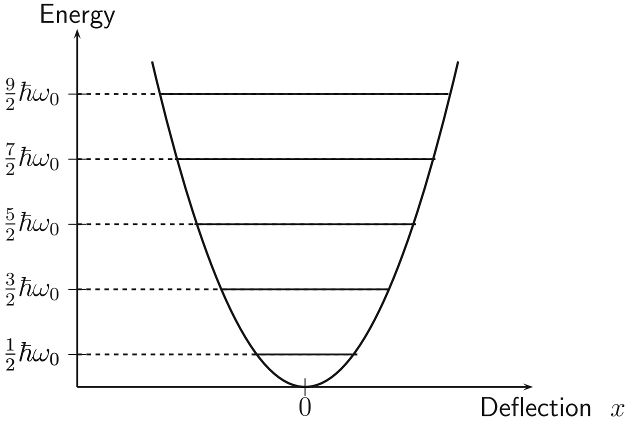


Fig. 5.1 The parabolic potential energy and the equidistant energy levels of the harmonic oscillator (to be continued upwards to infinity)

Because of

$$p(k,\ell) = m\dot{x}(k,\ell) = m\,i\,\omega(k,\ell)x(k,\ell), \tag{5.23}$$

we can calculate the respective matrix for $\boldsymbol{P}$:

$$\boldsymbol{P} = \begin{pmatrix} 0 & p(0,1) & 0 & 0 & 0 & \cdots \\ p(1,0) & 0 & p(1,2) & 0 & 0 & \cdots \\ 0 & p(2,1) & 0 & p(2,3) & 0 & \cdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}. \tag{5.24}$$

Let us now derive the shape of $\boldsymbol{X}$ and $\boldsymbol{P}$ in a more sophisticated way. From (5.10), we get

$$\boldsymbol{H}\boldsymbol{e}_n = \hbar\omega_0\boldsymbol{A}^\dagger\boldsymbol{A}\boldsymbol{e}_n + \frac{\hbar\omega_0}{2}\boldsymbol{e}_n = $$

$$= \hbar\omega_0\left(n + \frac{1}{2}\right)\boldsymbol{e}_n,$$

and therefore with (5.19)

$$\boldsymbol{A}^\dagger\boldsymbol{A}\boldsymbol{e}_n = n\,\boldsymbol{e}_n. \tag{5.25}$$

The matrix $\boldsymbol{A}^\dagger\boldsymbol{A}$ is obviously an Hermitian matrix; that is, its eigenvalues are all real and its eigenvectors form an orthogonal basis. Moreover, one can descend in the series of eigenvectors $\boldsymbol{e}_n$ by multiplying from the left with $\boldsymbol{A}$:

$$\boldsymbol{A}\boldsymbol{e}_n = \alpha_n \boldsymbol{e}_{n-1}. \tag{5.26}$$

We multiply this result by the complex conjugate from the left, which yields

$$\boldsymbol{e}_{\boldsymbol{n}}^{\dagger}\boldsymbol{A}^{\dagger}\boldsymbol{A}\boldsymbol{e}_n = \boldsymbol{e}_{\boldsymbol{n}}^{\dagger}\boldsymbol{e}_n\,|\alpha_n|^2 \overset{!}{=} |\alpha_n|^2\,, \tag{5.27}$$

in order for the length of the eigenvectors to be 1. With (5.25), we get from (5.27)

$$n\boldsymbol{e}_{\boldsymbol{n}}^{\dagger}\boldsymbol{e}_n = n = |\alpha_n|^2\,. \tag{5.28}$$

Choosing α_n as a real number, we get

$$\alpha_n = \sqrt{n}.$$

Equation (5.26) now reads

$$\boldsymbol{A}\boldsymbol{e}_n = \sqrt{n}\boldsymbol{e}_{n-1}. \tag{5.29}$$

Using (5.25), we find

$$\boldsymbol{A}^{\dagger}\boldsymbol{A}\boldsymbol{e}_n = n\,\boldsymbol{e}_n = \boldsymbol{A}^{\dagger}(\sqrt{n}\boldsymbol{e}_{n-1});$$

in other words,

$$\boldsymbol{A}^{\dagger}\boldsymbol{e}_{n-1} = \sqrt{n}\boldsymbol{e}_n,$$

or (shifting to the index n rather than $n-1$)

$$\boldsymbol{A}^{\dagger}\boldsymbol{e}_n = \sqrt{n+1}\boldsymbol{e}_{n+1}. \tag{5.30}$$

This calculation can be repeatedly applied, and we get

$$\boldsymbol{e}_n = \frac{1}{\sqrt{n}}\boldsymbol{A}^{\dagger}\boldsymbol{e}_{n-1} = \frac{\boldsymbol{A}^{\dagger}}{\sqrt{n}}\cdot\frac{\boldsymbol{A}^{\dagger}}{\sqrt{n-1}}\boldsymbol{e}_{n-2} = \cdots\,.$$

As a final result, we find

$$\underline{\underline{\boldsymbol{e}_n = \frac{1}{\sqrt{n!}}\left(\boldsymbol{A}^{\dagger}\right)^n\boldsymbol{e}_0.}} \tag{5.31}$$

The $\boldsymbol{e}_i$ are eigenvectors of an Hermitian matrix, therefore they are mutually orthogonal; that is,

$$\boldsymbol{e}_i^{\dagger}\boldsymbol{e}_j = \delta_{ij}.$$

We therefore obtain the elements $A_{i,k}$ $(i, k = 0, 1, 2, 3, \ldots)$ of the infinitely large matrix $\boldsymbol{A}$ from

$$A_{i,k} = \boldsymbol{e}_i^{\dagger}\boldsymbol{A}\boldsymbol{e}_k = (\boldsymbol{A}^{\dagger}\boldsymbol{e}_i)^{\dagger}\boldsymbol{e}_k \overset{(5.30)}{=} (\sqrt{i+1}\boldsymbol{e}_{i+1})^{\dagger}\boldsymbol{e}_k = \sqrt{i+1}\,\delta_{i+1,k};$$

that is,

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & \sqrt{2} & 0 & \\ 0 & 0 & 0 & \sqrt{3} & \ddots \\ \vdots & & & \ddots & \ddots \end{pmatrix}. \tag{5.32}$$

Accordingly, the elements $A^\dagger_{i,k}$ of the infinitely large matrix $\boldsymbol{A}^\dagger$ are obtained by evaluating

$$A^\dagger_{i,k} = \boldsymbol{e}_i^\dagger \boldsymbol{A}^\dagger \boldsymbol{e}_k = (\boldsymbol{A}\boldsymbol{e}_i)^\dagger \boldsymbol{e}_k \overset{(5.29)}{=} \sqrt{i}\,\boldsymbol{e}^\dagger_{i-1}\boldsymbol{e}_k = \sqrt{i}\,\delta_{i-1,k},$$

$$\boldsymbol{A}^\dagger = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & \\ 0 & \sqrt{2} & 0 & 0 & \\ 0 & 0 & \sqrt{3} & 0 & \ddots \\ \vdots & & \ddots & \ddots & \ddots \end{pmatrix}. \tag{5.33}$$

From (5.6) and (5.7) follows[2]

$$\widetilde{\boldsymbol{X}} = \frac{1}{2}(\boldsymbol{A} + \boldsymbol{A}^\dagger) \quad \text{and} \quad \widetilde{\boldsymbol{P}} = \frac{i}{2}(\boldsymbol{A}^\dagger - \boldsymbol{A}),$$

and thus finally

$$\boldsymbol{X} = \sqrt{\frac{2\hbar}{m\omega_0}}\,\widetilde{\boldsymbol{X}} = \sqrt{\frac{\hbar}{2m\omega_0}}(\boldsymbol{A} + \boldsymbol{A}^\dagger), \tag{5.34}$$

$$\boxed{\boldsymbol{X} = \sqrt{\frac{\hbar}{2m\omega_0}} \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & \cdots \\ 1 & 0 & \sqrt{2} & 0 & 0 & \cdots \\ 0 & \sqrt{2} & 0 & \sqrt{3} & 0 & \cdots \\ 0 & 0 & \sqrt{3} & 0 & \sqrt{4} & \cdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots \end{pmatrix}} \tag{5.35}$$

---

[2] Note that we get in fact

$$\boldsymbol{A}\boldsymbol{A}^\dagger - \boldsymbol{A}^\dagger\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 0 & \cdots \\ 0 & 2 & 0 & \cdots \\ 0 & 0 & 3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} - \begin{pmatrix} 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ 0 & 0 & 2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} = \boldsymbol{I}.$$

$$\boxed{\boldsymbol{X}^2 = \frac{\hbar}{2m\omega_0}\begin{pmatrix} 1 & 0 & \sqrt{2} & 0 & 0 & \cdots \\ 0 & 3 & 0 & \sqrt{2\cdot 3} & 0 & \cdots \\ \sqrt{2} & 0 & 5 & 0 & \sqrt{3\cdot 4} & \cdots \\ 0 & \sqrt{2\cdot 3} & 0 & 7 & 0 & \cdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots \end{pmatrix}}. \tag{5.36}$$

For $\boldsymbol{P}$, we get similar results:

$$\boldsymbol{P} = \sqrt{2\hbar m\omega_0}\widetilde{\boldsymbol{P}} = \sqrt{\frac{\hbar m\omega_0}{2}}(i\boldsymbol{A}^\dagger - i\boldsymbol{A})$$

$$\boxed{\boldsymbol{P} = i\sqrt{\frac{\hbar m\omega_0}{2}}\begin{pmatrix} 0 & -1 & 0 & 0 & 0 & \cdots \\ 1 & 0 & -\sqrt{2} & 0 & 0 & \cdots \\ 0 & \sqrt{2} & 0 & -\sqrt{3} & 0 & \cdots \\ 0 & 0 & \sqrt{3} & 0 & -\sqrt{4} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}}. \tag{5.37}$$

$$\boxed{\boldsymbol{P}^2 = \frac{\hbar m\omega_0}{2}\begin{pmatrix} 1 & 0 & -\sqrt{2} & 0 & 0 & \cdots \\ 0 & 3 & 0 & -\sqrt{2\cdot 3} & 0 & \cdots \\ -\sqrt{2} & 0 & 5 & 0 & -\sqrt{3\cdot 4} & \cdots \\ 0 & -\sqrt{2\cdot 3} & 0 & 7 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}}. \tag{5.38}$$

Note that their form is indeed as predicted in (5.22) and (5.24), which had originally been derived by Heisenberg.

Next, we are looking for the frequency matrix $\boldsymbol{\Omega}$. We have

$$\dot{X}(j,k) = i \cdot \omega(j,k) X(j,k),$$

and we can calculate the elements in the secondary diagonals (on both sides of the main diagonal) from the formula

$$P(j,k) = m \cdot i \cdot \omega(j,k) \cdot X(j,k); \tag{5.39}$$

that is,

$$\omega(j,k) = \frac{P(j,k)}{m \cdot i \cdot X(j,k)}. \tag{5.40}$$

This yields the value $\mp\omega_0$ for the secondary diagonals. In total, the frequency matrix looks like

$$\boldsymbol{\Omega} = \begin{pmatrix} 0 & -\omega_0 & 0 & 0 & 0 & \cdots \\ \omega_0 & 0 & -\omega_0 & 0 & 0 & \cdots \\ 0 & \omega_0 & 0 & -\omega_0 & 0 & \cdots \\ 0 & 0 & \omega_0 & 0 & -\omega_0 & \cdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots \end{pmatrix}. \tag{5.41}$$

According to (5.19), we have

$$E_n = \hbar\omega_0 \left(n + \frac{1}{2}\right) \quad \text{for} \quad n = 0, 1, 2, \ldots. \tag{5.42}$$

Also, we know

$$\omega(n, m) = 2\pi\nu(n, m) = \frac{1}{\hbar}(E_n - E_m).$$

Together with (5.42), this yields

$$\omega(n, m) = (n - m) \cdot \omega_0 \quad \text{for} \quad |n - m| = 1,$$

that is, a perfect agreement with the values in (5.41)!

5.2 Expectation Values and Variances

Let a harmonic oscillator be in the eigenstate $\boldsymbol{e}_j$ with the eigenvalue $(j + 1/2)\hbar\omega_0$. If we choose the energy matrix to be diagonal, the eigenvector components must be $\boldsymbol{e}_j(k) = \delta_{jk}$; that is, the jth component is equal to one and the others are zero. The expectation value of any observable $\boldsymbol{A}$ in the considered eigenstate is then given by

$$\langle A \rangle = \boldsymbol{e}_j{}^{\mathsf{T}} \boldsymbol{A} \boldsymbol{e}_j = a_{jj}.$$

For example, the expectation value of the observable $\boldsymbol{X}$ is equal to zero, because in (5.35) we find $a_{jj} = 0$ for all j. Similarly, the mean value of $\boldsymbol{P}$ is equal to zero. On the other hand, if we calculate the mean value of $\boldsymbol{X}^2$ and $\boldsymbol{P}^2$, it follows from (5.36) that

$$(\boldsymbol{X}^2)_{jj} = \frac{\hbar}{2m\omega_0}(2j - 1),$$

and from (5.38) that

$$(\boldsymbol{P}^2)_{jj} = \frac{\hbar m\omega_0}{2}(2j - 1).$$

The mean values are all zero, thus this value also represents the mean square deviation of the position, namely

$$(\Delta x)^2 = \frac{\hbar}{2m\omega_0}(2j-1).$$

Similarly, we get for the mean square deviation of the momentum

$$(\Delta p)^2 = \frac{\hbar m\omega_0}{2}(2j-1).$$

It follows that

$$(\Delta x \Delta p)^2 = \hbar^2 \left(j - \frac{1}{2}\right)^2,$$

thus Heisenberg's uncertainty relation is fulfilled:

$$\underline{\underline{\Delta x \Delta p \geq \frac{1}{2}\,\hbar.}}$$

In the ground state, neither Δx nor Δp is equal to zero. Instead, $\Delta x \Delta p = \frac{1}{2}\,\hbar > 0$. In this state, the system has the "zero-point energy" $E_0 = \frac{\hbar\omega_0}{2}$, that can never leave the oscillator.

5.3 Problems

5.1 **Commutation Relation**: Show that the standard commutation relations for $\boldsymbol{P}$ and $\boldsymbol{X}$ apply for the harmonic oscillator.

5.2 **Dimension** of $\boldsymbol{A}$: What is the physical dimension of $\boldsymbol{A} = \widetilde{\boldsymbol{X}} + i\widetilde{\boldsymbol{P}}$, as defined in (5.6)?

5.3 **Observables**: Is the matrix $\boldsymbol{A} = \widetilde{\boldsymbol{X}} + i\widetilde{\boldsymbol{P}}$ as defined in (5.6) an observable?

5.4 **Hamiltonian**: According to (5.10) we have $\boldsymbol{H} = \hbar\omega_0(\boldsymbol{N} + \frac{1}{2}\boldsymbol{I})$ with $\boldsymbol{N} \overset{\text{def}}{=} \boldsymbol{A}^\dagger\boldsymbol{A}$. Is $\boldsymbol{N}$ Hermitian?

5.5 **Commutator**: Show that for $\boldsymbol{A} = \widetilde{\boldsymbol{X}} + i\widetilde{\boldsymbol{P}}$, as defined in (5.6), $[\boldsymbol{A}^n, \boldsymbol{N}] = n\boldsymbol{A}^n$ holds.

5.6 **Form of $\boldsymbol{N}$**: What is the form of $\boldsymbol{N}$, and what are the eigenvalues?

5.7 **Form of $\boldsymbol{X}$, $\boldsymbol{X}^2$, and $\boldsymbol{X}^3$**: What do the matrix elements of $\boldsymbol{X}$, $\boldsymbol{X}^2$, and $\boldsymbol{X}^3$ look like? Calculate them by means of the eigenvalue equations of the matrices $\boldsymbol{A}^\dagger$ and $\boldsymbol{A}$.

Chapter 6
Angular Momentum

Abstract The general procedure for the *one-dimensional* harmonic oscillator is now extended to *three-dimensional* systems. In such systems, the three-dimensional angular momentum comes into play. It has an important role in the treatment of atoms and quantum mechanical problems with rotational symmetry. The generalization to three dimensions reached a first climax in Born's, Heisenberg's, and Jordan's famous "Three men work" (Born et al., ZS f. Physik, 1925, [6]). Green explains in [12]:

> Even before the discovery of matrix mechanics, Bohr realized that atomic spectra could only be explained if the angular momentum was limited to certain values, which are integral multiples of Planck's constant $\hbar$. But a full knowledge of the behaviour of the angular momentum of atomic systems was obtained only when matrix mechanics was developed.

Most of the matrices in this chapter are finite-dimensional $N \times N$-matrices with $N \in \mathbb{N} = \{0, 1, 2, 3, \ldots\}$.

6.1 The Matrix Vector of the Angular Momentum

If a mass point revolves around a fixed axis, it has an angular momentum which is defined as

$$\boldsymbol{\ell} \overset{\text{def}}{=} \boldsymbol{r} \times \boldsymbol{p}, \tag{6.1}$$

where

$$\boldsymbol{r} \overset{\text{def}}{=} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \tag{6.2}$$

is the distance vector of the mass point from the axis of rotation and

$$\boldsymbol{p} \overset{\text{def}}{=} m\,\dot{\boldsymbol{r}} \tag{6.3}$$

G. Ludyk, *Quantum Mechanics in Matrix Form*,
https://doi.org/10.1007/978-3-319-26366-3_6

is the momentum of the particle. Componentwise, the equation for the angular momentum (6.1) looks like

$$\ell_1 = x_2 p_3 - x_3 p_2,$$

$$\ell_2 = x_3 p_1 - x_1 p_3,$$

$$\ell_3 = x_1 p_2 - x_2 p_1.$$

In quantum mechanics, the components of the distance vector are defined by three matrices $\boldsymbol{X_1}$, $\boldsymbol{X_2}$, and $\boldsymbol{X_3}$, which are summarized in the matrix vector

$$\mathfrak{R} \overset{\text{def}}{=} \begin{pmatrix} \boldsymbol{X_1} \\ \boldsymbol{X_2} \\ \boldsymbol{X_3} \end{pmatrix}. \tag{6.4}$$

Accordingly, the matrix vector of the momentum is defined as

$$\mathfrak{P} = \begin{pmatrix} \boldsymbol{P_1} \\ \boldsymbol{P_2} \\ \boldsymbol{P_3} \end{pmatrix} \overset{\text{def}}{=} \begin{pmatrix} m\dot{\boldsymbol{X}}_1 \\ m\dot{\boldsymbol{X}}_2 \\ m\dot{\boldsymbol{X}}_3 \end{pmatrix}, \tag{6.5}$$

or, with the matrix vector for the velocity

$$\mathfrak{V} \overset{\text{def}}{=} \begin{pmatrix} \dot{\boldsymbol{X}}_1 \\ \dot{\boldsymbol{X}}_2 \\ \dot{\boldsymbol{X}}_3 \end{pmatrix}, \tag{6.6}$$

also as

$$\mathfrak{P} = m\mathfrak{V}. \tag{6.7}$$

Remember that for the matrices $\boldsymbol{X_i}$ and $\boldsymbol{P_i}$ the commutation relations according to (3.23) to (3.26) hold.

By defining the vector product for matrix vectors as

$$\mathfrak{A} \times \mathfrak{B} = \begin{pmatrix} \boldsymbol{A_1} \\ \boldsymbol{A_2} \\ \boldsymbol{A_3} \end{pmatrix} \times \begin{pmatrix} \boldsymbol{B_1} \\ \boldsymbol{B_2} \\ \boldsymbol{B_3} \end{pmatrix} \overset{\text{def}}{=} \begin{pmatrix} \boldsymbol{A_2 B_3 - A_3 B_2} \\ \boldsymbol{A_3 B_1 - A_1 B_3} \\ \boldsymbol{A_1 B_2 - A_2 B_1} \end{pmatrix},$$

we can show that

$$\mathfrak{A} \times \mathfrak{B} = -\mathfrak{B} \times \mathfrak{A},$$

but only if $\boldsymbol{A}_i$ commutes with $\boldsymbol{B}_j$ for $i \neq j$. Because $\boldsymbol{X}_i$ commutes with $\boldsymbol{P}_j$ for $i \neq j$, we obtain the definition:

Definition (*Matrix vector of angular momentum*)

$$\mathfrak{L} = \begin{pmatrix} \boldsymbol{L}_1 \\ \boldsymbol{L}_2 \\ \boldsymbol{L}_3 \end{pmatrix} \overset{\text{def}}{=} \mathfrak{R} \times \mathfrak{P} = -\mathfrak{P} \times \mathfrak{R}. \tag{6.8}$$

Note, however, that the components of angular momentum do *not* commute! For example, we have

$$[\boldsymbol{L}_1, \boldsymbol{L}_2] = \boldsymbol{L}_1\boldsymbol{L}_2 - \boldsymbol{L}_2\boldsymbol{L}_1 =$$

$$= (\boldsymbol{X}_2\boldsymbol{P}_3 - \boldsymbol{X}_3\boldsymbol{P}_2)(\boldsymbol{X}_3\boldsymbol{P}_1 - \boldsymbol{X}_1\boldsymbol{P}_3) - (\boldsymbol{X}_3\boldsymbol{P}_1 - \boldsymbol{X}_1\boldsymbol{P}_3)(\boldsymbol{X}_2\boldsymbol{P}_3 - \boldsymbol{X}_3\boldsymbol{P}_2)$$

$$= \boldsymbol{X}_2\boldsymbol{P}_3\boldsymbol{X}_3\boldsymbol{P}_1 + \boldsymbol{X}_3\boldsymbol{P}_2\boldsymbol{X}_1\boldsymbol{P}_3 - \boldsymbol{X}_3\boldsymbol{P}_1\boldsymbol{X}_2\boldsymbol{P}_3 - \boldsymbol{X}_1\boldsymbol{P}_3\boldsymbol{X}_3\boldsymbol{P}_2$$

$$= \boldsymbol{X}_1\boldsymbol{P}_2(\boldsymbol{X}_3\boldsymbol{P}_3 - \boldsymbol{P}_3\boldsymbol{X}_3) + \boldsymbol{X}_2\boldsymbol{P}_1(\boldsymbol{P}_3\boldsymbol{X}_3 - \boldsymbol{X}_3\boldsymbol{P}_3)$$

$$= (\boldsymbol{X}_1\boldsymbol{P}_2 - \boldsymbol{X}_2\boldsymbol{P}_1)\, i\, \hbar = i\, \hbar\, \boldsymbol{L}_3.$$

Similarly, we obtain

$$[\boldsymbol{L}_2, \boldsymbol{L}_3] = i\, \hbar\, \boldsymbol{L}_1$$

and

$$[\boldsymbol{L}_3, \boldsymbol{L}_1] = i\, \hbar\, \boldsymbol{L}_2.$$

As a consequence, the observables that are assigned to the components of the angular momentum $\mathfrak{L}$ can *not* be measured at the same time.

For "normal" vectors, a vector multiplied by itself is always zero:

$$\boldsymbol{a} \times \boldsymbol{a} = \boldsymbol{0}.$$

This no longer applies to the matrix vector product in quantum mechanics. It follows from the above that

$$\boxed{\mathfrak{L} \times \mathfrak{L} = i\, \hbar\, \mathfrak{L} \neq \boldsymbol{0}}. \tag{6.9}$$

For the two mutually attracting masses m_e and m_n of an electron and a nucleus in an atom, the following equations of motion apply.

$$m_e \frac{\mathrm{d}^2}{\mathrm{d}t^2} \boldsymbol{x}_e = \boldsymbol{f}, \quad m_n \frac{\mathrm{d}^2}{\mathrm{d}t^2} \boldsymbol{x}_n = -\boldsymbol{f}.$$

By dividing the equations of motion by the masses and subtracting them, we obtain for the relative distance

$$\boldsymbol{x} \overset{\text{def}}{=} \boldsymbol{x_e} - \boldsymbol{x_n} \tag{6.10}$$

an equation of motion that looks like

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}(\boldsymbol{x_e} - \boldsymbol{x_n}) = \left(\frac{1}{m_e} + \frac{1}{m_n}\right)\boldsymbol{f} = \frac{1}{m}\boldsymbol{f},$$

where the *reduced mass* m is defined by

$$m \overset{\text{def}}{=} \frac{m_e m_n}{m_e + m_n}.$$

With the relative velocity $\boldsymbol{v} \overset{\text{def}}{=} \dot{\boldsymbol{x}}$ and the relative momentum

$$\boldsymbol{p} \overset{\text{def}}{=} m\boldsymbol{v},$$

we finally get

$$m\frac{\mathrm{d}^2}{\mathrm{d}t^2}\boldsymbol{x} = \boldsymbol{f} = m\dot{\boldsymbol{v}} = \dot{\boldsymbol{p}}.$$

On the other hand, it follows from (6.10) by differentiation with respect to time that

$$\boldsymbol{v} = \boldsymbol{v_e} - \boldsymbol{v_n} = \frac{1}{m_e}\boldsymbol{p}_e - \frac{1}{m_n}\boldsymbol{p}_n = \frac{1}{m}\boldsymbol{p},$$

which provides an alternative formulation for the relative momentum:

$$\boldsymbol{p} = m\left(\frac{1}{m_e}\boldsymbol{p}_e - \frac{1}{m_n}\boldsymbol{p}_n\right). \tag{6.11}$$

In quantum mechanics, the relative position and the relative momentum of the electron and the nucleus in an atom are represented by the matrices:

$$\boldsymbol{X}_j = \boldsymbol{X}_{ej} - \boldsymbol{X}_{nj}\,,\ j = 1, 2, 3.$$

The velocities and momenta of the electron and the nucleus are represented by the matrices $\boldsymbol{V}_{ej}$ and $\boldsymbol{V}_{nj}$ as well as

$$\boldsymbol{V}_{ej} = \frac{1}{m_e}\boldsymbol{P}_{ej} \quad \text{and} \quad \boldsymbol{V}_{nj} = \frac{1}{m_n}\boldsymbol{P}_{nj}.$$

The relative momentum is defined as

$$\boldsymbol{P}_j = m\left(\frac{1}{m_e}\boldsymbol{P}_{ej} - \frac{1}{m_n}\boldsymbol{P}_{nj}\right) = \frac{1}{m_n + m_e}\left(m_n\boldsymbol{P}_{ej} - m_e\boldsymbol{P}_{nj}\right).$$

The matrices $\boldsymbol{X}_{ej}$ and $\boldsymbol{P}_{ej}$ for the electron commute with the matrices $\boldsymbol{X}_{nj}$ and $\boldsymbol{P}_{nj}$ for the nucleus. This leads to the following commutation relations for the relative position matrices and momentum matrices:

$$\boldsymbol{X}_j\boldsymbol{P}_j - \boldsymbol{P}_j\boldsymbol{X}_j = \frac{m_n}{m_n+m_e}\left(\boldsymbol{X}_{ej}\boldsymbol{P}_{ej} - \boldsymbol{P}_{ej}\boldsymbol{X}_{ej}\right) + \frac{m_e}{m_n+m_e}\left(\boldsymbol{X}_{nj}\boldsymbol{P}_{nj} - \boldsymbol{P}_{nj}\boldsymbol{X}_{nj}\right) =$$

$$= \left(\frac{m_n}{m_n+m_e} + \frac{m_e}{m_n+m_e}\right) i\hbar\boldsymbol{I} = i\hbar\boldsymbol{I}.$$

Thus, the relative matrices satisfy the same conditions as in the case of a single mass.

If we define the scalar product of two matrix vectors as

$$(\mathfrak{A}\cdot\mathfrak{B}) \overset{\text{def}}{=} \boldsymbol{A}_1\boldsymbol{B}_1 + \boldsymbol{A}_2\boldsymbol{B}_2 + \boldsymbol{A}_3\boldsymbol{B}_3,$$

then we get

$$(\mathfrak{L}\cdot\mathfrak{L}) \overset{\text{def}}{=} \boldsymbol{L}^2 = \boldsymbol{L}_1^2 + \boldsymbol{L}_2^2 + \boldsymbol{L}_3^2.$$

6.2 Eigenvalues and Eigenvectors of $\boldsymbol{L}^2$ and $\boldsymbol{L}_3$

6.2.1 *Commutativity of $\boldsymbol{L}^2$ and $\boldsymbol{L}_3$*

First, we show that for the matrix $\boldsymbol{L}^2$ and all matrices $\boldsymbol{L}_i$, $i = 1, 2, 3$ the following relation holds,

$$\boldsymbol{L}^2\boldsymbol{L}_i - \boldsymbol{L}_i\boldsymbol{L}^2 = \boldsymbol{0};$$

that is, $\boldsymbol{L}^2$ commutes with all matrices $\boldsymbol{L}_i$, and therefore $\boldsymbol{L}^2$ and $\boldsymbol{L}_i$ can be measured simultaneously (which is not the case for the $\boldsymbol{L}_i$ among themselves). We prove this result for the two matrices $\boldsymbol{L}^2$ and $\boldsymbol{L}_3$. With the abbreviation

$$[\boldsymbol{A}, \boldsymbol{B}] \overset{\text{def}}{=} \boldsymbol{AB} - \boldsymbol{BA},$$

we get

$$[\boldsymbol{L}_3, \boldsymbol{L}_1^2] = \underbrace{\boldsymbol{L}_3\boldsymbol{L}_1\boldsymbol{L}_1 - \boldsymbol{L}_1\boldsymbol{L}_3\boldsymbol{L}_1}_{[\boldsymbol{L}_3,\boldsymbol{L}_1]\boldsymbol{L}_1} + \underbrace{\boldsymbol{L}_1\boldsymbol{L}_3\boldsymbol{L}_1 - \boldsymbol{L}_1\boldsymbol{L}_1\boldsymbol{L}_3}_{\boldsymbol{L}_1[\boldsymbol{L}_3,\boldsymbol{L}_1]} =$$

$$= i\hbar(\boldsymbol{L}_2\boldsymbol{L}_1 + \boldsymbol{L}_1\boldsymbol{L}_2). \tag{6.12}$$

Similarly, we find that

$$[\boldsymbol{L}_3, \boldsymbol{L}_2^2] = -i\hbar(\boldsymbol{L}_1\boldsymbol{L}_2 + \boldsymbol{L}_2\boldsymbol{L}_1). \tag{6.13}$$

Furthermore, it is clear that

$$[\boldsymbol{L_3}, \boldsymbol{L_3^2}] = \boldsymbol{0}. \tag{6.14}$$

Adding (6.12), (6.13), and (6.14) gives the desired result. The general rule for the commutativity is

$$\boldsymbol{L_i L^2} = \boldsymbol{L^2 L_i},\ i = 1, 2, 3;$$

that is,

$$[\boldsymbol{L_i}, \boldsymbol{L^2}] = \boldsymbol{0}.$$

6.2.2 Eigenvalues and Eigenvectors

We now want to determine the eigenvalues and eigenvectors of the two matrices $\boldsymbol{L^2}$ and $\boldsymbol{L_3}$. $\boldsymbol{L_3}$ is preferred over $\boldsymbol{L_1}$ and $\boldsymbol{L_2}$, inasmuch as we later observe the effect of a magnetic field on atoms, where the field is aligned along the x_3- or z-direction.

Generally speaking, two commuting matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ have the same set of eigenvectors. To see this, let $\boldsymbol{e_A}$ be an eigenvector of $\boldsymbol{A}$ to the eigenvalue λ_A; that is, $\boldsymbol{A e_A} = \lambda_A \boldsymbol{e_A}$. Because $\boldsymbol{A}$ and $\boldsymbol{B}$ commute, we have

$$\boldsymbol{ABe_A} = \boldsymbol{BAe_A} = \boldsymbol{B}(\lambda_A \boldsymbol{e_A}) = \lambda_A \boldsymbol{Be_A}.$$

Hence $\boldsymbol{Be_A}$ is an eigenvector of $\boldsymbol{A}$ with the same eigenvalue λ_A, therefore $\boldsymbol{Be_A}$ must be a multiple of the eigenvector $\boldsymbol{e_A}$:

$$\boldsymbol{Be_A} = \lambda_B \boldsymbol{e_A}.$$

As claimed, the matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ have the same eigenvectors.

The two matrices $\boldsymbol{L_3}$ and $\boldsymbol{L^2}$ commute, therefore they have the same set of eigenvectors. Let us single out one of them, $\boldsymbol{e}$. In addition, we extract the factor $\hbar^2$ from the eigenvalue, such that λ is a purely numerical value[1]

$$\boldsymbol{L^2 e} = \hbar^2 \lambda\, \boldsymbol{e}, \tag{6.15}$$

and

$$\boldsymbol{L_3 e} = \hbar\, \mu\, \boldsymbol{e}. \tag{6.16}$$

Similar to the harmonic oscillator, we can now define the so-called *creation* and *annihilation operators* that allow us to jump from a known eigenvector to the next:

---

[1] Planck's constant $\hbar$ has the dimension $energy \cdot time$. The angular momentum has the dimension $length \cdot mass \cdot length \cdot time^{-1}$. The energy has the dimension $mass \cdot length^2 \cdot time^{-2}$. Therefore, $\hbar$ has the dimension $mass \cdot length^2 \cdot time^{-1}$, that is, the same dimension as the angular momentum. The squared angular momentum then has the same dimension as $\hbar^2$.

$$\boldsymbol{L}_+ \overset{\text{def}}{=} \boldsymbol{L}_1 + i\,\boldsymbol{L}_2, \tag{6.17}$$

and

$$\boldsymbol{L}_- \overset{\text{def}}{=} \boldsymbol{L}_1 - i\,\boldsymbol{L}_2. \tag{6.18}$$

The matrices $\boldsymbol{L}_\pm$ commute with the matrix $\boldsymbol{L}^2$, because

$$\boldsymbol{L}^2\boldsymbol{L}_\pm - \boldsymbol{L}_\pm\boldsymbol{L}^2 = \boldsymbol{L}^2\boldsymbol{L}_1 \pm i\,\boldsymbol{L}^2\boldsymbol{L}_2 - \boldsymbol{L}_1\boldsymbol{L}^2 \mp i\,\boldsymbol{L}_2\boldsymbol{L}^2 =$$

$$= \underbrace{[\boldsymbol{L}^2, \boldsymbol{L}_1]}_{\boldsymbol{0}} + i\,\underbrace{[\boldsymbol{L}^2, \boldsymbol{L}_2]}_{\boldsymbol{0}} = \boldsymbol{0}.$$

On the other hand, we have

$$[\boldsymbol{L}_3, \boldsymbol{L}_\pm] = \pm\hbar\,\boldsymbol{L}_\pm, \tag{6.19}$$

because

$$[\boldsymbol{L}_3, \boldsymbol{L}_\pm] = [\boldsymbol{L}_3, \boldsymbol{L}_1] \pm i[\boldsymbol{L}_3, \boldsymbol{L}_2] = i\hbar\boldsymbol{L}_2 \mp i^2\hbar\boldsymbol{L}_1 =$$

$$= \hbar(\pm\boldsymbol{L}_1 + i\boldsymbol{L}_2) = \pm\hbar\,\boldsymbol{L}_\pm.$$

It also follows that

$$[\boldsymbol{L}_+, \boldsymbol{L}_-] = 2\hbar\boldsymbol{L}_3. \tag{6.20}$$

Solving (6.20) for $\boldsymbol{L}_+\boldsymbol{L}_-$ yields

$$\boldsymbol{L}_+\boldsymbol{L}_- = \boldsymbol{L}^2 - \boldsymbol{L}_3(\boldsymbol{L}_3 - \hbar\boldsymbol{I}) \tag{6.21}$$

and

$$\boldsymbol{L}_-\boldsymbol{L}_+ = \boldsymbol{L}^2 - \boldsymbol{L}_3(\boldsymbol{L}_3 + \hbar\boldsymbol{I}). \tag{6.22}$$

6.2.3 Maximum and Minimum Eigenvalues

Instead of (6.15), we can also write

$$(\boldsymbol{L}_1^2 + \boldsymbol{L}_2^2)\boldsymbol{e} + \hbar^2\mu^2\boldsymbol{e} = \hbar^2\lambda\boldsymbol{e}.$$

If we multiply this with $\boldsymbol{e}^\top$ from the left, we get for the normalized eigenvectors (i.e., $|\boldsymbol{e}| = 1$)

$$\boldsymbol{e}^\top(\boldsymbol{L}_1^2 + \boldsymbol{L}_2^2)\boldsymbol{e} = \hbar^2(\lambda - \mu^2). \tag{6.23}$$

Because the matrices $\boldsymbol{L_1}$ and $\boldsymbol{L_2}$ are Hermitian matrices, the eigenvalues are all real. Therefore, the left side of (6.23) must be positive: $(\lambda - \mu^2) \geq 0$, or $\lambda \geq \mu^2$. Let us now multiply (6.15) with the matrix $\boldsymbol{L}_\pm$ from the left. Because $\boldsymbol{L}_\pm$ commutes with $\boldsymbol{L^2}$, we obtain

$$\boldsymbol{L}_\pm \boldsymbol{L}^2 \boldsymbol{e} = \underline{\underline{\boldsymbol{L}^2 \boldsymbol{L}_\pm \boldsymbol{e} = \hbar^2 \lambda \boldsymbol{L}_\pm \boldsymbol{e}}}; \tag{6.24}$$

that is, $\boldsymbol{L}_\pm \boldsymbol{e}$ is also an eigenvector of $\boldsymbol{L}^2$ with the same eigenvalue $\hbar^2\lambda$. Next, we obtain with (10.21) for $\boldsymbol{L}_3 \boldsymbol{L}_\pm \boldsymbol{e}$

$$\underline{\underline{\boldsymbol{L}_3 \boldsymbol{L}_\pm \boldsymbol{e}}} = ([\boldsymbol{L_3}, \boldsymbol{L}_\pm] + \boldsymbol{L}_\pm \boldsymbol{L_3})\boldsymbol{e} =$$

$$= (\pm\hbar \boldsymbol{L}_\pm + \boldsymbol{L}_\pm \boldsymbol{L}_3)\boldsymbol{e} = \underline{\underline{\hbar(1 \pm \mu)\boldsymbol{L}_\pm \boldsymbol{e}}}. \tag{6.25}$$

In other words, $\boldsymbol{L}_\pm \boldsymbol{e}$ is also an eigenvector of $\boldsymbol{L_3}$. However, the associated eigenvalue changes to $\hbar(1 \pm \mu)$. As in the case of the harmonic oscillator, a multiple application of $\boldsymbol{L}_+$ to the eigenvalue equation subsequently increases the eigenvalue of $\boldsymbol{L_3}$. Of course, this has to stop at some point, inasmuch as otherwise the condition $\lambda \geq \mu^2$ would be violated. Let $\bar{m}$ be the maximum value of λ that does not violate $\lambda \geq \mu^2$. With the definition $\boldsymbol{e_m} \overset{\text{def}}{=} \boldsymbol{L_+^m e}$, we then get

$$\boldsymbol{L_3 e_m} = \hbar \bar{m} \boldsymbol{e_m}. \tag{6.26}$$

Because $\boldsymbol{L_+}$ and $\boldsymbol{L_3}$ commute, it follows that

$$\boldsymbol{L_3(L_+e_m)} = \hbar(1 + \bar{m})\boldsymbol{(L_+e_m)}.$$

But an eigenvalue $\hbar(1 + \bar{m})$ is not possible, because $\bar{m}$ was by definition the largest value of μ! Therefore $\boldsymbol{(L_+e_m) = 0}$ must apply. Under consideration of

$$\boldsymbol{L_-L_+ = L^2 - L_3^2} - \hbar \boldsymbol{L_3},$$

this equation multiplied by the matrix $\boldsymbol{L_-}$ from the left finally gives

$$\boldsymbol{L_-(L_+e_m) = 0},$$

or

$$(\boldsymbol{L^2 - L_3^2} - \hbar \boldsymbol{L_3})\boldsymbol{e}_m = \boldsymbol{0}.$$

Therefore, we get

$$\hbar^2(\lambda - \bar{m}^2 - \bar{m})\boldsymbol{e}_m = \boldsymbol{0}.$$

Inasmuch as the eigenvector $\boldsymbol{e}_m$ cannot be a zero vector, the value in brackets must be zero; that is,

$$\lambda = \bar{m}(\bar{m}+1). \tag{6.27}$$

If we now multiply the second eigenvalue equation (6.26) with the matrix $\boldsymbol{L}_-$ from the left, we obtain

$$\boldsymbol{L}_-\boldsymbol{L}_3\boldsymbol{e}_m = \hbar\bar{m}\boldsymbol{L}_-\boldsymbol{e}_m,$$

or with (6.25)

$$\boldsymbol{L}_3(\boldsymbol{L}_-\boldsymbol{e}_m) = \hbar(\bar{m}-1)(\boldsymbol{L}_-\boldsymbol{e}_m).$$

For the largest eigenvalue $\bar{m}$, we get

$$\boldsymbol{L}^2\boldsymbol{e}_m = \hbar^2\lambda\boldsymbol{e}_m,$$

or

$$\boldsymbol{L}^2(\boldsymbol{L}_-\boldsymbol{e}_m) = \hbar^2\lambda(\boldsymbol{L}_-\boldsymbol{e}_m).$$

In other words, $\boldsymbol{L}_-\boldsymbol{e}_m$ is an eigenvector of $\boldsymbol{L}^2$ and $\boldsymbol{L}_3$ with the eigenvalues $\hbar^2\lambda$ and $\hbar(\bar{m}-1)$, respectively. Similarly, n-times multiplication by $\boldsymbol{L}_-$ yields

$$\boldsymbol{L}_3(\boldsymbol{L}_-^n\boldsymbol{e}_m) = \hbar(\bar{m}-n)(\boldsymbol{L}_-^n\boldsymbol{e}_m).$$

However, there must be a lowest value for m, because otherwise $\lambda \geq m^2$ would be violated. The minimum value of m without violation of $\lambda \geq m^2$ is $\hbar(\bar{m}-n)$. Then

$$\boldsymbol{L}_3(\boldsymbol{L}_-^{n+1}\boldsymbol{e}_m) = \hbar(\bar{m}-(n+1))(\boldsymbol{L}_-^{n+1}\boldsymbol{e}_m) = \boldsymbol{0}.$$

Because $\hbar(\bar{m}-(n+1)) \neq 0$, $(\boldsymbol{L}_-^{n+1}\boldsymbol{e}_m) = \boldsymbol{0}$ must hold; that is, $(\boldsymbol{L}_-\boldsymbol{L}_-^n\boldsymbol{e}_m) = \boldsymbol{0}$ and therefore also $(\boldsymbol{L}_+\boldsymbol{L}_-\boldsymbol{L}_-^n\boldsymbol{e}_m) = \boldsymbol{0}$. Replacing $\boldsymbol{L}_+\boldsymbol{L}_-$, we obtain

$$(\boldsymbol{L}^2 - \boldsymbol{L}_3^2 + \hbar\boldsymbol{L}_3)\boldsymbol{L}_-^n\boldsymbol{e}_m = \boldsymbol{0},$$

or

$$\hbar^2(\lambda - (\bar{m}-n)^2 + (\bar{m}-n))\boldsymbol{L}_-^n\boldsymbol{e}_m = \boldsymbol{0}.$$

Because $\boldsymbol{L}_-^n\boldsymbol{e}_m \neq \boldsymbol{0}$, we get the condition

$$\lambda - (\bar{m}-n)^2 + (\bar{m}-n) = 0.$$

With the value of $\lambda = \bar{m}(\bar{m}+1)$ according to (6.27), we get

$$\bar{m}(\bar{m}+1) - (\bar{m}-n)^2 + (\bar{m}-n) = (n+1)(2\bar{m}-n) = 0.$$

Due to $(n+1) \neq 0, 2\bar{m}-n = 0$ (i.e., $\bar{m} = n/2$) must hold, where n is the number of steps from the maximum eigenvalue $\hbar^2\bar{m}$ to the minimal eigenvalue $\hbar^2(\bar{m}-n)$. This number n is always an integer (including zero). If we set ℓ for $n/2$, $\hbar\ell$ and $-\hbar\ell$ are the maximum and minimum eigenvalues of $\boldsymbol{L_3}$. For a given value of ℓ, the eigenvalues of $\boldsymbol{L_3}$ are therefore equal to $\hbar m$, where $m = \ell, \ell - 1, \ldots, -\ell$. In particular, there are $2\ell + 1$ eigenvalues of $\boldsymbol{L_3}$. Because n is always an integer, the possible values of ℓ are $0, 1/2, 1, 3/2, \ldots$. Note that these fractional quantum numbers automatically occurred in our treatment of angular momentum.

If we denote $\boldsymbol{e_{\ell m}}$ the common eigenvector of $\boldsymbol{L^2}$ and $\boldsymbol{L_3}$ with eigenvalues $\hbar^2\ell(\ell + 1)$ and $\hbar m$, respectively, we finally obtain the following theorem.

Theorem *For the matrices $\boldsymbol{L^2}$ and $\boldsymbol{L_3}$, the eigenvalue equations are*

$$\boldsymbol{L^2}\boldsymbol{e_{\ell m}} = \hbar^2\ell(\ell + 1)\boldsymbol{e_{\ell m}} \tag{6.28}$$

and

$$\boldsymbol{L_3}\boldsymbol{e_{\ell m}} = \hbar m\boldsymbol{e_{\ell m}}, \tag{6.29}$$

where the allowed quantum numbers *are the Angular momentum quantum number,*

$$\ell = 0,\ 1/2,\ 1,\ 3/2, \ldots,$$

and the Magnetic quantum number,

$$m = -\ell,\ -\ell + 1, \ldots,\ \ell - 1,\ \ell.$$

For a given ℓ, the matrix $\boldsymbol{L_3}$ has a spectrum with $\ell + 1$ eigenvalues

$$-\hbar\ell, \quad \hbar(-\ell + 1), \ldots, \quad \hbar(\ell - 1), \quad \hbar\ell.$$

The matrix is a square $(2\ell + 1) \times (2\ell + 1)$-matrix. Therefore in the Heisenberg matrix mechanics also *finitely large* matrices occur in a Hilbert space!

6.2.4 Orientation of the Angular Momentum Vectors

From the above theorem, we obtain for $|\boldsymbol{L}|$ and the z-component of angular momentum $\boldsymbol{L}$

$$|\boldsymbol{L}| = \hbar\sqrt{\ell(\ell + 1)}, \tag{6.30}$$

$$L_3 = \hbar m. \tag{6.31}$$

Because it is not possible in quantum mechanics to determine the components (L_1, L_2) in addition to the length $|\boldsymbol{L}|$ and the z-component L_3 at the same time,

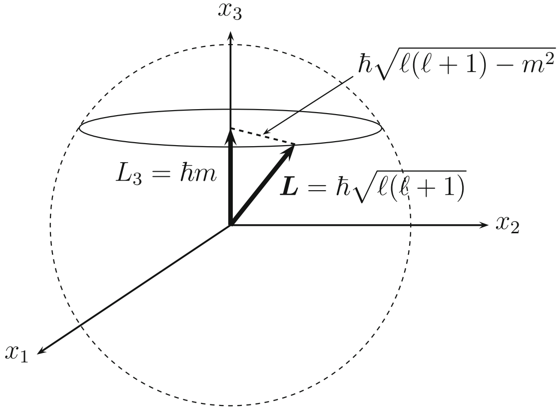


Fig. 6.1 The angular momentum $\boldsymbol{L}$ has a component along the x_3 axis with the allowed values $L_3 = \hbar m,\ -\ell \le m \le \ell$, and undetermined components L_1 and L_2 in the x_1 - x_2 plane

the angular momentum vector $\boldsymbol{L}$ can never be exactly parallel to the z-axis of the coordinate system. In such a case, the x- and the y-components would be exactly zero and thus accurately determined. The only statement that can be made about the x- and y-components is that they make up a circular path together, because

$$L_1^2 + L_2^2 = |\boldsymbol{L}|^2 - L_3^2 = \hbar^2[\ell(\ell+1) - m^2], \tag{6.32}$$

where the right side of this equation is constant for given values of the quantum numbers ℓ and m. The circle has the radius $\hbar\sqrt{\ell(\ell+1)-m^2}$ (see Fig. 6.1).

Our system with the angular momentum $|\boldsymbol{L}| = \hbar\sqrt{\ell(\ell+1)}$ has a well-defined component along the axis x_3, with allowed values $L_3 = \hbar m\ ,(-\ell \le m \le \ell)$, and it has indefinite components L_1 and L_2 in the x_1–x_2 plane. The angular momentum is therefore quantized, and ℓ is called the *angular momentum quantum number*. The quantum number m specifies the x_3-component of the angular momentum. Because m is bounded by the value $2\ell + 1$, the x_3-component of the angular momentum is limited to $2\ell + 1$ discrete values for a given ℓ. This limitation of the angular momentum is also called *space quantization*. This name results from the vector representation of the angular momentum, where the angular momentum is defined by a vector of length $\hbar\sqrt{\ell(\ell+1)}$ and a direction that is determined by its x_3-component of length $\hbar m$; see Fig. 6.2. It is not possible to determine the x_1- and x_2-components of the angular momentum from the quantum numbers ℓ and m.

6.2.5 The Matrices $\boldsymbol{L^2}$ and $\boldsymbol{L_3}$

We can use the eigenvectors as a complete set of orthonormal vectors for constructing a matrix that represents the angular momentum. Because the two matrices $\boldsymbol{L^2}$ and $\boldsymbol{L_3}$ commute, they are both *diagonal matrices* in this representation. In contrast, the

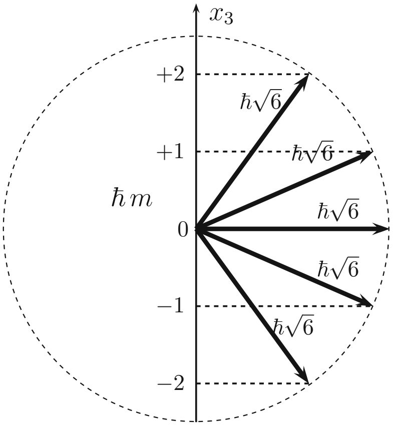


Fig. 6.2 The five (i.e., $2\ell+1$) permitted directions of the angular momentum for $\ell = 2$. The length of the vector is $\hbar\sqrt{\ell(\ell+1)} = \hbar\sqrt{6}$

two matrices $\boldsymbol{L_1}$ and $\boldsymbol{L_2}$ do not commute with the matrix $\boldsymbol{L_3}$, therefore they will not be diagonal. If we multiply the eigenvalue equations as specified in the above theorem with the transposed normalized eigenvector $\boldsymbol{e}_{\boldsymbol{\ell m}}^{\mathsf{T}}$ from the left, we obtain for the diagonal elements

$$\boldsymbol{e}_{\boldsymbol{\ell m}}^{\mathsf{T}} \boldsymbol{L}^{\boldsymbol{2}} \boldsymbol{e}_{\boldsymbol{\ell m}} = L^2_{mm} = \hbar^2\, \ell(\ell+1) \tag{6.33}$$

and

$$\boldsymbol{e}_{\boldsymbol{\ell m}}^{\mathsf{T}} \boldsymbol{L_3} \boldsymbol{e}_{\boldsymbol{\ell m}} = L_{3,mm} = \hbar m. \tag{6.34}$$

We now characterize matrices for the value ℓ by a superscript number, for example, ${\boldsymbol{L_3}}^{(\ell)}$. The matrices $\boldsymbol{L}^{\boldsymbol{2(\ell)}}$ then have the simple diagonal form

$$\underline{\underline{\boldsymbol{L}^{\boldsymbol{2(\ell)}} = \hbar^2\, \ell(\ell+1)\boldsymbol{I}.}} \tag{6.35}$$

For $\ell = \frac{1}{2}$, for example, we obtain the diagonal matrices

$$\boldsymbol{L}^{\boldsymbol{2}^{(1/2)}} = \hbar^2 \begin{pmatrix} 3/4 & 0 \\ 0 & 3/4 \end{pmatrix}, \tag{6.36}$$

$${\boldsymbol{L_3}}^{(1/2)} = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{6.37}$$

For $\ell = 1$, the diagonal matrices are

$$\boldsymbol{L}^{\boldsymbol{2(1)}} = \hbar^2 \begin{pmatrix} \mathbf{2} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{2} \end{pmatrix}, \boldsymbol{L}_{\boldsymbol{3}}^{\boldsymbol{(1)}} = \hbar \begin{pmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{-1} \end{pmatrix},$$

6.2.6 The Matrices $\boldsymbol{L}_+$, $\boldsymbol{L}_-$, $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$

We know

$$\boldsymbol{L}_3\boldsymbol{L}_+\boldsymbol{e}_{\ell m} = \hbar(m+1)\boldsymbol{L}_+\boldsymbol{e}_{\ell m}. \tag{6.38}$$

For the eigenvalue $\hbar(m+1)$, this eigenvalue equation can also be written as

$$\boldsymbol{L}_3\boldsymbol{e}_{\ell,m+1} = \hbar(m+1)\boldsymbol{e}_{\ell,m+1}. \tag{6.39}$$

Because the eigenvalues of $\boldsymbol{L}_3$ in the (6.38) and (6.39) are the same, the eigenvectors can only differ by a multiplicative factor α_m:

$$\boldsymbol{L}_+\boldsymbol{e}_{\ell,m} = \alpha_m\,\boldsymbol{e}_{\ell,m+1}. \tag{6.40}$$

Similarly, we obtain

$$\boldsymbol{L}_-\boldsymbol{e}_{\ell m} = \beta_m\,\boldsymbol{e}_{\ell,m-1}, \tag{6.41}$$

where

$$\alpha_m = \boldsymbol{e}_{\ell,m+1}^{\mathsf{T}}\boldsymbol{L}_+\boldsymbol{e}_{\ell m} \tag{6.42}$$

or

$$\alpha_m^* = \boldsymbol{e}_{\ell,m}^{\mathsf{T}}\boldsymbol{L}_-\boldsymbol{e}_{\ell,m+1}, \tag{6.43}$$

and

$$\beta_m = \boldsymbol{e}_{\ell,m-1}^{\mathsf{T}}\boldsymbol{L}_-\boldsymbol{e}_{\ell,m} \tag{6.44}$$

or

$$\beta_{m+1} = \boldsymbol{e}_{\ell,m}^{\mathsf{T}}\boldsymbol{L}_-\boldsymbol{e}_{\ell,m+1}. \tag{6.45}$$

A comparison of (6.43) and (6.45) provides

$$\alpha_m^* = \beta_{m+1}. \tag{6.46}$$

Multiplication of (6.40) by the matrix $\boldsymbol{L}_-$ from the left yields

$$\boldsymbol{L}_-\boldsymbol{L}_+\boldsymbol{e}_{\ell m} = \alpha_m\,\boldsymbol{L}_-\boldsymbol{e}_{\ell,m+1}. \tag{6.47}$$

Replacing $\boldsymbol{L}_-\boldsymbol{L}_+$ by $\boldsymbol{L}^2 - \boldsymbol{L}_3^2 - \hbar\boldsymbol{L}_3$ and using (6.41) leads to

$$(\boldsymbol{L}^2 - \boldsymbol{L}_3^2 - \hbar\boldsymbol{L}_3)\boldsymbol{e}_{\ell m} = \alpha_m\beta_{m+1}\boldsymbol{e}_{\ell m}$$

or

$$(\ell(\ell+1) - m^2 - m)\hbar^2\boldsymbol{e}_{\ell m} = |\alpha_m|^2\,\boldsymbol{e}_{\ell m}.$$

This implies

$$\alpha_m = [\ell(\ell+1) - m(m+1)]^{1/2}\,\hbar. \tag{6.48}$$

With this α_m, we obtain

$$\boldsymbol{L_+ e_{\ell m}} = [\ell(\ell+1) - m(m+1)]^{1/2}\,\hbar \boldsymbol{e_{\ell,m+1}}, \tag{6.49}$$

or

$$\boldsymbol{e_{\ell' m'}}^{\intercal} \boldsymbol{L_+ e_{\ell m}} = [\ell(\ell+1) - m(m+1)]^{1/2}\,\hbar\,\delta_{\ell\ell'}\delta_{m',m+1}. \tag{6.50}$$

Similarly, we obtain

$$\boldsymbol{e_{\ell' m'}}^{\intercal} \boldsymbol{L_- e_{\ell m}} = [\ell(\ell+1) - m(m-1)]^{1/2}\,\hbar\,\delta_{\ell\ell'}\delta_{m',m-1}. \tag{6.51}$$

The equations (6.50) and (6.51) provide all the elements of the matrices $\boldsymbol{L_+}$ and $\boldsymbol{L_-}$. The Kronecker-Delta functions indicate that all nonvanishing matrix elements occur in blocks along the diagonal $\ell' = \ell$. The block matrices belonging to the values $\ell = 0$, $\frac{1}{2}$, and 1 are shown below. The matrices $\boldsymbol{L_1}$ and $\boldsymbol{L_2}$ can now be derived from the following relations.

$$\boldsymbol{L_1} = \frac{1}{2}(\boldsymbol{L_+} + \boldsymbol{L_-}) \quad \text{and} \quad \boldsymbol{L_2} = \frac{1}{2i}(\boldsymbol{L_+} - \boldsymbol{L_-}). \tag{6.52}$$

For $\ell = 0$, we get

$$\boldsymbol{L_+ = L_- = L_1 = L_2 = 0}. \tag{6.53}$$

For $\ell = 1/2$, we get

$$\boldsymbol{L_+^{(1/2)}} = \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \boldsymbol{L_-^{(1/2)}} = \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

$$\boldsymbol{L_1^{(1/2)}} = \frac{1}{2}\hbar \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad \boldsymbol{L_2^{(1/2)}} = \frac{1}{2}\hbar \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{6.54}$$

Finally, for $j = 1$ we have

$$\boldsymbol{L_+^{(1)}} = \hbar \begin{pmatrix} 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2} \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \boldsymbol{L_-^{(1)}} = \hbar \begin{pmatrix} 0 & 0 & 0 \\ \sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{pmatrix},$$

$$\boldsymbol{L_1^{(1)}} = \frac{1}{\sqrt{2}}\hbar \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{and} \quad \boldsymbol{L_2^{(1)}} = \frac{1}{\sqrt{2}}\hbar \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}. \tag{6.55}$$

Note that all matrices $\boldsymbol{L}_1$, $\boldsymbol{L}_2$, and $\boldsymbol{L}_3$ are Hermitian matrices; that is, $\boldsymbol{L}_i^\dagger = \bar{\boldsymbol{L}}_i^\intercal = \boldsymbol{L}_i$ for all i. Also,

$$(\boldsymbol{L}_1^{(\ell)})^2 + (\boldsymbol{L}_2^{(\ell)})^2 + (\boldsymbol{L}_3^{(\ell)})^2 =$$

$$= \frac{1}{2}\hbar^2 \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} + \frac{1}{2}\hbar^2 \begin{pmatrix} 1 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 1 \end{pmatrix} + \hbar^2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} =$$

$$= \hbar^2 \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \boldsymbol{L}^{2(\ell)}.$$

A given ℓ-value defines a $(2\ell+1) \times (2\ell+1)$ matrix. This dimension reflects that the space is spanned by the $2\ell+1$ different m-values. If we allow all possible ℓ-values, we obtain an infinitely large matrix consisting of $(2\ell+1) \times (2\ell+1)$-blocks along the main diagonal. Their general form can be written as a direct sum ($\alpha = 1, 2, 3, +$ or $-$):

$$\boldsymbol{L}_\alpha = \bigoplus\nolimits_{\ell=0}^{\infty} \boldsymbol{L}_\alpha^{(\ell)} = \begin{pmatrix} \boldsymbol{L}_\alpha^{(0)} & \boldsymbol{0} & \boldsymbol{0} & \cdots \\ \boldsymbol{0} & \boldsymbol{L}_\alpha^{(1)} & \boldsymbol{0} & \cdots \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{L}_\alpha^{(2)} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

For integer values of ℓ, we obtain matrices such as

$$\boldsymbol{L}_1 = \frac{\hbar}{2} \left(\begin{array}{c|ccc|ccccc|c}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \cdots \\ \cline{2-10}
0 & 0 & \sqrt{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \cdots \\
0 & \sqrt{2} & 0 & \sqrt{2} & 0 & 0 & 0 & 0 & 0 & 0 \cdots \\
0 & 0 & \sqrt{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \cdots \\ \hline
0 & 0 & 0 & 0 & 0 & \sqrt{4} & 0 & 0 & 0 & 0 \cdots \\
0 & 0 & 0 & 0 & \sqrt{4} & 0 & \sqrt{6} & 0 & 0 & 0 \cdots \\
0 & 0 & 0 & 0 & 0 & \sqrt{6} & 0 & \sqrt{6} & 0 & 0 \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & \sqrt{6} & 0 & \sqrt{4} & 0 \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{4} & 0 & 0 \cdots \\ \hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ddots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}\right),$$

$$L_2 = i\frac{\hbar}{2}\left(\begin{array}{c|ccc|ccccc|cc}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ \hline
0 & 0 & \sqrt{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & -\sqrt{2} & 0 & \sqrt{2} & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & -\sqrt{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ \hline
0 & 0 & 0 & 0 & 0 & \sqrt{4} & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & -\sqrt{4} & 0 & \sqrt{6} & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & -\sqrt{6} & 0 & \sqrt{6} & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & -\sqrt{6} & 0 & \sqrt{4} & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -\sqrt{4} & 0 & 0 & \cdots \\ \hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ddots & \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \ddots
\end{array}\right)$$

and

$$L_3 = \hbar\left(\begin{array}{c|ccc|ccccc|cc}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ \hline
0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ \hline
0 & 0 & 0 & 0 & -2 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & \cdots \\ \hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ddots & \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \ddots
\end{array}\right).$$

With these matrices, we obtain

$$L^2 = L_1^2 + L_2^2 + L_3^2 = \hbar^2\left(\begin{array}{c|ccc|ccccc|cc}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ \hline
0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ \hline
0 & 0 & 0 & 0 & 6 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 6 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 6 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 6 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 6 & 0 & \cdots \\ \hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ddots & \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \ddots
\end{array}\right).$$

6.3 Problems

6.1 Commutation Relation: What is $\mathfrak{X} \cdot \mathfrak{P} - \mathfrak{P} \cdot \mathfrak{X}$ for the harmonic oscillator?

6.2 Matrix Vectors: Prove the identity

$$(\mathfrak{A} \times \mathfrak{B}) \cdot \mathfrak{C} = \mathfrak{A} \cdot (\mathfrak{B} \times \mathfrak{C}).$$

6.3 Matrices $\boldsymbol{L}_+$ and $\boldsymbol{L}_-$: Are $\boldsymbol{L}_+$ and $\boldsymbol{L}_-$ Hermitian?

6.4 Eigenvector for $\boldsymbol{L}_+$ and $\boldsymbol{L}_-$: Let $\boldsymbol{e}(j, m)$ be normalized, common eigenvectors of $\boldsymbol{L}^2$ and $\boldsymbol{L}_3$. Show that

$$\boldsymbol{L}_+\boldsymbol{e}(j, m) = \hbar\sqrt{(j - m)(j + m + 1)}\,\boldsymbol{e}(j, m + 1),$$

and

$$\boldsymbol{L}_-\boldsymbol{e}(j, m) = \hbar\sqrt{(j + m)(j - m + 1)}\,\boldsymbol{e}(j, m - 1).$$

Chapter 7
Wolfgang Pauli and the Hydrogen Atom

Abstract It was up to the physicist Wolfgang Pauli to apply the new matrix quantum mechanics to the hydrogen atom successfully.

7.1 Basic Matrices and Matrix Vectors

In his work, Pauli [18] first introduced some matrices that we already know from the angular momentum. This is not surprising, however, because the hydrogen atom essentially consists of a nucleus and one electron orbiting around it. At least, that was the general image in the older atomic theory.

The Cartesian coordinates are again represented by the matrices $\boldsymbol{X}_1$, $\boldsymbol{X}_2$, $\boldsymbol{X}_3$ and combined into the matrix vector

$$\mathfrak{R} \overset{\text{def}}{=} \begin{pmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \\ \boldsymbol{X}_3 \end{pmatrix},$$

which satisfies the relation

$$\boldsymbol{R}^2 = (\mathfrak{R} \cdot \mathfrak{R}) = \boldsymbol{X}_1^2 + \boldsymbol{X}_2^2 + \boldsymbol{X}_3^2. \tag{7.1}$$

Next, Pauli defined the momentum matrices $\boldsymbol{P}_1 \overset{\text{def}}{=} m\dot{\boldsymbol{X}}_1$, $\boldsymbol{P}_2 \overset{\text{def}}{=} m\dot{\boldsymbol{X}}_2$, and $\boldsymbol{P}_3 \overset{\text{def}}{=} m\dot{\boldsymbol{X}}_3$, which are summarized in the matrix vector

$$\mathfrak{P} \overset{\text{def}}{=} \begin{pmatrix} \boldsymbol{P}_1 \\ \boldsymbol{P}_2 \\ \boldsymbol{P}_3 \end{pmatrix} = m\frac{\mathrm{d}}{\mathrm{d}t}\mathfrak{R}.$$

G. Ludyk, *Quantum Mechanics in Matrix Form*,
https://doi.org/10.1007/978-3-319-26366-3_7

Also, he derived the relations[1] from equations (3.23) to (3.26)

$$\boldsymbol{P}_\rho \boldsymbol{P}_\sigma - \boldsymbol{P}_\sigma \boldsymbol{P}_\rho = \mathbf{0}, \quad \boldsymbol{X}_\rho \boldsymbol{X}_\sigma - \boldsymbol{X}_\sigma \boldsymbol{X}_\rho = \mathbf{0}, \tag{7.2}$$

$$\boldsymbol{P}_\rho \boldsymbol{X}_\sigma - \boldsymbol{X}_\sigma \boldsymbol{P}_\rho = \begin{cases} \mathbf{0} & \text{for } \rho \neq \sigma \\ \frac{h}{2\pi i}\boldsymbol{I} & \text{for } \rho = \sigma. \end{cases} \tag{7.3}$$

He further assumed that for an arbitrary function $\boldsymbol{F}$ of $\boldsymbol{R}, \boldsymbol{X_1}, \boldsymbol{X_2}, \boldsymbol{X_3}$, this relation is valid for all i (see (3.19) and (3.27)):

$$\boldsymbol{P_i}\boldsymbol{F} - \boldsymbol{F}\boldsymbol{P_i} = \frac{h}{2\pi i}\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{X}_i} \tag{7.4}$$

In particular, it is true for $\boldsymbol{F} = \boldsymbol{R}$:

$$\boldsymbol{P_i}\boldsymbol{R} - \boldsymbol{R}\boldsymbol{P_i} = \frac{h}{2\pi i}\boldsymbol{X}_i \boldsymbol{R}^{-1}. \tag{7.5}$$

With these relations, together with the energy equation for a single particle,

$$\boldsymbol{E} = \frac{1}{2m}\mathfrak{P}^2 + \boldsymbol{F}(\boldsymbol{X_1}, \boldsymbol{X_2}, \boldsymbol{X_3}), \tag{7.6}$$

where $\boldsymbol{E}$ is a diagonal matrix, as well as the Heisenberg equation for any quantity $\boldsymbol{\Phi}$,

$$\frac{h}{2\pi i}\dot{\boldsymbol{\Phi}} = \boldsymbol{E}\boldsymbol{\Phi} - \boldsymbol{\Phi}\boldsymbol{E},$$

he obtained the equations of motion for all i as

$$\frac{\mathrm{d}\boldsymbol{P}_i}{\mathrm{d}t} = -\frac{\partial \boldsymbol{E}}{\partial \boldsymbol{X}_i}. \tag{7.7}$$

---

[1] Pauli summarizes these relationships as

$$\boldsymbol{R}\mathfrak{R} = \mathfrak{R}\boldsymbol{R}$$

and

$$\mathfrak{P}\boldsymbol{R} - \boldsymbol{R}\mathfrak{P} = \frac{h}{2\pi i}\mathfrak{R}/\boldsymbol{R},$$

which of course is not formally correct for both row-vector matrices and for column-vector matrices. Rather, one must write, for example, for column-vector matrices $\mathfrak{R}$ and $\mathfrak{P}$:

$$(\boldsymbol{I_3} \otimes \boldsymbol{R})\mathfrak{R} = \mathfrak{R}\boldsymbol{R}$$

and

$$\mathfrak{P}\boldsymbol{R} - (\boldsymbol{I_3} \otimes \boldsymbol{R})\mathfrak{P} = \frac{h}{2\pi i}\mathfrak{R}\boldsymbol{R^{-1}},$$

where $\otimes$ is the *Kronecker product*.

Hereby, Pauli postulated the existence of a matrix $\boldsymbol{R}$ that satisfies the relations (7.1) and (7.5). Next, he introduced the vector matrix $\mathfrak{L}$ corresponding to the angular momentum of the particle with respect to the origin. Similar to (6.8), with $m\mathfrak{V}$ instead of $\mathfrak{P}$, we get the following definition.

Definition The angular momentum of a particle is

$$\mathfrak{L} \overset{\text{def}}{=} m(\mathfrak{R} \times \mathfrak{V}) = -m(\mathfrak{V} \times \mathfrak{R}). \tag{7.8}$$

7.2 Introduction of the Matrix Vector $\mathfrak{A}$

For the further calculations, we need the so-called *Laplace–Runge–Lenz vector*. This vector was introduced by Lenz[2] into quantum theory. An atom contains an electron with mass m and charge e. The electron is attracted by the fixed nucleus with the charge $+Ze$ due to the Coulomb force. The matrix $\boldsymbol{E}$ is set as

$$\boldsymbol{E} = \frac{1}{2m}\mathfrak{P}^2 - Ze^2\boldsymbol{R}^{-1}. \tag{7.9}$$

Note that $\boldsymbol{E}$ is supposed to be diagonal. The equations of motion corresponding to (7.7) are

$$\dot{\mathfrak{P}} = m\ddot{\mathfrak{R}} = -Ze^2\mathfrak{R}\boldsymbol{R}^{-1}. \tag{7.10}$$

We can now define the Lenz-Matrix-Vector.

Definition The Lenz-Matrix-Vector is given by

$$\mathfrak{A} \overset{\text{def}}{=} \frac{1}{Ze^2m}\frac{1}{2}(\mathfrak{L} \times \mathfrak{P} - \mathfrak{P} \times \mathfrak{L}) + \mathfrak{R}\boldsymbol{R}^{-1}. \tag{7.11}$$

Note that the symmetrized vector matrix difference $\frac{1}{2}(\mathfrak{L} \times \mathfrak{P} - \mathfrak{P} \times \mathfrak{L})$ was used for the vector product $\boldsymbol{\ell} \times \boldsymbol{p}$ in the Lenz-Vector, in order for the vector matrix components $\boldsymbol{A}_i$ to be anti-symmetrical.

In a somewhat lengthy, but trivial calculation (see Problem 7.2), Pauli also showed that

---

[2] See Appendix D.

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathfrak{R}\boldsymbol{R}^{-1}) = \frac{1}{2m}\left[\mathfrak{L} \times (\mathfrak{R}\boldsymbol{R}^{-3}) - (\mathfrak{R}\boldsymbol{R}^{-3}) \times \mathfrak{L}\right]. \tag{7.12}$$

Using this relation, one can show (see Problem 7.3) that the time derivative of $\mathfrak{A}$ is equal to zero. In other words, $\mathfrak{A}$ is a constant matrix vector over time.

Summing up, we now have three time-constant matrix vectors or matrices $\mathfrak{A}$, $\mathfrak{L}$, and $\boldsymbol{E}$ that describe the system (i.e., the atom) completely. We do not need the coordinate matrices $\boldsymbol{X}_i$ and $\boldsymbol{R}$ anymore. Let us summarize the governing equations for $\mathfrak{A}$, $\mathfrak{L}$, and $\boldsymbol{E}$.

$$\mathfrak{L} \times \mathfrak{L} = i\hbar\mathfrak{L}, \tag{7.13}$$

$$[\boldsymbol{A}_i, \boldsymbol{L}_i] = \boldsymbol{0}, \tag{7.14}$$

$$[\boldsymbol{A}_i, \boldsymbol{L}_j] = i\hbar\varepsilon_{ijk}\boldsymbol{A}_k, \text{for} \quad i \neq j \neq k, \tag{7.15}$$

$$\mathfrak{A} \cdot \mathfrak{L} = \mathfrak{L} \cdot \mathfrak{A} = \boldsymbol{0}, \tag{7.16}$$

$$\mathfrak{A} \times \mathfrak{A} = \frac{h}{2\pi i}\frac{2}{mZ^2e^4}\mathfrak{L}\boldsymbol{E}, \tag{7.17}$$

$$\mathfrak{A}^2 = \frac{2}{mZ^2e^4}\boldsymbol{E}\left(\mathfrak{L}^2 + \frac{h^2}{4\pi^2}\boldsymbol{I}\right) + \boldsymbol{I}. \tag{7.18}$$

Equation (7.13) is identical to equation (6.9). Equations (7.14) and (7.15) are analogous to (7.2) and (7.3) in form. Equation (7.18) is analogous to the classical equation (D.5). These equations are proofed in Appendix H.

Because the matrices $\boldsymbol{L}_i$ and $\boldsymbol{A}_i$ commute with the matrix $\boldsymbol{E}$, the quantities that are represented by them can be measured at a given energy value. Once the atom has a specific energy, the variables that correspond to the matrices $\boldsymbol{L}_i$ and $\boldsymbol{A}_i$ describe the simplified system completely. Inasmuch as the matrices $\boldsymbol{X}_i$ and $\boldsymbol{P}_i$ do not commute with $\boldsymbol{E}$, they cannot be measured for a given energy value.

Let ϵ be this fixed energy value; that is, $\boldsymbol{E} = \epsilon \cdot \boldsymbol{I}$. We introduce yet another matrix vector

$$\mathfrak{K} \overset{\text{def}}{=} \sqrt{-\frac{mZ^2e^4}{2\epsilon}}\,\mathfrak{A}.$$

Because of (7.16), we certainly also have

$$\mathfrak{L} \cdot \mathfrak{K} = \boldsymbol{0}.$$

Furthermore, (7.18) implies

$$\boldsymbol{\mathfrak{K}}^2 = \left(\boldsymbol{\mathfrak{L}}^2 + \frac{h^2}{\pi^2}\boldsymbol{I}\right) + \frac{mZ^2e^4}{2\epsilon}\boldsymbol{I}.$$

Relation (7.17) now reads

$$\boldsymbol{\mathfrak{K}} \times \boldsymbol{\mathfrak{K}} = \frac{ih}{2\pi}\boldsymbol{\mathfrak{L}}.$$

We further define the matrix vectors

$$\boldsymbol{\mathfrak{M}} \overset{\text{def}}{=} \frac{1}{2\hbar}(\boldsymbol{\mathfrak{L}} + \boldsymbol{\mathfrak{K}})$$

and

$$\boldsymbol{\mathfrak{N}} \overset{\text{def}}{=} \frac{1}{2\hbar}(\boldsymbol{\mathfrak{L}} - \boldsymbol{\mathfrak{K}}).$$

Based on the relations (7.13) to (7.18), it is easy to see that the following relations hold.

$$\boldsymbol{\mathfrak{M}}^2 - \boldsymbol{\mathfrak{N}}^2 = \frac{1}{\hbar^2}(\boldsymbol{\mathfrak{L}} \cdot \boldsymbol{\mathfrak{K}}) = \boldsymbol{0},$$

$$2(\boldsymbol{\mathfrak{M}}^2 + \boldsymbol{\mathfrak{N}}^2) = \frac{1}{\hbar^2}(\boldsymbol{\mathfrak{L}}^2 + \boldsymbol{\mathfrak{K}}^2) = -\left(1 + \frac{mZ^2e^4}{2\hbar^2\epsilon}\right)\boldsymbol{I},$$

$$\boldsymbol{\mathfrak{M}} \times \boldsymbol{\mathfrak{M}} = i\boldsymbol{\mathfrak{M}},$$

and

$$\boldsymbol{\mathfrak{N}} \times \boldsymbol{\mathfrak{N}} = i\boldsymbol{\mathfrak{N}}.$$

Therefore, we get

$$\boldsymbol{\mathfrak{M}}^2 = \boldsymbol{\mathfrak{N}}^2$$

and

$$4\,\boldsymbol{\mathfrak{M}}^2 = -\left(1 + \frac{mZ^2e^4}{2\hbar^2\epsilon}\right)\boldsymbol{I}. \tag{7.19}$$

For negative values of ϵ (as assumed by Niels Bohr in his atom model), the square root $\sqrt{-\frac{mZ^2e^4}{2\epsilon}}$ is real. In this case, the matrices $\boldsymbol{K}_i$, $\boldsymbol{M}_i$ and $\boldsymbol{N}_i$ represent real entities. The commutation relations for the $\boldsymbol{M}_i$ are the same as for the matrices $\boldsymbol{L}_i$ representing the angular momentum (except a factor $\hbar$). For this reason, the matrix

$$\boldsymbol{\mathfrak{M}}^2 = \boldsymbol{M}_1^2 + \boldsymbol{M}_2^2 + \boldsymbol{M}_3^2$$

can only have the eigenvalues

$$\ell(\ell+1), \ell = 1, \frac{3}{2}, 2, \ldots.$$

From (7.19), it follows that

$$4\ell(\ell+1) = -1 - \frac{mZ^2e^4}{2\hbar^2\epsilon};$$

that is,

$$-\frac{mZ^2e^4}{2\hbar^2\epsilon} = 4\ell(\ell+1) + 1 = (2\ell+1)^2.$$

This yields for the energy value

$$\epsilon = -\frac{m(Ze^2)^2}{2\hbar^2(2\ell+1)^2}.$$

If we introduce the quantum number

$$n \overset{\text{def}}{=} 2\ell + 1,$$

then n must be one of the numbers

$$1, 2, 3, 4, \ldots,$$

and the possible energy values are (with the usual notation E instead of ϵ)

$$\boxed{E_n = -\frac{m(Ze^2)^2}{2\hbar^2 n^2}}. \tag{7.20}$$

Did you notice that these are the exact same values as predicted in Bohr's atom model (1.29)? We solved the eigenvalue problem of the hydrogen atom! For $Z = 1$, (7.20) defines its energy spectrum. n is called the *principal quantum number*. In addition to the principal quantum number, there is still the *angular momentum quantum number* ℓ and the *magnetic quantum number* m (see the chapter on angular momentum). Note that this equation was derived without describing the hydrogen atom in detail, such as electrons orbiting around the nucleus.

7.3 The Hydrogen Spectrum

In Sect. 2.1 we have already provided some information on the hydrogen spectrum. Back then, however, the relationships were more or less an educated guess; see Balmer. Based on the results of Pauli's calculations, we can now give a more profound summary.

For an arbitrary nuclear charge (with Z protons), we get an energy

$$E_n = -\frac{me^4}{8\varepsilon_0^2 h^2}\frac{Z^2}{n^2} = -13,6\,\frac{Z^2}{n^2}\,\mathrm{eV}, \tag{7.21}$$

where ϵ_0 is the so-called *permittivity of free space*. For the energy difference between the n_1th and the n_2th state, we obtain

$$\Delta E = E_{n_2} - E_{n_1} = \frac{me^4 Z^2}{8\varepsilon_0^2 h^2}\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right). \tag{7.22}$$

For $n_2 > n_1$, this energy difference is positive; that is, the total energy of the system is increased by external energy input. Otherwise energy is emitted. This so-called *Rydberg formula* was found by Johannes Rydberg[3] in 1888. Rydberg had no knowledge of any atomic model, but found this formula based on the observed line spectra only. A few years earlier, Balmer[4] had found the famous Balmer formula for the case $n_1 = 2$ in the hydrogen atom ($Z = 1$),

$$\lambda = A\left(\frac{n^2}{n^2-4}\right) = A\left(\frac{n^2}{n^2-2^2}\right),$$

which covers the visible region of the spectrum lines. The empirical constant is $A = 364.56\,\mathrm{nm} = 3645.6 \times 10^{-10}\,\mathrm{m}$. For the explanation of the spectra, we are interested in the frequency. According to Einstein and Planck, it is $E = h\nu$. For a jump from the n_1th to the n_2th state ($n_1 > n_2$), the frequency of the emitted radiation is therefore

$$\nu = \frac{c}{\lambda} = \frac{me^4}{8\varepsilon_0^2 h^3}\left(\frac{1}{n_2^2} - \frac{1}{n_1^2}\right) = cR\left(\frac{1}{n_2^2} - \frac{1}{n_1^2}\right), \tag{7.23}$$

with the *Rydberg constant*

$$R = \frac{me^4}{8\varepsilon_0^2 h^2} = 10\,973\,731.568\,539\,(55)\,m^{-1}.$$

---

[3] Johannes Robert Rydberg, 1854–1919, Swedish physicist.

[4] Johann Jakob Balmer, 1825–1898, Swiss mathematics teacher.

If we set n_1 to infinity in (7.22), we obtain the energy that is needed to move an electron from infinity to the state n_2, so the total energy of the state n_2.

7.4 Problems

7.1 **Time Derivative of $\mathfrak{P}$**: Under which conditions is (7.10) equal to $\dot{\mathfrak{P}} = -Ze^2\mathfrak{R}\boldsymbol{R}^{-1}$?

7.2 **Time Invariant Matrix**: Show the relation (7.12),

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathfrak{R}\boldsymbol{R}^{-1}) = \frac{1}{2m}\left[\mathfrak{L} \times (\mathfrak{R}\boldsymbol{R}^{-3}) - (\mathfrak{R}\boldsymbol{R}^{-3}) \times \mathfrak{L}\right].$$

7.3 **Constancy of $\mathfrak{A}$**: Show that the Lenz-Matrix-Vector $\mathfrak{A}$ is constant.

Chapter 8
Spin

Abstract We introduce the spin based on symmetry considerations. Its effect is described by spinors and Pauli matrices. Also, spin-orbit coupling is investigated.

8.1 Magnetic Fields and Light

In the nineteenth century, the influence of magnetic fields on light was studied extensively. Based on the idea of classical physics that light is an electromagnetic wave caused by atomic vibrations, in 1892 Lorentz[1] theoretically derived that the spectral lines are split threefold once the radiating atom is placed in a magnetic field. In 1896, the Dutch physicist Zeeman[2] confirmed this splitting in an experiment. Subsequent measurements of the splitting showed that it even complies with Lorentz's formula if it is assumed that not the entire atom, but only the much lighter electron vibrates. Already back then, electrons were assumed to be part of the atom. Thanks to this so-called *normal Zeeman effect*, the electron hypothesis was considerably strengthened.

In addition to the normal Zeeman effect, however, a great number of observations were made where more than three lines appear. This so-called *anomalous* Zeeman effect was an inexplicable phenomenon both for classical physics and Bohr's atom model and stimulated further theoretical investigation.

8.2 Derivation of the Zeeman Effect (Without Spin)

According to classical physics, an electron with the mass m and the electric charge e that moves on a circular path with the angular momentum $\boldsymbol{L}$ has a *magnetic moment*

$$\boldsymbol{p}_m = -\frac{e}{2m}\boldsymbol{L}.$$

---

[1] Hendrik Antoon Lorentz, 1853–1928, Dutch mathematician and physicist, Nobel Prize 1902.

[2] Pieter Zeeman, 1865–1943, Dutch physicist, Nobel Prize 1902.

G. Ludyk, *Quantum Mechanics in Matrix Form*,
https://doi.org/10.1007/978-3-319-26366-3_8

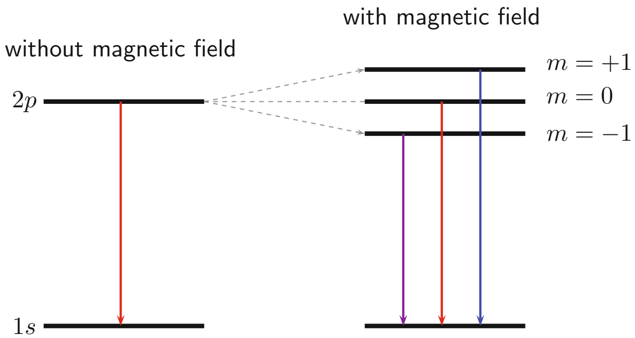


Fig. 8.1 Splitting of spectral lines due to the Zeeman effect. Photons with three different energies can be emitted

The factor $\frac{e}{2m}$ is called the *gyromagnetic ratio*. Under the influence of a magnetic field $\mathfrak{b}$, there is an additional contribution to the potential energy of the electron in the Hamiltonian function to

$$E_m = -\boldsymbol{p_m} \cdot \mathfrak{b} = \frac{e}{2m}(L_3 \cdot b),$$

where we assumed the magnetic field to be parallel to the x_3-axis, having the magnitude b. If we transfer this relationship into quantum mechanics, using the eigenvalue $\hbar m$ of $\boldsymbol{L_3}$, we get

$$E_m = \frac{e}{2m} \cdot b \cdot \hbar m = \mu_{Bohr} \cdot b \cdot m,$$

with the *Bohr magneton*

$$\mu_{Bohr} \overset{\text{def}}{=} \frac{e\hbar}{2m} = 9.2732 \cdot 10^{-24} \frac{J}{T}.$$

Overall, we now get the energy

$$E_{n,m} \overset{\text{def}}{=} E_n + \mu_{Bohr}\, b\, m.$$

The so-called Zeeman splitting then yields an energy difference for $\Delta m = \pm 1$ (Fig. 8.1):

$$\Delta E = E_{n,m+1} - E_{n,m} = \pm \mu_{Bohr}\, b.$$

For the so-called one-electron atoms [21] (i.e., *one* electron on the outermost electron shell), the standard terminology for the energy levels $\ell = 0, 1, 2, 3$ is s, p, d, and f. For the hydrogen atom and other one-electron atoms, these letters are preceded by a number that indicates the total energy level. For example, the lowest energy level in the hydrogen atom is $1s$, the next $2s$ and $2p$, then $3s$, $3p$ and $3d$, and so on. For each n, the value of ℓ ranges from 0 to $n-1$, and for each ℓ we have $2\ell+1$ values of m. The total number of states with an energy value of E_n is therefore

$$\sum_{\ell=0}^{n-1}(2\ell+1) = 2\frac{n(n-1)}{2} + n = n^2.$$

The alkali metals lithium, sodium, potassium, and so on are one-electron atoms that consist of $Z-1$ inner electrons and one external electron. The transitions of the latter between the energy levels are responsible for the spectral lines. If no external fields are present, the energy levels depend on the angular momentum quantum number ℓ as well as the principal quantum number n. Due to the spherical symmetry, they do not depend on the x_3-component $\hbar m$ of the angular momentum. In other words, there is one energy level for each n, ℓ, and m. However, studies of the spectra revealed that (except for the s-level) all levels occur twice. For example, the so-called D-line of sodium, which is generated by the transition $3p \to 3s$, occurs twice with the wavelengths 589.6 and 589.0 nm. It is for this very reason that Pauli predicted a fourth quantum number for electrons in atoms that has two values (except for the s-level). In 1925, Uhlenbeck[3] and Goudsmit[4] proposed that the doubling of the energy level is created by an internal angular momentum of the electron. The component in the direction of the angular momentum $\boldsymbol{L}$ of the electron orbit around the nucleus can only assume two values, and its interaction with the weak magnetic field formed by the electrons that orbit around the nucleus, all except the s category, splits into two levels. All components of the inner angular momentum $\boldsymbol{S}$ can have $2\ell+1$ values, such that the size s corresponding to ℓ assumes the unusual value $\frac{1}{2}$. This inner angular momentum was called the *spin* of the electron. How is a spin of size $\hbar/2$ generated? For an angular momentum of this amount and the classical electron radius of $r_E = \alpha\hbar/m_e c$, a peripheral speed would be required that is more than 70 times the speed of light! From a classical perspective, the spin can therefore not be explained as an angular momentum. Let us approach the solution to this question via symmetry considerations.

8.3 Symmetry Considerations

Historically, classical mechanics supplied quantum mechanics with observable quantities and their properties [21]. In some cases, this contradicts Heisenberg's idea to use only *measurable*, that is, observable, quantities, for example, when analyzing the motion of an electron around a nucleus (which cannot be observed). However, many of the entities required in quantum mechanics are introduced solely by *symmetry considerations*.

---

[3] George Uhlenbeck, Dutch/American physicist, 1900–1988.

[4] Samuel Goudsmit, Dutch/American physicist, 1902–1978.

The *symmetry principle* states that a law of nature does not change, for example, if we change our viewpoint. In other words, if a law is valid here (locus $\boldsymbol{x}$), then the law is also valid there (at the location $\boldsymbol{x}+\boldsymbol{a}$).

The unity operator $\boldsymbol{I}$ represents a trivial symmetry that does not alter the state. Next, there is a special class of symmetries that can be represented by linear unitary operators $\boldsymbol{U}$ that are arbitrarily close to $\boldsymbol{I}$. Such operators can be written as a *Taylor series*

$$\boldsymbol{U}(\epsilon) \overset{\text{def}}{=} \boldsymbol{I} + \epsilon \cdot \left.\frac{\mathrm{d}\boldsymbol{U}}{\mathrm{d}\epsilon}\right|_{\epsilon=0} + \mathcal{O}(\epsilon^2), \tag{8.1}$$

where ϵ is arbitrarily small. In order for $\boldsymbol{U}$ to be unitary (i.e., $\boldsymbol{U}\boldsymbol{U}^\dagger = \boldsymbol{I}$), we must require

$$\boldsymbol{U}\boldsymbol{U}^\dagger = \left(\boldsymbol{I} + \epsilon \cdot \left.\frac{\mathrm{d}\boldsymbol{U}}{\mathrm{d}\epsilon}\right|_{\epsilon=0} + \mathcal{O}(\epsilon^2)\right)\left(\boldsymbol{I} + \epsilon \cdot \left.\frac{\mathrm{d}\boldsymbol{U}^\dagger}{\mathrm{d}\epsilon}\right|_{\epsilon=0} + \mathcal{O}(\epsilon^2)\right) \overset{!}{=} \boldsymbol{I}.$$

This leads to

$$\boldsymbol{I} + \epsilon \underbrace{\left.\left[\frac{\mathrm{d}\boldsymbol{U}}{\mathrm{d}\epsilon} + \frac{\mathrm{d}\boldsymbol{U}^\dagger}{\mathrm{d}\epsilon}\right]\right|_{\epsilon=0}}_{\overset{!}{=}\boldsymbol{0}} + \mathcal{O}(\epsilon^2) \overset{!}{=} \boldsymbol{I}.$$

Let us now introduce the notation

$$\frac{\mathrm{d}\boldsymbol{U}}{\mathrm{d}\epsilon} = i \cdot \boldsymbol{G}, \tag{8.2}$$

with the (symmetric) *generator*

$$\boldsymbol{G} = \boldsymbol{G}^\dagger. \tag{8.3}$$

Then we get indeed

$$\left[\frac{\mathrm{d}\boldsymbol{U}}{\mathrm{d}\epsilon} + \frac{\mathrm{d}\boldsymbol{U}^\dagger}{\mathrm{d}\epsilon}\right] = i \cdot \boldsymbol{G} - i \cdot \boldsymbol{G} = \boldsymbol{0}.$$

With the definition $\epsilon \overset{\text{def}}{=} \theta/N$, (8.1) now reads

$$\boldsymbol{U}(\theta/N) = \boldsymbol{I} + i \cdot \theta/N \cdot \boldsymbol{G} + \mathcal{O}(\epsilon^2).$$

Now let θ be fixed and N get large. Then we can obtain the transformation according to θ by applying N times the infinitesimal transformation according to θ/N:

$$\boldsymbol{U}(\theta) = (\boldsymbol{I} + i \cdot \theta/N \cdot \boldsymbol{G})^N.$$

For $N \to \infty$, this becomes

$$\underline{\underline{\boldsymbol{U}(\theta)}} = \lim_{N\to\infty} (\boldsymbol{I} + i \cdot \theta/N \cdot \boldsymbol{G})^N = \underline{\underline{e^{i\,\theta \boldsymbol{G}}}}. \tag{8.4}$$

Note that we left the higher-order terms $\mathcal{O}(\epsilon^2)$ aside in this rather heuristic derivation. However, it can be shown in a mathematically rigorous way that this is indeed a feasible definition for *any size* of θ ! In particular, the first terms of the series representation

$$e^{\boldsymbol{X}} = \sum_{m=0}^{\infty} \frac{\boldsymbol{X}^m}{m!},$$

with $\boldsymbol{X} = i\varepsilon\boldsymbol{G}$, indeed yield the desired result (8.1) for small ϵ.

It is well known from mathematics that a similarity transformation leaves eigenvalues and eigenvectors unchanged. Now we want to show that there is a unitary transformation such that a coordinate shift occurs according to

$$\boldsymbol{x} \Rightarrow \boldsymbol{x} + \boldsymbol{a}, \text{ for all } \boldsymbol{x},$$

where $\boldsymbol{a}$ is an arbitrary three-dimensional vector. In the case of the Heisenberg matrices $\boldsymbol{X}_j$ of the matrix vector

$$\boldsymbol{\mathfrak{X}} = \begin{pmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \\ \boldsymbol{X}_3 \end{pmatrix},$$

the equivalent condition is

$$\boldsymbol{X}_j \Rightarrow \boldsymbol{X}_j + a_j\,\boldsymbol{I}\,,\; j = 1, 2, 3,$$

or

$$\boldsymbol{\mathfrak{X}} \Rightarrow \boldsymbol{\mathfrak{X}} + \boldsymbol{a} \otimes \boldsymbol{I}.$$

In other words, we are looking for a unitary matrix $\boldsymbol{U}(\boldsymbol{a})$ such that

$$\boldsymbol{U}(\boldsymbol{a})\boldsymbol{X}_j\boldsymbol{U}(\boldsymbol{a})^\dagger = \boldsymbol{X}_j + a_j\,\boldsymbol{I}. \tag{8.5}$$

For infinitesimally small a_j, $\boldsymbol{U}$ must have a shape like that in (8.1). Therefore, it seems useful to start with the ansatz

$$\boldsymbol{U}(\boldsymbol{a}) = \boldsymbol{I} + \frac{i}{\hbar}\,(a_1\boldsymbol{P}_1 + a_2\boldsymbol{P}_2 + a_3\boldsymbol{P}_3) + \mathcal{O}(\boldsymbol{a}^2). \tag{8.6}$$

The dimension of $\hbar$ is $length^2 \cdot mass/time$. It turns out later that $\boldsymbol{P}$ represents the momentum (with dimension $length \cdot mass/time$) and that the a_i have the dimension $length$. It is therefore convenient to add the $\hbar$ in order to render $\boldsymbol{U}$ a dimensionless transformation matrix. For an infinitesimal three-vector $\boldsymbol{a}$ the condition (8.5) requires that

$$i[\boldsymbol{P}_j \cdot a, \boldsymbol{X}_j]/\hbar \overset{!}{=} a_j \boldsymbol{I}.$$

$\boldsymbol{P}_j$ commutes with itself, and of course also with any function of $\boldsymbol{P}_j$. Therefore, we get

$$\boldsymbol{U}(\boldsymbol{a})\boldsymbol{P}_j\boldsymbol{U}^\dagger(\boldsymbol{a}) = \boldsymbol{P}_j\boldsymbol{U}(\boldsymbol{a})\boldsymbol{U}^\dagger(\boldsymbol{a}) = \boldsymbol{P}_j. \tag{8.7}$$

In other words, $\boldsymbol{P}_j$ remains unchanged.

Does this transformation also shift the position vector $\boldsymbol{x}' = \boldsymbol{x} + \boldsymbol{a}$; that is, $\mathfrak{X}' = \mathfrak{X} + \boldsymbol{a} \otimes \boldsymbol{I}$? We have

$$\boldsymbol{X}'_j(\boldsymbol{a}) = \boldsymbol{U}(\boldsymbol{a})\boldsymbol{X}_j\boldsymbol{U}^\dagger(\boldsymbol{a}).$$

Let us now form the derivative with respect to a_k:

$$\frac{\partial}{\partial a_k}\boldsymbol{X}'_j(\boldsymbol{a}) = \left[\frac{\partial}{\partial a_k}\boldsymbol{U}(\boldsymbol{a})\right]\boldsymbol{X}_j\boldsymbol{U}^\dagger(\boldsymbol{a}) + \boldsymbol{U}(\boldsymbol{a})\boldsymbol{X}_j\frac{\partial}{\partial a_k}\boldsymbol{U}^\dagger(\boldsymbol{a}). \tag{8.8}$$

Differentiating (8.4) by a_k and exploiting the fact that all three $\boldsymbol{P}_j$ matrices commute with each other, one obtains

$$\frac{\partial}{\partial a_k}\boldsymbol{U}(\boldsymbol{a}) = \frac{i}{\hbar}\boldsymbol{U}(\boldsymbol{a})\boldsymbol{P}_k, \quad \frac{\partial}{\partial a_k}\boldsymbol{U}^\dagger(\boldsymbol{a}) = -\frac{i}{\hbar}\boldsymbol{P}_k\boldsymbol{U}^\dagger(\boldsymbol{a}).$$

Putting this result in (8.8) yields

$$\frac{\partial}{\partial a_k}\boldsymbol{X}'_j(\boldsymbol{a}) = -\frac{i}{\hbar}\boldsymbol{U}(\boldsymbol{a})\underbrace{\left[\boldsymbol{X}_j, \boldsymbol{P}_k\right]}_{i\,\hbar\,\delta_{jk}\boldsymbol{I}}\boldsymbol{U}^\dagger(\boldsymbol{a}) = \delta_{jk}\boldsymbol{I}, \tag{8.9}$$

where we used the known fact

$$\boldsymbol{X}_j\boldsymbol{P}_k - \boldsymbol{P}_k\boldsymbol{X}_j = \left[\boldsymbol{X}_j, \boldsymbol{P}_k\right] = i\,\hbar\,\delta_{jk}\boldsymbol{I}.$$

We can now integrate (8.9) to get

$$\boldsymbol{X}'_j(\boldsymbol{a}) = \boldsymbol{X}'_j(0) + a_j\boldsymbol{I}.$$

Because of $\boldsymbol{U}(0) = \boldsymbol{I}$, we have $\boldsymbol{X}'_j(0) = \boldsymbol{X}_j$. Therefore, we indeed find that

$$\mathfrak{X}'(\boldsymbol{a}) = \mathfrak{X} + \boldsymbol{a} \otimes \boldsymbol{I}.$$

The most exciting result of this whole consideration is that we found the previously called momentum matrix

$$\mathfrak{P} = \begin{pmatrix} \boldsymbol{P}_1 \\ \boldsymbol{P}_2 \\ \boldsymbol{P}_3 \end{pmatrix}$$

to be a generator of a symmetry operation, without having to use classical mechanics as a justification.

8.4 Symmetry and the Spin

Consider now an infinitesimal *rotation* of a physical system around the x_3-axis by an angle $\delta\varphi$. The components of a vector $\boldsymbol{v}$ then change according to

$$v_1' = v_1 + \delta\varphi v_2; \quad v_2' = v_2 - \delta\varphi v_1 \quad \text{und} \quad v_3' = v_3.$$

In vector form, this reads

$$\boldsymbol{v}' = \boldsymbol{v} - \boldsymbol{\omega} \times \boldsymbol{v}, \tag{8.10}$$

where $\boldsymbol{\omega}$ is defined as

$$\boldsymbol{\omega} \overset{\text{def}}{=} \begin{pmatrix} 0 \\ 0 \\ \delta\varphi \end{pmatrix}.$$

If we want this infinitesimal rotation to be performed by means of a generator $\boldsymbol{G}$ via a transformation $\boldsymbol{U}(\delta\varphi) = \boldsymbol{I} + i\boldsymbol{G}\delta\varphi + \mathcal{O}(\delta\varphi^2)$, we must start with

$$\boldsymbol{G} = \frac{1}{\hbar}(\boldsymbol{\omega} \cdot \boldsymbol{\mathfrak{J}}) = \frac{1}{\hbar}(\omega_1 \boldsymbol{J}_1 + \omega_2 \boldsymbol{J}_2 + \omega_3 \boldsymbol{J}_3). \tag{8.11}$$

It will turn out that $\boldsymbol{\mathfrak{J}}$ is the *total angular momentum*, a 3-vector matrix, consisting of the three matrices $\boldsymbol{J}_1$, $\boldsymbol{J}_2$, and $\boldsymbol{J}_3$. Applying the transformation to $\boldsymbol{V_j}$ yields

$$\boldsymbol{V_j}' = (\boldsymbol{I} - i\boldsymbol{G})\boldsymbol{V_j}(\boldsymbol{I} + i\boldsymbol{G}) = \boldsymbol{V_j} + i(\boldsymbol{V_j}\boldsymbol{G} - \boldsymbol{G}\boldsymbol{V_j}) + \mathcal{O}(\delta\varphi^2), \tag{8.12}$$

or with (8.11)

$$\boldsymbol{V_j}' - \boldsymbol{V_j} = i(\boldsymbol{V_j}\boldsymbol{G} - \boldsymbol{G}\boldsymbol{V_j}) = i[\boldsymbol{V_j}, \boldsymbol{G}] = \frac{i}{\hbar}[\boldsymbol{V_j}, \boldsymbol{\omega} \cdot \boldsymbol{\mathfrak{J}}]. \tag{8.13}$$

Multiplying (8.13) with a_1, a_2, and a_3, and summing up the three equations, we get

$$\boldsymbol{a} \cdot (\boldsymbol{\mathfrak{V}}' - \boldsymbol{\mathfrak{V}}) = \frac{1}{i\hbar}[\boldsymbol{a} \cdot \boldsymbol{\mathfrak{V}}, \boldsymbol{\omega} \cdot \boldsymbol{\mathfrak{J}}], \tag{8.14}$$

where

$$\boldsymbol{a} \cdot \boldsymbol{\mathfrak{V}} = a_1 \boldsymbol{V_1} + a_2 \boldsymbol{V_2} + a_3 \boldsymbol{V_3}$$

and

$$\boldsymbol{\omega} \cdot \boldsymbol{\mathfrak{J}} = \omega_1 \boldsymbol{J}_1 + \omega_2 \boldsymbol{J}_2 + \omega_3 \boldsymbol{J}_3.$$

From (8.10) follows

$$\frac{1}{i\hbar}[\boldsymbol{a}\cdot\boldsymbol{\mathfrak{V}},\boldsymbol{\omega}\cdot\boldsymbol{\mathfrak{J}}]=\boldsymbol{a}\cdot(\boldsymbol{\omega}\times\boldsymbol{\mathfrak{V}}). \tag{8.15}$$

In compliance with (see also Problem 6.2)

$$\boldsymbol{a}\cdot(\boldsymbol{b}\times\boldsymbol{c})=(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{c},$$

we get the relationship

$$\frac{1}{i\hbar}[\boldsymbol{a}\cdot\boldsymbol{\mathfrak{V}},\boldsymbol{\omega}\cdot\boldsymbol{\mathfrak{J}}]=\boldsymbol{a}\cdot(\boldsymbol{\omega}\times\boldsymbol{\mathfrak{V}})=(\boldsymbol{a}\times\boldsymbol{\omega})\cdot\boldsymbol{\mathfrak{V}}. \tag{8.16}$$

If $\boldsymbol{a}$ and $\boldsymbol{\omega}$ are parallel to each other, it follows that

$$[\boldsymbol{V}_j,\boldsymbol{J}_j]=\boldsymbol{0},\ \ j=1,2\text{ or }3. \tag{8.17}$$

If $\boldsymbol{a}$ and $\boldsymbol{\omega}$ are perpendicular, however, it follows that

$$\frac{1}{i\hbar}[\boldsymbol{V}_j,\boldsymbol{J}_k]=\begin{cases}+\boldsymbol{V}_\ell & \text{if } j,k,\ell \text{are cyclic permutations of } 1,2,3,\\ -\boldsymbol{V}_\ell & \text{if } j,k,\ell \text{ are anti-cyclic permutations.}\end{cases} \tag{8.18}$$

We can summarize this as[5]

$$\boxed{[\boldsymbol{V}_j,\boldsymbol{J}_k]=i\hbar\sum_\ell \epsilon_{jk\ell}\boldsymbol{V}_\ell}. \tag{8.19}$$

The same is true for $\boldsymbol{V}_j=\boldsymbol{J}_j=\boldsymbol{L}_j$, thus

$$\boxed{[\boldsymbol{L}_j,\boldsymbol{L}_k]=i\hbar\sum_\ell \epsilon_{jk\ell}\boldsymbol{L}_\ell}. \tag{8.20}$$

Now we are ready to turn to the spin. For that purpose, let us define a vector matrix $\boldsymbol{\mathfrak{S}}$ according to

$$\boldsymbol{\mathfrak{S}}\overset{\text{def}}{=}\boldsymbol{\mathfrak{J}}-\boldsymbol{\mathfrak{L}};$$

---

[5] The Levi-Civita symbol $\varepsilon_{jk\ell}$ is defined as

$$\varepsilon_{jk\ell}=\begin{cases}+1 & \text{if } (j,k,\ell) \text{ is } (1,2,3),\ (2,3,1) \text{ or } (3,1,2),\\ -1 & \text{if } (j,k,\ell) \text{ is } (3,2,1),\ (1,3,2) \text{ or } (2,1,3),\\ \ \ 0 & \text{if } j=k \text{ or } k=\ell \text{ or } \ell=j\end{cases}$$

that is,

$$\mathfrak{J} = \mathfrak{L} + \mathfrak{S}. \tag{8.21}$$

By subtracting (8.20) from (8.19), after substituting $\boldsymbol{L}_j$ for $\boldsymbol{V}_j$ in (8.19), we find

$$[\boldsymbol{S_i}, \boldsymbol{L_j}] = \boldsymbol{0}. \tag{8.22}$$

Together with the other equations above, we obtain

$$\boxed{[\boldsymbol{S_j}, \boldsymbol{S_k}] = i\hbar \sum_{\ell} \epsilon_{jk\ell} \boldsymbol{S_\ell}}. \tag{8.23}$$

In other words, $\mathfrak{S}$ also behaves as an angular momentum and is called the *spin*. The spin angular momentum $\mathfrak{S}$ of a closed system, therefore, is the proportion of the total angular momentum $\mathfrak{J}$ that *is not* due to an orbital angular momentum $\mathfrak{L}$.

Goudsmith and Uhlenbeck suggested in 1925 to assign such an additional angular eigenmomentum called spin to the electron. It must have a half-integral angular momentum quantum number $s = \frac{1}{2}$ in order for the magnetic spin quantum number to be limited to two possible values $m_s = \pm\frac{1}{2}$. The spin accounts for a twofold or, together with an orbital angular momentum $\ell \geq 1$, a higher splitting of the spectrum.

8.5 Spin-$\frac{1}{2}$ Systems and Spinors

We can obtain the matrices representing $\boldsymbol{S_1}$, $\boldsymbol{S_2}$, and $\boldsymbol{S_3}$ from (6.37) and (6.54):

$$\boldsymbol{S_1} \overset{\text{def}}{=} \boldsymbol{L}_1^{(1/2)} = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \boldsymbol{S_2} \overset{\text{def}}{=} \boldsymbol{L}_2^{(1/2)} = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix},$$

$$\text{and} \quad \boldsymbol{S_3} \overset{\text{def}}{=} \boldsymbol{L}_3^{(1/2)} = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{8.24}$$

Using the *Pauli matrices*[6] $\boldsymbol{\sigma_i}$, we can write

$$\boldsymbol{S}_i = \frac{\hbar}{2}\boldsymbol{\sigma_i}, \tag{8.25}$$

where the Pauli matrices have the form

$$\boldsymbol{\sigma}_1 \overset{\text{def}}{=} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \boldsymbol{\sigma}_2 \overset{\text{def}}{=} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \text{and} \quad \boldsymbol{\sigma}_2 \overset{\text{def}}{=} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{8.26}$$

---

[6] We stick to the traditional notation for these matrices.

The eigenvalues of these matrices are ±1, and we have

$$\boldsymbol{\sigma}_1^2 = \boldsymbol{\sigma}_2^2 = \boldsymbol{\sigma}_3^2 = \boldsymbol{I}_2. \tag{8.27}$$

Pauli was the first to recognize the need for two-component state vectors for describing the observed properties of atomic spectra.

The spin-(1/2) matrices $\boldsymbol{S_1}$, $\boldsymbol{S_2}$, and $\boldsymbol{S_3}$ have eigenvalues and two-dimensional, normalized eigenvectors as follows.

$$\boldsymbol{S_1} = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \text{ has eigenvalues } \pm\frac{\hbar}{2} \text{ and eigenvectors } \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ \pm 1 \end{pmatrix}$$

,

$$\boldsymbol{S_2} = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \text{ has eigenvalues } \pm\frac{\hbar}{2} \text{ and eigenvectors } \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ \pm i \end{pmatrix}$$

,

$$\boldsymbol{S_3} = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \text{ has eigenvalues } \pm\frac{\hbar}{2} \text{ and eigenvectors } \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

.

8.6 Adding Angular Momenta

8.6.1 Clebsch–Gordan Coefficients

We now consider a system with two sources of angular momenta, which we call

$$\mathfrak{J}' = \begin{pmatrix} \boldsymbol{J}_1' \\ \boldsymbol{J}_2' \\ \boldsymbol{J}_3' \end{pmatrix} \text{ and } \mathfrak{J}'' = \begin{pmatrix} \boldsymbol{J}_1'' \\ \boldsymbol{J}_2'' \\ \boldsymbol{J}_3'' \end{pmatrix}.$$

The system may be a single particle with a spin and an angular momentum, or may consist of two particles with either spin or angular momentum. What are the commutator rules for the total angular momentum? As for all angular momenta, the following eigenvalue equations apply.

$$\boldsymbol{J}'^{\boldsymbol{2}}\boldsymbol{e}' = j'(j'+1)\hbar^2\boldsymbol{e}', \tag{8.28}$$

$$\boldsymbol{J}_{\boldsymbol{3}}'\boldsymbol{e}' = m'\hbar\boldsymbol{e}', \tag{8.29}$$

and

$$\boldsymbol{J}''^2\boldsymbol{e}'' = j''(j''+1)\hbar^2\boldsymbol{e}'', \tag{8.30}$$

$$\boldsymbol{J}_3''\boldsymbol{e}'' = m''\hbar\boldsymbol{e}'', \tag{8.31}$$

where

$$m' = j', j'-1, \dots, -j'; \quad \text{und} \quad m'' = j'', j''-1, \dots, -j''.$$

For independent angular momentum sources, we certainly have

$$[\boldsymbol{J}_k', \boldsymbol{J}_\ell''] = \boldsymbol{0} \quad \text{for all } k, \ell \in \{1,2,3\}, \text{and} \quad [\boldsymbol{J}'^2, \boldsymbol{J}''^2] = \boldsymbol{0}. \tag{8.32}$$

However,

$$[\boldsymbol{J}_j', \boldsymbol{J}_k'] = i\hbar\epsilon_{jk\ell}\boldsymbol{J}_\ell', \tag{8.33}$$

and

$$[\boldsymbol{J}_j'', \boldsymbol{J}_k''] = i\hbar\epsilon_{jk\ell}\boldsymbol{J}_\ell''. \tag{8.34}$$

Is the sum

$$\mathfrak{J} = \begin{pmatrix} \boldsymbol{J}_1 \\ \boldsymbol{J}_2 \\ \boldsymbol{J}_3 \end{pmatrix} \stackrel{\text{def}}{=} \mathfrak{J}' + \mathfrak{J}''$$

also an angular momentum that satisfies the usual commutation rules? In fact, we get

$$\begin{aligned} [\boldsymbol{J}_1, \boldsymbol{J}_2] &= [\boldsymbol{J}_1' + \boldsymbol{J}_1'', \boldsymbol{J}_2' + \boldsymbol{J}_2''] = \\ &= [\boldsymbol{J}_1', \boldsymbol{J}_2'] + [\boldsymbol{J}_1'', \boldsymbol{J}_2''] + [\boldsymbol{J}_1', \boldsymbol{J}_2''] + [\boldsymbol{J}_1'', \boldsymbol{J}_2'] = \\ &= i\hbar\boldsymbol{J}_3' + i\hbar\boldsymbol{J}_3'' + \boldsymbol{0} + \boldsymbol{0} = \\ &= i\hbar\boldsymbol{J}_3. \end{aligned} \tag{8.35}$$

The same applies to all cyclic permutations of the indices, therefore $\mathfrak{J} = \mathfrak{J}' + \mathfrak{J}''$ is indeed an angular momentum. Its quantum numbers j and m_j can have the following values. Generally we know, that for given j,

$$-j \le m \le j \quad \text{and} \quad m_{max} = j.$$

Because $m = m' + m''$, the maximum value of m for all j is $j' + j''$; that is, the maximum value of $m' + m''$. This must be the maximum value of j also, for otherwise there would be a larger value of $m' + m''$. Thus

$$j_{max} = j' + j''.$$

With the minimum value of $j_{min} = |j' - j''|$, given at the end of this subsection, we obtain

$$|j' - j''| \leq j \leq |j' + j''|,$$

$$m_j = -j, -j+1, \ldots, j.$$

Summing up, $\mathfrak{J}$ is an angular momentum with the length $\sqrt{j(j+1)}\hbar$, where j can be an integer or half-integer, and the x_3-component has the value $m_j\hbar$ (with $m_j = j, j-1, \ldots, -j$).

Each angular momentum has its own eigenspace that is spanned by the eigenvectors $\boldsymbol{e}'_i$ and $\boldsymbol{e}''_i$, respectively. In the basis of the eigenvectors $\boldsymbol{e}'_i$, $\boldsymbol{J}'_j$ has a simple diagonal form. The same applies to $\boldsymbol{J}''_j$. The eigenvectors of the overall system can be assembled of eigenvectors for the subsystems (see Chap. 10) according to

$$\boldsymbol{e}_i = \boldsymbol{e}'_j \otimes \boldsymbol{e}''_k.$$

However, these vectors are generally not eigenvectors to

$$\boldsymbol{J}^2 = \boldsymbol{J}_1^2 + \boldsymbol{J}_2^2 + \boldsymbol{J}_3^2,$$

As a consequence, this matrix is not diagonal in this basis. Therefore, one better shifts from the complete set of commuting matrices $\boldsymbol{J'}^2$, $\boldsymbol{J'}_3$, $\boldsymbol{J''}^2$, $\boldsymbol{J''}_3$ with the eigenstates $\boldsymbol{e}'_j \otimes \boldsymbol{e}''_k$ to the complete set of commuting matrices $\boldsymbol{J}^2$, $\boldsymbol{J}_3$, $\boldsymbol{J'}^2$, $\boldsymbol{J''}^2$, with the eigenvectors $\boldsymbol{e}(j, m_j, j', j'')$. In this new basis, all four matrices $\boldsymbol{J}^2$, $\boldsymbol{J}_3$, $\boldsymbol{J'}^2$, and $\boldsymbol{J''}^2$ can be diagonalized simultaneously. The new eigenvectors satisfy the following eigenvalue equations:

$$\boldsymbol{J}^2\boldsymbol{e}(j, m_j, j', j'') = \hbar^2 j(j+1)\, \boldsymbol{e}(j, m_j, j', j''),$$

$$\boldsymbol{J}_3\boldsymbol{e}(j, m_j, j', j'') = \hbar m_j\, \boldsymbol{e}(j, m_j, j', j'').$$

How can we obtain the new basis vectors $\boldsymbol{e}(j, m_j, j', j'')$ if the basis vectors $\boldsymbol{e}'_j \otimes \boldsymbol{e}''_k$ are given? The elements of one basis must be a linear combination of the elements of the other basis. In the basis with the basis vectors $\boldsymbol{e}'_j \otimes \boldsymbol{e}''_k$, the new basis vector $\boldsymbol{e}(j, m_j, j', j'')$ looks like

$$\boldsymbol{e}(j, m_j, j', j'') = \sum_{a,b} C(a, b;\, j, m_j, j', j'')\boldsymbol{e}'_a \otimes \boldsymbol{e}''_b. \tag{8.36}$$

The coefficients $C(a, b;\, j, m_j, j', j'')$ can be calculated by multiplying (8.36) from the left with the transposed basis vector $\boldsymbol{e'}_a \otimes \boldsymbol{e''}_b$ (remember that the basis vectors $\boldsymbol{e}'_a \otimes \boldsymbol{e}''_b$ are orthogonal). These coefficients are called *Clebsch*[7]*-Gordan*[8]*-coefficients*

---

[7] Rudolf Friedrich Alfred Clebsch, 1833–1872, German mathematician.

[8] Paul Albert Gordan, 1837–1912, German mathematician.

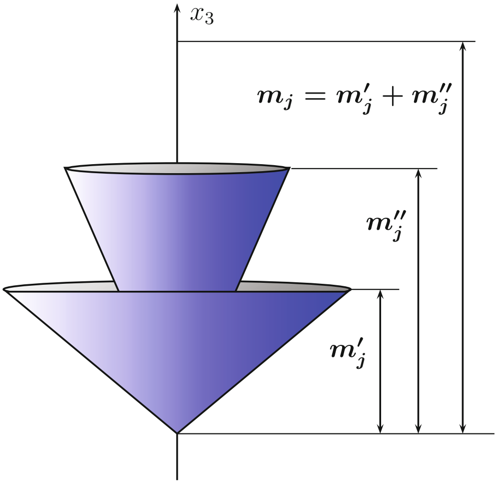


Fig. 8.2 The relationship $m_j = m'_j + m''_j$

and are defined by

$$\boxed{C(a, b; j, m_j, j', j'') = (\boldsymbol{e}'_a \otimes \boldsymbol{e}''_b)^\dagger \boldsymbol{e}(j, m_j, j', j'')}. \tag{8.37}$$

Now which values j can exist for the system for a given j' and j''? Because $\boldsymbol{J}^2$ commutes with its own components, it commutes in particular with $\boldsymbol{J_3 = J'_3 + J''_3}$. This shows us that we can specify the value of m_j and j at the same time. The allowed values for m_j follow immediately from the relationship

$$\boldsymbol{J_3 = J'_3 + J''_3};$$

that is (see Fig. 8.2),

$$m_j = m'_j + m''_j. \tag{8.38}$$

To determine the allowed values of j, we first note that the complete number of possible values for the uncoupled systems is

$$(2j' + 1)(2j'' + 1) = 4j'j'' + 2j' + 2j'' + 1.$$

There is only one state in which both components reach their maximum value, namely $m'_j = j'$ and $m''_j = j''$ (which entails $m_j = j' + j''$). However, the maximum value of m_j is j by definition, therefore the maximum value of j is $j = j' + j''$. Note that the Clebsch–Gordan coefficients only differ from zero if $|j' - j''| \leq j \leq j' + j''$ and $m_j = m' + m''$.

8.6.2 Clebsch–Gordan Coefficients on the Internet

You can find the following formula for the calculation of the Clebsch–Gordan coefficients on the Internet:

$$C(m', m''; j, m_j, j', j'') =$$

$$= \delta_{m_j, m'+m''} \left(\frac{(2j+1)(j'+j''-j)!(j'-j''+j)!(j+j''-j')!}{(j'+j''+j+1)!} \right)^{\frac{1}{2}} \times$$

$$\sum_n \left(\frac{(-1)^n [(j'+m'')!(j'-m')!(j'+m'')!(j''-m'')!(j+m_j)!(j-m_j)!]^{\frac{1}{2}}}{n!(j'+j''-j-n)!(j'-m'-n)!(j''+m''-n)!(j-j''+m'+n)!(j-j'-m''+n)!} \right).$$

The sum runs over all n such that the factorials are always well-defined and positive. The formula is available at
http://www.spektrum.de/lexikon/physik/clebsch-gordan-koeffizienten/2438.
A downloadable PDF table of Clebsch–Gordan coefficients can be found under
http://pdg.lbl.gov/2011/reviews/rpp2011-rev-clebsch-gordan-coefs.pdf.
You can also download a calculator under http://www.volya.net/index.php.

8.7 Spin-Orbit Coupling

This stands for the observation in atomic, nuclear, and elementary particle physics that a particle's direction of the spin with respect to its orbital angular momentum affects the energy of the particle. Good examples are electrons in the atomic shell, where the spin-orbit coupling leads to the splitting of the spectral lines and thus contributes to the fine structure of atomic spectra.

It is easiest to describe spin-orbit coupling of the electrons in a semi-classical model. Maxwell's theory and the special theory of relativity require that a magnetic field act on an electron when it circulates in the electric field of the atomic nucleus. After all, a circular motion of the nucleus is perceived in the electron's frame of reference. Due to the charge of the nucleus, this movement represents a circulating current that generates a magnetic field parallel to the angular momentum vector. Because the electron with its intrinsic angular momentum (spin) has a magnetic moment in its resting frame of reference as well, the spin direction parallel to the field yields a lower energy and vice versa. Hence, a single energy level is split into two levels, and we find two optical spectra that are slightly shifted compared to the original lines. The total angular momentum of the electron is a combination of the spin and the orbital part:

$$\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}.$$

The usual quantum mechanical properties must be valid:

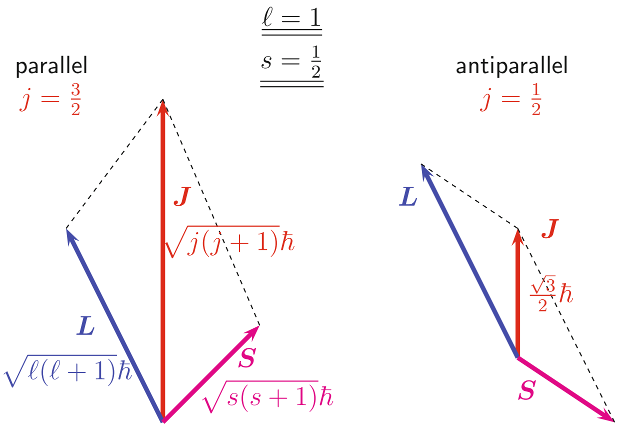


Fig. 8.3 Spin-orbit coupling

$$\boldsymbol{J}^2 \boldsymbol{e}_{jm_j} = j(j+1)\hbar^2 \boldsymbol{e}_{jm_j},$$

$$\boldsymbol{J}_3 \boldsymbol{e}_{jm_j} = m_j \hbar \boldsymbol{e}_{jm_j},$$

$$-j \leq m_j \leq j,$$

where $\boldsymbol{e}_{jm_j}$ is a common eigenvector of $\boldsymbol{J}^2$ and $\boldsymbol{J}_3$.

As we already found out, in the hydrogen atom the total spin quantum number can have two values (depending on the coupling):

$$j = \ell + \frac{1}{2}(\text{parallel}).$$

$$j = \ell - \frac{1}{2}(\text{antiparallel}).$$

Figure 8.3 shows this result for $\ell = 1$ in vector form.

The following section gives a relatively simple description of the spin-orbit interaction for an electron that is bound to an atom. We use semi-classical electrodynamics and nonrelativistic quantum mechanics. The results agree quite well with the observations.

Energy of the Magnetic Dipole Moment

The energy of a magnetic dipole moment in a magnetic field is given by

$$\Delta E = -\boldsymbol{\mu} \cdot \mathfrak{b}$$

where $\boldsymbol{\mu}$ is the magnetic dipole moment of the particle and $\mathfrak{b}$ is the magnetic induction of the magnetic field.

Magnetic Field

Let us first consider the magnetic field. An electron with mass m_e and charge e that moves with velocity $\boldsymbol{v}$ in an electric field $\boldsymbol{\mathfrak{e}}$ generates the magnetic field

$$\boldsymbol{\mathfrak{b}} = \frac{\boldsymbol{\mathfrak{e}} \times \boldsymbol{v}}{c^2}.$$

If the electric field is generated by the potential $U(r)$, we have

$$\boldsymbol{\mathfrak{e}} = -\frac{\boldsymbol{r}}{r}\frac{\partial U(r)}{\partial r}.$$

The unit vector $\boldsymbol{r}/r$ implies a radial direction of the electric field. With this, we get

$$\boldsymbol{\mathfrak{b}} = -\frac{1}{rc^2}\frac{\partial U(r)}{\partial r}\boldsymbol{r} \times \boldsymbol{v}.$$

Therefore, if $\boldsymbol{r}$ and $\boldsymbol{v}$ are in the x_1, x_2 plane, $\boldsymbol{\mathfrak{b}}$ points in x_3-direction. Now remember the angular momentum of a particle,

$$\boldsymbol{\ell} = \boldsymbol{r} \times \boldsymbol{p},$$

and the relation $\boldsymbol{p} = m_e \boldsymbol{v}$. Then

$$\boldsymbol{\mathfrak{b}} = \frac{-1}{m_e c^2 r}\frac{\partial U(r)}{\partial r}\boldsymbol{\ell}. \tag{8.39}$$

In other words, the magnetic field $\boldsymbol{\mathfrak{b}}$ is parallel to the orbital angular momentum $\boldsymbol{\ell}$ of the particle.

Magnetic Dipole Moment of the Electron

The magnetic dipole moment of the electron is given by

$$\boldsymbol{\mu} = -\frac{g_s \mu_B}{\hbar}\boldsymbol{s} = -\frac{g_s e}{2m_e}\boldsymbol{s}, \tag{8.40}$$

where $\boldsymbol{s}$ is the spin angular momentum vector,

$$\mu_B \overset{\text{def}}{=} \frac{e\hbar}{2m_e} = 9.274 \cdot 10^{-24}\frac{J}{T}$$

is the Bohr magneton, and $g_S = 2,002319304 \approx 2$ is the electron spin g-factor[9] We see that the magnetic dipole moment is anti-parallel to the spin angular momentum.

---

[9] The g-factor (Lande factor) is the ratio of the measured magnetic moment to the magnetic moment, which we would expect from classical physics. It comes into play when considering quantum electrodynamics (Dirac).

The potential due to the spin orbit interaction consists of two parts. The Larmor[10]-part has to do with the interaction of the magnetic moment of the electron with the magnetic field of the nucleus in the coordinate system of the electron. The second contribution is related to the so-called Thomas[11] precession.

In a homogeneous magnetic field with the magnetic induction density $\mathfrak{b}$, a particle with a magnetic dipole moment $\boldsymbol{\mu}$ experiences the angular momentum $\boldsymbol{m} = \boldsymbol{\mu} \times \mathfrak{b}$. The work required for the rotation of a magnetic dipole by an angle $\mathrm{d}\varphi$ is given by

$$\mathrm{d}W = -|\boldsymbol{m}| \cdot \mathrm{d}\varphi = -|\boldsymbol{\mu}| \cdot |\mathfrak{b}| \sin\varphi \cdot \mathrm{d}\varphi,$$

where φ is the angle between $\boldsymbol{\mu}$ and $\mathfrak{b}$ (the minus sign is because φ decreases by the action of the angular momentum $\boldsymbol{m}$). This work equals the decrease of potential energy E_{pot} of the system; that is,

$$\mathrm{d}E_{pot} = -\mathrm{d}W = |\boldsymbol{\mu}| \cdot |\mathfrak{b}| \sin\varphi \cdot \mathrm{d}\varphi.$$

Integrating the equation, we get

$$E_{pot} = -|\boldsymbol{\mu}| \cdot |\mathfrak{b}| \cos\varphi = -\boldsymbol{\mu} \cdot \mathfrak{b}.$$

This contribution is also called *Larmor* interaction energy. Substituting the expressions for the magnetic moment and the magnetic field leads to

$$\Delta E_L = -\boldsymbol{\mu} \cdot \mathfrak{b} = \frac{2\mu_B}{\hbar m_e e c^2} \frac{1}{r} \frac{\partial U(r)}{\partial r} \boldsymbol{\ell} \cdot \boldsymbol{s}.$$

Next, we have to consider the *Thomas precession correction* that is caused by the curved path of the electron. In 1926, Llewellyn Thomas computed the splitting in the fine structure of the atomic spectrum in a relativistic framework. The Thomas precession rate, $\boldsymbol{\Omega}_T$, is related to the angular frequency of the orbital motion ω of a spinning particle as follows.

$$\boldsymbol{\Omega}_T = \boldsymbol{\omega}(1 - \gamma),$$

where $\gamma = 1/\sqrt{1 - (v/c)^2}$ is the Lorentz factor of the moving particle. The Hamiltonian that causes the spin precession $\boldsymbol{\Omega}_T$ is given by[12]

$$\Delta E_T = \boldsymbol{\Omega}_T \cdot \boldsymbol{S}.$$

In the first order of $(v/c)^2$, this yields

---

[10] Joseph Larmor, 1857–1942, Irish physicist.

[11] Llewellyn Thomas, 1903–1992, British physicist.

[12] Note that we now moved back from the matrix vectors to operators.

$$\Delta E_T = -\frac{\mu_B}{\hbar m_e e c^2}\frac{1}{r}\frac{\partial U(r)}{\partial r}\boldsymbol{L}\cdot\boldsymbol{S}.$$

Total Interaction Energy

The total spin-orbit potential in an external electrostatic potential therefore has the form

$$\Delta E \equiv \Delta E_L + \Delta E_T = \frac{\mu_B}{\hbar m_e e c^2}\frac{1}{r}\frac{\partial U(r)}{\partial r}(\boldsymbol{L}\cdot\boldsymbol{S}).$$

Note that the net effect of the Thomas precession is a reduction of the Larmor interaction energy by a factor $\frac{1}{2}$, which became famous as the “Thomas half”.

Evaluation of the Energy Shift

With all the above approximations, we can now determine the energy shift in this model. In particular, we want to find a vector basis that diagonalizes both H_0 (the nonperturbed Hamilton function) and H. For this purpose, we first define the total angular momentum matrix (operator)

$$\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}.$$

The scalar product of $\boldsymbol{J}$ with itself yields

$$\boldsymbol{J^2 = L^2 + S^2} + 2\boldsymbol{L}\cdot\boldsymbol{S},$$

and therefore

$$\boldsymbol{L}\cdot\boldsymbol{S} = \frac{1}{2}(\boldsymbol{J}^2 - \boldsymbol{L}^2 - \boldsymbol{S}^2).$$

It is easy to show that the five matrices $\boldsymbol{H_0}$, $\boldsymbol{J^2}$, $\boldsymbol{L^2}$, $\boldsymbol{S^2}$, and $\boldsymbol{J_z}$ all commute with each other as well as with $\boldsymbol{H}$. Therefore, the basis that we are looking for is also a simultaneous eigenbasis of these five matrices (i.e., a basis in which all five matrices are diagonal). Eigenvectors of this basis have five quantum numbers:

- n (the “principal quantum number”)
- j (the “total angular momentum quantum number”)
- ℓ (the “orbital angular momentum quantum number”)
- s (the “spin quantum number”)
- j_3 (the “x_3-component of total angular momentum”)

The expectation value of $\boldsymbol{L}\cdot\boldsymbol{S}$ is

$$\langle \boldsymbol{L}\cdot\boldsymbol{S}\rangle = \frac{1}{2}(\langle \boldsymbol{J}^2\rangle - \langle \boldsymbol{L}^2\rangle - \langle \boldsymbol{S}^2\rangle) = \frac{\hbar^2}{2}(j(j+1) - \ell(\ell+1) - s(s+1)).$$

To calculate the final energy shift, we can say that

$$\Delta E = \frac{\beta}{2}(j(j+1) - \ell(\ell+1) - s(s+1)),$$

where

$$\beta = \frac{-\mu_B}{m_e\,ec^2}\left\langle \frac{1}{r}\frac{\partial U(r)}{\partial r}\right\rangle.$$

For a nucleus with charge Ze, the Coulomb potential is

$$U(r) = \frac{Ze}{4\pi\epsilon_0 r}$$

Therefore,

$$\frac{\partial U}{\partial r} = \frac{Ze}{4\pi\epsilon_0}\left(\frac{\partial 1/r}{\partial r}\right) = \frac{-Ze}{4\pi\epsilon_0 r^2}. \tag{8.41}$$

For hydrogen, the expectation value for $\frac{1}{r^3}$ is given by (see, e.g., [1] p. 251)

$$\left\langle \frac{1}{r^3}\right\rangle = \frac{2}{a^3 n^3 \ell(\ell+1)(2\ell+1)},$$

where $a = \hbar/Z\alpha m_e c$ is the Bohr radius divided by the nuclear charge Ze. For hydrogen, we therefore get the result

$$\beta(n,l) = \frac{\mu_0}{4\pi} g_s \mu_B^2 \frac{1}{n^3 a_0^3 \ell(\ell+1/2)(\ell+1)}.$$

For any hydrogenlike atom with Z protons, one can show that

$$\boxed{\beta(n,l) = Z^4 \frac{\mu_0}{4\pi} g_s \mu_B^2 \frac{1}{n^3 a_0^3 \ell(\ell+1)(2\ell+1)}}. \tag{8.42}$$

This leads to an energy shift for the individual energy levels as

$$\boxed{\Delta E = Z^4\,\frac{\mu_0 g_s \mu_B^2}{8\pi a_0^3}\,\frac{j(j+1) - \ell(\ell+1) - s(s+1)}{n^3 \ell(\ell+1)(2\ell+1)}}. \tag{8.43}$$

In general, the magnitude of this energy shift is $10^{-4} eV$. Note that the splitting due to spin-orbit coupling increases with increasing atomic number (namely as Z^4). Inasmuch as spin-orbit coupling is possible only for $\ell \geq 1$, the presence of ℓ in the denominator causes no issues.

For an electron in the p-shell, for example, the difference between the energies $j = \frac{3}{2}$ and $j = \frac{1}{2}$ is only

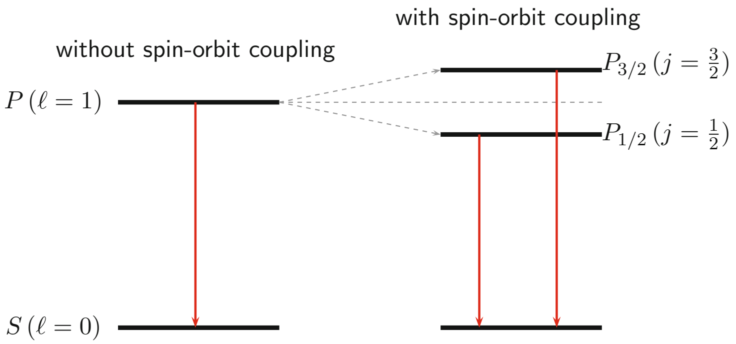


Fig. 8.4 Splitting of spectral lines due to spin-orbit coupling. Photons can be emitted at two different energies

$$Z^4\,\frac{\mu_0 g_s \mu_B^2}{8\pi a_0^3 n^3},$$

which is negligible with increasing n. For an electron in the p-shell (i.e., for $\ell = 1$), j can only assume the two values $1 + \frac{1}{2} = \frac{3}{2}$ and $1 - \frac{1}{2} = \frac{1}{2}$. For $j = \frac{3}{2}$, we find $2j + 1 = 4$ degenerate states with an energy that is proportional to $j(j+1) - \ell(\ell + 1) - s(s+1) = 1$. For $j = \frac{1}{2}$, there are $2j + 1 = 2$ degenerate states with an energy of $j(j+1) - \ell(\ell+1) - s(s+1) = -2$. Due to $4 \cdot 1 + 2 \cdot (-2) = 0$, the center of the distribution of energy levels is unchanged (see Fig. 8.4). Also, the center of energy levels that are split due to spin-orbit coupling is always the same as without splitting, because the interference comes from within the atom (and not from the environment).

8.8 Problems

8.1 Exponential Function of a Pauli Matrix: Evaluate the exponential function of the Pauli matrix $\boldsymbol{\sigma}_1$.

8.2 Eigenvalues and Eigenvectors of the Pauli Matrices: Evaluate the eigenvalues and eigenvectors of the Pauli matrices $\boldsymbol{\sigma}_1$, $\boldsymbol{\sigma}_2$ and $\boldsymbol{\sigma}_3$.

8.3 Transformation of the Pauli Matrices to Diagonal Form: How can you transform the Pauli matrices to diagonal form?

8.4 Clebsch–Gordan Coefficients: What are the Clebsch–Gordan coefficients of two spin-$\frac{1}{2}$ particles?

8.5 Clebsch–Gordan Coefficients: The orbital angular momentum $\boldsymbol{L}$ and the spin $\boldsymbol{S}$ of an electron are coupled to the total angular momentum $\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}$. What are the coupled states and the Clebsch–Gordan coefficients?

Chapter 9
Atoms in Electromagnetic Fields

Abstract We examine once again how external magnetic and electric fields affect the energy levels and thus the spectra of atoms. Now, however, we take the spin into account. The Zeeman effect reflects the response to magnetic fields.

9.1 Normal Zeeman Effect

If we apply an external magnetic field to an atom, it will interact with the magnetic dipole moment of the atom. The atomic magnetic dipole moment contains contributions from both the orbital and the spin angular momenta of the electron. The normal Zeeman effect (without spin) has already been discussed in Sect. 8.2. Therefore, two pictures (namely Figs. 9.1 and 9.2) should suffice to remind you of the results.

9.2 Anomalous Zeeman Effect

9.2.1 Weak Field Limit

In fact, the anomalous Zeeman effect is much more common than the normal Zeeman effect. Here, however, the spectral lines are split into more than three lines in a very complicated way, often an even number (quartet, sextet, etc.). To explain this effect, we must consider the spin. This angular momentum $\boldsymbol{s}$ of the electron that amounts to $\frac{1}{2}\hbar$ cannot be explained by classical physics. Although it is only half as large as the unit $\hbar$ of the orbital angular momentum, the spin carries the same strength in terms of the magnetic effect (1 Bohr magneton). In the anomalous Zeeman effect, we thus have to include both the orbital and the spin magnetism.

Let us now calculate the complete magnetic moment. The Hamiltonian for the interaction of the magnetic field $\mathfrak{b}$ with the orbital and spin angular momenta $\boldsymbol{L}$ and $\boldsymbol{S}$ is

$$\boldsymbol{H} = -\boldsymbol{\mu}_L \cdot \mathfrak{b} - \boldsymbol{\mu}_S \cdot \mathfrak{b} = -\gamma(\boldsymbol{L} + 2\boldsymbol{S}) \cdot \mathfrak{b}, \tag{9.1}$$

G. Ludyk, *Quantum Mechanics in Matrix Form*,
https://doi.org/10.1007/978-3-319-26366-3_9

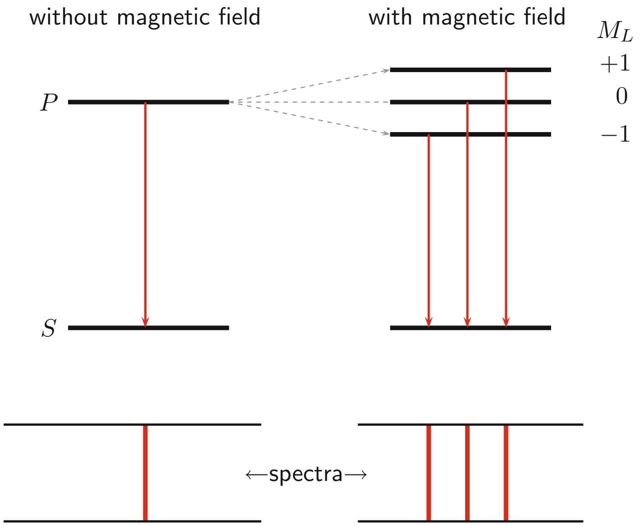


Fig. 9.1 Splitting of spectral lines in the normal Zeeman effect. Photons can be emitted with three different energies

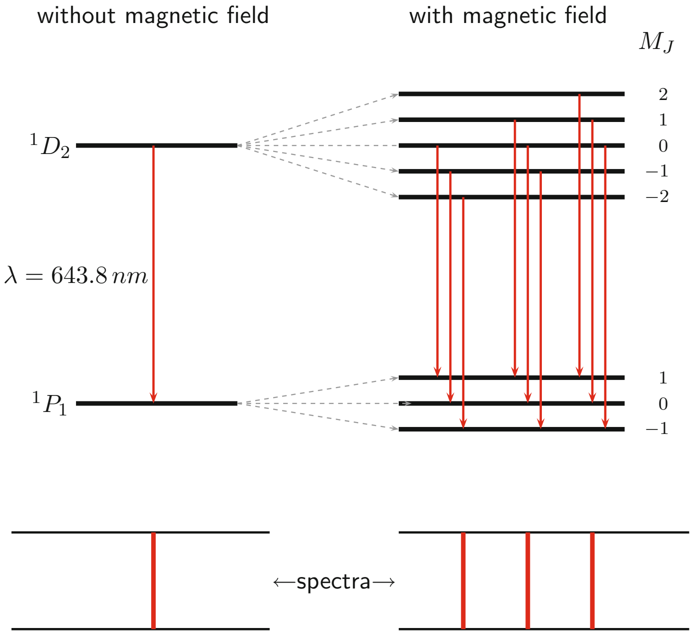


Fig. 9.2 Splitting of spectral lines due to the normal Zeeman effect in cadmium. There are three line triplets (with almost similar wavelengths within the triplet)

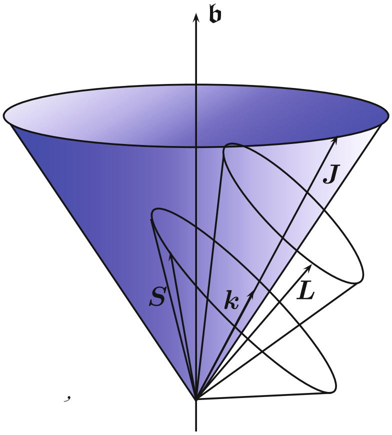


Fig. 9.3 Vector diagram for the calculation of the Landé g-factor

where we chose $\gamma = 2$ for the spin. What does this Hamiltonian look like in proportion to $\boldsymbol{J}$? We write

$$\boldsymbol{H} = -g_J(L, S)\gamma \boldsymbol{J} \cdot \mathfrak{b}, \tag{9.2}$$

where $g_J(L, S)$ is a constant that depends on $\boldsymbol{L}$, $\boldsymbol{S}$, and $\boldsymbol{J}$. Hereby, we assume that the Hamilton functions in (9.1) and (9.2) share at least their values on the diagonals.

In Fig. 9.3, three precession movements can be seen, namely $\boldsymbol{L}$ around $\boldsymbol{J}$, $\boldsymbol{S}$ around $\boldsymbol{J}$, and $\boldsymbol{J}$ around $\mathfrak{b}$. The effective magnetic moment can be determined by projecting $\boldsymbol{L}$ onto $\boldsymbol{J}$ and $\boldsymbol{J}$ onto $\mathfrak{b}$, and then the same for $\boldsymbol{S}$. Let $\boldsymbol{k}$ be a unit vector pointing in the direction of $\boldsymbol{J}$, that is, $\boldsymbol{k} \overset{\text{def}}{=} \boldsymbol{J}/|\boldsymbol{J}|$. Then the projections are

$$\boldsymbol{L} \cdot \mathfrak{b} \longrightarrow (\boldsymbol{L} \cdot \boldsymbol{k})(\boldsymbol{k} \cdot \mathfrak{b}) = \frac{(\boldsymbol{L} \cdot \boldsymbol{J})(\boldsymbol{J} \cdot \mathfrak{b})}{|\boldsymbol{J}|^2}, \tag{9.3}$$

$$\boldsymbol{S} \cdot \mathfrak{b} \longrightarrow (\boldsymbol{S} \cdot \boldsymbol{k})(\boldsymbol{k} \cdot \mathfrak{b}) = \frac{(\boldsymbol{S} \cdot \boldsymbol{J})(\boldsymbol{J} \cdot \mathfrak{b})}{|\boldsymbol{J}|^2}.$$

Because $\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}$, it follows that

$$2\boldsymbol{L} \cdot \boldsymbol{J} = \boldsymbol{J}^2 + \boldsymbol{L}^2 - \boldsymbol{S}^2 \quad \text{and} \quad 2\boldsymbol{S} \cdot \boldsymbol{J} = \boldsymbol{J}^2 + \boldsymbol{S}^2 - \boldsymbol{L}^2.$$

Substituting this in (9.1) and using the quantum mechanical variables (e.g., $\boldsymbol{J}^2$ is replaced by $J(J + 1)\hbar^2$), we obtain

$$\boldsymbol{H} = -\gamma(\boldsymbol{L} + 2\boldsymbol{S}) \cdot \mathfrak{b} =$$

$$= -\gamma\left(1 + \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)}\right)\boldsymbol{J} \cdot \mathfrak{b}. \tag{9.4}$$

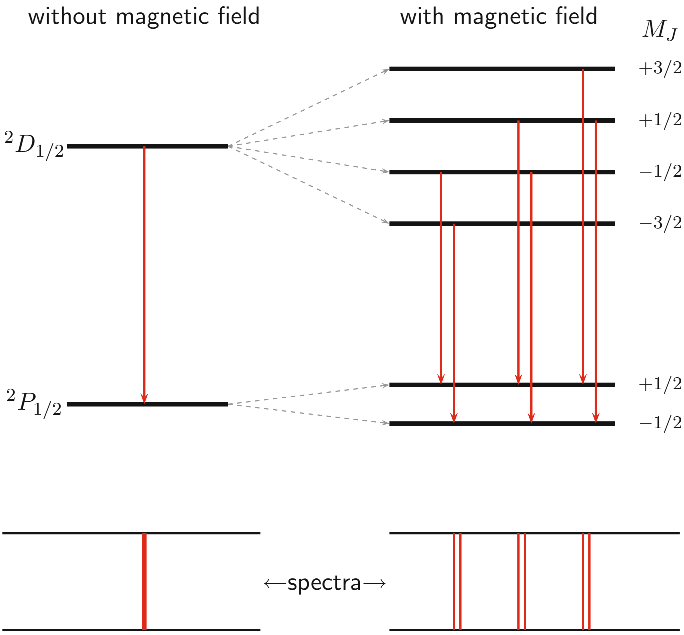


Fig. 9.4 Splitting of spectral lines due to the anomalous Zeeman effect. The various g-values make the spectrum more complicated than in the normal Zeeman effect (Fig. 9.2)

Compared with (9.2), we get the *Landé g-factor*

$$g_J(L, S) \overset{\text{def}}{=} 1 + \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)}. \tag{9.5}$$

If $S = 0$, then $g_J(L, S) = 1$, because J is then equal to L. In this case, the magnetic moment is independent of L, and we get the normal Zeeman effect as before; that is, all lines are shifted by the same amount. For $S \neq 0$, the value $g_J(L, S)$ depends on the values of L and S; that is, different lines are shifted by different amounts.

Figure 9.4 shows the anomalous Zeeman effect where the magnetic field is applied to the transition $^2D_{3/2} \longrightarrow ^2 P_{1/2}$. The Landé g-factor is computed for each level separately, and the level energy is then shifted proportional to its g-value. Hereby, the selection rule[1] $\Delta M_j = 0, \pm 1$ decides which transitions are allowed. For the level of $^2D_{3/2}$, we get $L = 2,\ S = \frac{1}{2}$, and $J = \frac{3}{2}$. This implies $g_{3/2}(2, \frac{1}{2}) = \frac{4}{5}$. For the lower level $^2P_{1/2}$, we find $g_{1/2}(1, \frac{1}{2}) = \frac{2}{3}$. The shift therefore amounts to $\frac{4}{5}\mu_B b$ in $^2D_{3/2}$, and $\frac{2}{3}\mu_B b$ in $^2P_{1/2}$, respectively. The six allowed transitions, namely three doublets, are shown in Fig. 9.4.

---

[1] In physics and chemistry, a selection rule constrains the possible transitions of a system from one quantum state to another.

9.2.2 Strong Magnetic Field

In strong magnetic fields ($B > 1$ Tesla), the coupling of the magnetic moments to the applied field can be stronger than the spin-orbit coupling. In such a case, the total spin $\boldsymbol{S}$ and the total orbital angular momentum $\boldsymbol{L}$ do not couple to $\boldsymbol{J}$, but precess independently around the axis of the applied magnetic field.

When the magnetic field increases, deviations from the equidistance of the splitting occur in the anomalous Zeeman effect. In fact, some of the individual lines approach each other such that the result resembles the normal Zeeman effect, with only triple splitting. This Paschen-Back effect case is called,[2,3]

9.3 Problem

9.1 Zeeman effect: Into how many lines does an energy line for $\ell = 2$ split up once a magnetic field is applied?

---

[2] Friedrich Paschen, 1865–1947, German physicist.

[3] Ernst Emil Alexander Back, 1881–1959, German physicist.

Chapter 10
Many Particle Systems

Abstract In this chapter, systems of distinguishable and indistinguishable particles are discussed in detail. Besides, the new concept of the entangled state is introduced. Also, the Pauli principle for the occupation of atomic shells is derived, and its application is discussed with respect to an explanation of the atomic structure.

10.1 Composed Systems

10.1.1 Systems with Two Distinguishable Particles

Let us consider a composite system consisting of two particles which do not interact with each other. The matrix $\boldsymbol{A}$ describes an observable of the particle 1, and $\boldsymbol{B}$ an observable of the particle 2. Any behaviour of a particle can therefore be described without reference to the other particle.

We can now prepare a state of the particle 1 in which the observable belonging to $\boldsymbol{A}$ has a unique value with probability 1. The corresponding state vector is an eigenvector of $\boldsymbol{A}$. A similar state vector exists for $\boldsymbol{B}$. The probability for the result of a simultaneous measurement of such observables that only relate to one or the other particle must be the product of the individual probabilities. For an overall representation of the system, this strongly suggests that all possible products of the components of the two state vectors occur in a common representation of both particles. Let the two state vectors be $\boldsymbol{a}$ and $\boldsymbol{b}$. We then obtain a vector containing all possible combinations of $a_i b_j$ by using the *Kronecker product* of the two vectors:

$$\boldsymbol{a} \otimes \boldsymbol{b} = \begin{pmatrix} a_1 \boldsymbol{b} \\ a_2 \boldsymbol{b} \\ a_3 \boldsymbol{b} \\ \vdots \end{pmatrix} = \begin{pmatrix} a_1 b_1 \\ a_1 b_2 \\ a_1 b_3 \\ \vdots \end{pmatrix}.$$

If $\boldsymbol{A}$ is a $N \times N$ matrix, and $\boldsymbol{B}$ a $M \times M$ matrix, the state vector $\boldsymbol{a}$ belongs to the N-dimensional Hilbert space $\mathcal{H}_1$, and $\boldsymbol{b}$ is an element of the M-dimensional Hilbert

G. Ludyk, *Quantum Mechanics in Matrix Form*,
https://doi.org/10.1007/978-3-319-26366-3_10

space $\mathcal{H}_2$. The Kronecker-product $\boldsymbol{a} \otimes \boldsymbol{b}$ then belongs to the $N \times M$-dimensional Hilbert space $\mathcal{H}_{1,2}$, which is spanned by the Kronecker products $\boldsymbol{u_A} \otimes \boldsymbol{u_B}$ of the N eigenvectors $\boldsymbol{u_A}(i)$ of $\boldsymbol{A}$ and the M eigenvectors $\boldsymbol{u_B}(j)$ of $\boldsymbol{B}$. As an abbreviation for the newly established Hilbert space, we write

$$\mathcal{H}_{1,2} = \mathcal{H}_1 \otimes \mathcal{H}_2.$$

If the set $\{\boldsymbol{u}_A\}$ of vectors defines an orthonormal basis for $\mathcal{H}_1$, and the set $\{\boldsymbol{u}_B\}$ of vectors is an orthonormal basis for $\mathcal{H}_2$, then the set of pairs $\{(\boldsymbol{u}_A, \boldsymbol{u}_B)\}$ forms an orthonormal basis for the product space $\mathcal{H}_{1,2}$. For the scalar product on $\mathcal{H}_{1,2}$, we find

$$\begin{aligned}(\boldsymbol{u}_A(i) \otimes \boldsymbol{u}_B(m))^\dagger\,(\boldsymbol{u}_A(j) \otimes \boldsymbol{u}_B(n)) &= \left(\boldsymbol{u}_A(i)^\dagger \boldsymbol{u}_A(j)\right) \otimes \left(\boldsymbol{u}_B(m)^\dagger \boldsymbol{u}_B(n)\right) = \\ &= \left(\boldsymbol{u}_A(i)^\dagger \boldsymbol{u}_A(j)\right) \cdot \left(\boldsymbol{u}_B(m)^\dagger \boldsymbol{u}_B(n)\right). \qquad (10.1)\end{aligned}$$

Note that this is a normal product of two numbers, since the Kronecker product of two scalars on the right side of the equation is just an ordinary multiplication.

The expectation value $\langle \boldsymbol{A} \rangle$ for the observable $\boldsymbol{A}$ in a system that is in the state $\boldsymbol{\xi}$ is (as has been defined in (4.5))

$$\langle \boldsymbol{A} \rangle = \boldsymbol{\xi}^\dagger \boldsymbol{A}\,\boldsymbol{\xi}.$$

For a two-particle system, we can define $\boldsymbol{A}_{1,2} = \boldsymbol{A} \otimes \boldsymbol{B}$ for any two observables $\boldsymbol{A}$ and $\boldsymbol{B}$. If the subsystems are in the states $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$, respectively (the overall system is therefore in the state $\boldsymbol{\xi}_1 \otimes \boldsymbol{\xi}_2$), the expectation value is given by

$$\begin{aligned}\langle \boldsymbol{A}_{1,2} \rangle = (\boldsymbol{\xi}_1 \otimes \boldsymbol{\xi}_2)^\dagger (\boldsymbol{A} \otimes \boldsymbol{B})(\boldsymbol{\xi}_1 \otimes \boldsymbol{\xi}_2) &= (\boldsymbol{\xi}_1^\dagger \boldsymbol{A}\,\boldsymbol{\xi}_1) \otimes (\boldsymbol{\xi}_2^\dagger \boldsymbol{B}\,\boldsymbol{\xi}_2) = \\ &= (\boldsymbol{\xi}_1^\dagger \boldsymbol{A}\,\boldsymbol{\xi}_1)(\boldsymbol{\xi}_2^\dagger \boldsymbol{B}\,\boldsymbol{\xi}_2) = \langle \boldsymbol{A} \rangle \langle \boldsymbol{B} \rangle. \qquad (10.2)\end{aligned}$$

In other words, the expectation value of the composite system equals the product of the expectation values of the observables in the subsystems.

One important note has to be made on the matrices in $\mathcal{H}_{1,2}$. In general, we have

$$(\boldsymbol{A} \otimes \boldsymbol{B})(\boldsymbol{\xi}_1 \otimes \boldsymbol{\xi}_2) = (\boldsymbol{A}\,\boldsymbol{\xi}_1) \otimes (\boldsymbol{B}\,\boldsymbol{\xi}_2). \qquad (10.3)$$

For $\boldsymbol{B} = \boldsymbol{I}_M$, (10.3) would read

$$(\boldsymbol{A} \otimes \boldsymbol{I}_M)(\boldsymbol{\xi}_1 \otimes \boldsymbol{\xi}_2) = (\boldsymbol{A}\,\boldsymbol{\xi}_1) \otimes \boldsymbol{\xi}_2.$$

If we want to apply the matrices $\boldsymbol{A}$ or $\boldsymbol{B}$ (which only act on $\mathcal{H}_1$ and $\mathcal{H}_2$) in the composite system, we have to extend their effect to $\mathcal{H}_{1,2}$ by setting

$$\boldsymbol{A}_{1,2} \overset{\text{def}}{=} \boldsymbol{A} \otimes \boldsymbol{I}_M \quad \text{and} \quad \boldsymbol{B}_{1,2} \overset{\text{def}}{=} \boldsymbol{I}_N \otimes \boldsymbol{B}. \qquad (10.4)$$

With this definition, we get indeed

$$\boldsymbol{A}_{1,2}(\boldsymbol{\xi}_1 \otimes \boldsymbol{\xi}_2) = (\boldsymbol{A}\,\boldsymbol{\xi}_1) \otimes \boldsymbol{\xi}_2,$$

or

$$\boldsymbol{B}_{1,2}\boldsymbol{A}_{1,2}(\boldsymbol{\xi}_1 \otimes \boldsymbol{\xi}_2) = (\boldsymbol{A}\,\boldsymbol{\xi}_1) \otimes (\boldsymbol{B}\,\boldsymbol{\xi}_2).$$

One more interesting fact is provided by the following theorem.

Theorem *If* $\boldsymbol{A}$ *has the eigenvalues* a_i *and* $\boldsymbol{B}$ *has the eigenvalues* b_j*, then the so-called Kronecker sum*

$$\boldsymbol{A} \oplus \boldsymbol{B} \overset{\text{def}}{=} (\boldsymbol{A} \otimes \boldsymbol{I}_M) + (\boldsymbol{I}_N \otimes \boldsymbol{B}) \in \mathbb{R}^{N \cdot M \times N \cdot M} \tag{10.5}$$

has the eigenvalues

$$a_i + b_j.$$

If $\boldsymbol{a}_i$ *is an eigenvector of* $\boldsymbol{A}$ *and* $\boldsymbol{b}_j$ *is an eigenvector of* $\boldsymbol{B}$*, then*

$$\boldsymbol{a}_i \otimes \boldsymbol{b}_j$$

is an eigenvector of $\boldsymbol{A} \oplus \boldsymbol{B}$*.*

Proof $[(\boldsymbol{A} \otimes \boldsymbol{I}_M) + (\boldsymbol{I}_N \otimes \boldsymbol{B})](\boldsymbol{a}_i \otimes \boldsymbol{b}_j) = (\boldsymbol{A}\boldsymbol{a}_i \otimes \boldsymbol{b}_j) + (\boldsymbol{a}_i \otimes \boldsymbol{B}\boldsymbol{b}_j) =$
$= (a_i\boldsymbol{a}_i \otimes \boldsymbol{b}_j) + (\boldsymbol{a}_i \otimes b_j\boldsymbol{b}_j) =$
$= \underbrace{(a_i \otimes 1)}_{a_i}(\boldsymbol{a}_i \otimes \boldsymbol{b}_j) + \underbrace{(1 \otimes b_j)}_{b_j}(\boldsymbol{a}_i \otimes \boldsymbol{b}_j) = (a_i + b_j)(\boldsymbol{a}_i \otimes \boldsymbol{b}_j)$ **q.e.d.**

10.1.2 Systems with N Distinguishable Subsystems

For more than two subsystems, for example N, the Hilbert space $\mathcal{H}^{(N)}$ is the product space of N single Hilbert spaces:

$$\mathcal{H}^{(N)} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3 \otimes \cdots \otimes \mathcal{H}_N.$$

A system state is then given by

$$\boldsymbol{\xi}^{(N)} = \boldsymbol{\xi}_1 \otimes \boldsymbol{\xi}_2 \otimes \boldsymbol{\xi}_3 \otimes \cdots \otimes \boldsymbol{\xi}_N.$$

Instead of the matrices $\boldsymbol{A}_i$ that only act on $\mathcal{H}_i$, the corresponding matrices $\boldsymbol{A}_i^{(N)}$ that act on $\mathcal{H}^{(N)}$ have to be used:

$$\boldsymbol{A}_i^{(N)} = \boldsymbol{I}_{(1)} \otimes \cdots \otimes \boldsymbol{I}_{(i-1)} \otimes \boldsymbol{A}_i \otimes \boldsymbol{I}_{(i+1)} \otimes \cdots \otimes \boldsymbol{I}_{(N)}.$$

Then we get in fact

$$\boldsymbol{A}_i^{(N)} \boldsymbol{\xi}^{(N)} = \boldsymbol{\xi}_1 \otimes \cdots \otimes \boldsymbol{\xi}_{i-1} \otimes \boldsymbol{A}_i \boldsymbol{\xi}_i \otimes \boldsymbol{\xi}_{i+1} \otimes \cdots \otimes \boldsymbol{\xi}_N,$$

i.e. the operator only acts on $\boldsymbol{\xi}_i$ in the subspace $\mathcal{H}_i$. Consequently, the expectation value in the composite system is defined by

$$\left\langle \boldsymbol{A}^{(N)} \right\rangle = \langle \boldsymbol{A}_1 \rangle \langle \boldsymbol{A}_2 \rangle \cdots \langle \boldsymbol{A}_N \rangle .$$

In other words: The expectation value for a combination of individual observables is equal to the product of the expectation values for each subsystem observable.

Like in the case of two particles, the extended Kronecker sum

$$\bigoplus_{i=1}^{N} \boldsymbol{A}_i = \sum_{i=1}^{N} \boldsymbol{A}_i^{(N)}$$

has the eigenvalues

$$\sum_i a_i^{(N)}$$

if the $\boldsymbol{A}_i$ have the eigenvalues a_i. The proof is very similar to the above theorem for only two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$.

10.1.3 Entangled Systems

Let us again start with two systems in the Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, connected to $\mathcal{H}_A \otimes \mathcal{H}_B$. If the system A is in the state $\boldsymbol{\xi}_A$ and the System B in state $\boldsymbol{\xi}_B$, the state of the composite system is $\boldsymbol{\xi}_A \otimes \boldsymbol{\xi}_B$. Let $\{\boldsymbol{u}_A(i)\}$ and $\{\boldsymbol{u}_B(j)\}$ be a orthonormal basis of $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively. Then the composite state can be rewritten as

$$\boldsymbol{\xi}_A \otimes \boldsymbol{\xi}_B = \left(\sum_i a_i \boldsymbol{u}_A(i) \right) \otimes \left(\sum_j b_j \boldsymbol{u}_B(j) \right).$$

States of the composite system that can be represented in such a product form are called *separable states*. Not all possible states in $\mathcal{H}_A \otimes \mathcal{H}_B$ are separable states, however! The most general state in the composite system has the representation

$$\boldsymbol{\xi}_{AB} = \sum_{i,j} c_{ij}(\boldsymbol{u}_A(i) \otimes \boldsymbol{u}_B(j)).$$

Obviously this state is only separable if $c_{ij} = a_i \cdot b_j$, yielding $\boldsymbol{\xi}_A = \sum_i a_i \boldsymbol{u}_A(i)$ and $\boldsymbol{\xi}_B = \sum_j b_j \boldsymbol{u}_B(j)$. The state is inseparable if any one of the c_{ij} cannot be factored into a product $a_i \cdot b_j$. If a state is not separable, we call it an *entangled state*. E.g., if $\boldsymbol{u}_A(1)$ and $\boldsymbol{u}_A(2)$ are two basis vectors in $\mathcal{H}_A$ and $\boldsymbol{u}_B(1)$ and $\boldsymbol{u}_B(2)$ two basis vectors in $\mathcal{H}_B$, the following state is an entangled state:

$$\boldsymbol{\xi}_{AB} = [\boldsymbol{u}_A(1) \otimes \boldsymbol{u}_B(2) - \boldsymbol{u}_A(2) \otimes \boldsymbol{u}_B(1)]/\sqrt{2}. \tag{10.6}$$

To see this, we write down the general condition for a state in $\mathcal{H}_A \otimes \mathcal{H}_B$ (which is $2 \times 2 = 4$-dimensional) to be a separable state:

$$\begin{aligned} \boldsymbol{\xi}_{AB} &= c_{11}(\boldsymbol{u}_A(1) \otimes \boldsymbol{u}_B(1)) + c_{12}(\boldsymbol{u}_A(1) \otimes \boldsymbol{u}_B(2)) + \\ &\quad + c_{21}(\boldsymbol{u}_A(2) \otimes \boldsymbol{u}_B(1)) + c_{22}(\boldsymbol{u}_A(2) \otimes \boldsymbol{u}_B(2)) \overset{!}{=} \\ &= a_1 b_1(\boldsymbol{u}_A(1) \otimes \boldsymbol{u}_B(1)) + a_1 b_2(\boldsymbol{u}_A(1) \otimes \boldsymbol{u}_B(2)) + \\ &\quad + a_2 b_1(\boldsymbol{u}_A(2) \otimes \boldsymbol{u}_B(1)) + a_2 b_2(\boldsymbol{u}_A(2) \otimes \boldsymbol{u}_B(2)). \end{aligned} \tag{10.7}$$

This yields a necessary condition for the state to be separable:

$$c_{11} \cdot c_{22} = c_{12} \cdot c_{21}. \tag{10.8}$$

In the case of (10.6), however, we find

$$c_{12} \cdot c_{21} = -1/2,$$

but

$$c_{11} \cdot c_{22} = 0 \neq -1/2.$$

The state $\boldsymbol{\xi}_{AB}$ in (10.6) is therefore not separable, i.e., the state is entangled. Entangled states are unique to quantum mechanics. They do not exist in classical mechanics and have no classical analogue! In quantum mechanics, however, they are fundamental in modern applications such as *quantum communication* and *quantum computing*.

10.2 Indistinguishable Subsystems

The most prominent examples for indistinguishable, i.e. identical particles are electrons in the atomic shell, or the protons and neutrons in the atomic nucleus. When the structure of the electron shells in an atom and the resulting setup of the periodic table of the elements was derived in the Bohr–Sommerfeld quantum theory, it had just been assumed that the energy levels of an atom cannot be occupied by an arbitrary number of electrons. Today we know that this is a consequence of the *Pauli principle*, according to which two electrons can never simultaneously occupy one and the same state. As we shall see soon, this principle is a direct consequence of the fundamental indistinguishability of the electrons!

Identical particles have identical particle properties (such as mass, spin, or charge). Of course, the state of each particle may be different and may also change over time. Still, an individual assignment such as

$$\text{Particle } i \Leftrightarrow \text{State } \boldsymbol{\xi}_i$$

is not possible! Instead, only a general assignment

$$\{\text{Set of all particles, } i = 1, \ldots, N\} \Leftrightarrow N\text{-particle state } \boldsymbol{\xi}^{(N)}.$$

Here, the overall state is again a composed state, i.e.

$$\boldsymbol{\xi}^{(N)} = \boldsymbol{\xi}_1 \otimes \boldsymbol{\xi}_2 \otimes \cdots \otimes \boldsymbol{\xi}_N. \tag{10.9}$$

However, all expectation values are only allowed to refer to the *totality* of the particles. In other words, they must look like

$$\boldsymbol{\xi}^{(N)^\dagger} \boldsymbol{A}_N \boldsymbol{\xi}^{(N)}. \tag{10.10}$$

Moreover, this result must not depend on the order of the sub-states $\boldsymbol{\xi}_i$ in (10.9). This gives us the condition

$$\begin{aligned}&(\boldsymbol{\xi}_1 \otimes \cdots \otimes \boldsymbol{\xi}_i \otimes \cdots \otimes \boldsymbol{\xi}_j \otimes \cdots \otimes \boldsymbol{\xi}_N)^\dagger \boldsymbol{A}_N (\boldsymbol{\xi}_1 \otimes \cdots \otimes \boldsymbol{\xi}_i \otimes \cdots \otimes \boldsymbol{\xi}_j \otimes \cdots \otimes \boldsymbol{\xi}_N) \overset{!}{=} \\ &(\boldsymbol{\xi}_1 \otimes \cdots \otimes \boldsymbol{\xi}_j \otimes \cdots \otimes \boldsymbol{\xi}_i \otimes \cdots \otimes \boldsymbol{\xi}_N)^\dagger \boldsymbol{A}_N (\boldsymbol{\xi}_1 \otimes \cdots \otimes \boldsymbol{\xi}_j \otimes \cdots \otimes \boldsymbol{\xi}_i \otimes \cdots \otimes \boldsymbol{\xi}_N).\end{aligned} \tag{10.11}$$

This condition must hold for any permutation of the state vectors.

Permutations

In Appendix B, a *permutation relationship* is specified as

$$\boldsymbol{U}_{s\times p}(\boldsymbol{B} \otimes \boldsymbol{A})\boldsymbol{U}_{q\times t} = \boldsymbol{A} \otimes \boldsymbol{B}, \quad \text{if } \boldsymbol{A} \in \mathbb{R}^{p\times q} \text{ and } \boldsymbol{B} \in \mathbb{R}^{s\times t}.$$

In the present case, we have to deal with the interchange of vectors, i.e.

$$\boldsymbol{U}_{s\times p}(\boldsymbol{b} \otimes \boldsymbol{a})\boldsymbol{U}_{1\times 1} = \underline{\underline{\boldsymbol{a} \otimes \boldsymbol{b} = \boldsymbol{U}_{s\times p}(\boldsymbol{b} \otimes \boldsymbol{a})}}, \quad \text{if } \boldsymbol{a} \in \mathbb{R}^{p\times 1} \text{ and } \boldsymbol{b} \in \mathbb{R}^{s\times 1}, \tag{10.12}$$

since $\boldsymbol{U}_{1\times 1} = 1$.

We now derive as an example what the permutation matrix must look like for the case of three state vectors $\boldsymbol{\xi}_1 \otimes \boldsymbol{\xi}_2 \otimes \boldsymbol{\xi}_3$, when the first and third state vector are interchanged. This is done in three steps. Since all subsystems are indistinguishable, we must assume that the state vectors all have the same dimension. In our example, the dimension is 2, i.e., $\boldsymbol{\xi}_i \in \mathbb{R}^2$. In the first step, $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ are interchanged. This is achieved by multiplying $\boldsymbol{\xi}_1 \otimes \boldsymbol{\xi}_2 \otimes \boldsymbol{\xi}_3$ with the *transposition matrix*[1] $\boldsymbol{T}_{12} = (\boldsymbol{U}_{2\times 2} \otimes \boldsymbol{I}_2)$, where $\boldsymbol{U}_{2\times 2} \in \mathbb{R}^{4\times 4}$ has the form

$$\boldsymbol{U}_{2\times 2} = \left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right),$$

i.e., it has only one 1 and otherwise only zeros in each row and in each column (like all permutation matrices). The transposition matrix $\boldsymbol{T}_{12} \in \mathbb{R}^{8\times 8}$ then has the form

$$\boldsymbol{T}_{12} = \boldsymbol{U}_{2\times 2} \otimes \boldsymbol{I}_2 = \left(\begin{array}{cc|cc} \boldsymbol{I_2} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{I_2} & \mathbf{0} \\ \hline \mathbf{0} & \boldsymbol{I_2} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{I_2} \end{array}\right).$$

In fact, we get

$$(\boldsymbol{U}_{2\times 2} \otimes \boldsymbol{I}_2)(\boldsymbol{\xi}_1 \otimes \boldsymbol{\xi}_2 \otimes \boldsymbol{\xi}_3) = \left(\begin{array}{cc|cc} \boldsymbol{I_2} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{I_2} & \mathbf{0} \\ \hline \mathbf{0} & \boldsymbol{I_2} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{I_2} \end{array}\right) \left(\begin{array}{c} \xi_{11}\xi_{21}\boldsymbol{\xi}_3 \\ \xi_{11}\xi_{22}\boldsymbol{\xi}_3 \\ \hline \xi_{12}\xi_{21}\boldsymbol{\xi}_3 \\ \xi_{12}\xi_{22}\boldsymbol{\xi}_3 \end{array}\right) =$$

$$= \left(\begin{array}{c} \xi_{11}\xi_{21}\boldsymbol{\xi}_3 \\ \xi_{12}\xi_{21}\boldsymbol{\xi}_3 \\ \hline \xi_{11}\xi_{22}\boldsymbol{\xi}_3 \\ \xi_{12}\xi_{22}\boldsymbol{\xi}_3 \end{array}\right) = \left(\begin{array}{c} \xi_{21}(\boldsymbol{\xi}_1 \otimes \boldsymbol{\xi}_3) \\ \hline \xi_{22}(\boldsymbol{\xi}_1 \otimes \boldsymbol{\xi}_3) \end{array}\right) = \boldsymbol{\xi}_2 \otimes \boldsymbol{\xi}_1 \otimes \boldsymbol{\xi}_3.$$

In a next step, we multiply this intermediate result with $\boldsymbol{T}_{23} = (\boldsymbol{I}_2 \otimes \boldsymbol{U}_{2\times 2})$, which leads to

$$\boldsymbol{T}_{23}(\boldsymbol{\xi}_2 \otimes \boldsymbol{\xi}_1 \otimes \boldsymbol{\xi}_3) = \boldsymbol{\xi}_2 \otimes \boldsymbol{\xi}_3 \otimes \boldsymbol{\xi}_1.$$

In a third step, we multiply the result with $\boldsymbol{T}_{12} = (\boldsymbol{U}_{2\times 2} \otimes \boldsymbol{I}_2)$ and finally obtain the desired result

$$\boldsymbol{T}_{12}(\boldsymbol{\xi}_2 \otimes \boldsymbol{\xi}_3 \otimes \boldsymbol{\xi}_1) = \boldsymbol{\xi}_3 \otimes \boldsymbol{\xi}_2 \otimes \boldsymbol{\xi}_1.$$

---

[1] If only two states are exchanged, the permutation is called *transposition matrix*.

In total, the permutation matrix $\boldsymbol{P}_{1,3}$ has the form

$$\boldsymbol{P}_{1,3} = \boldsymbol{T}_{12}\boldsymbol{T}_{23}\boldsymbol{T}_{12} = (\boldsymbol{U}_{2\times2} \otimes \boldsymbol{I}_2)(\boldsymbol{I}_2 \otimes \boldsymbol{U}_{2\times2})(\boldsymbol{U}_{2\times2} \otimes \boldsymbol{I}_2) =$$

$$= \left(\begin{array}{cc|cc|cc|cc} 1&0&0&0&0&0&0&0\\ 0&0&0&0&1&0&0&0\\ \hline 0&0&1&0&0&0&0&0\\ 0&0&0&0&0&0&1&0\\ \hline 0&1&0&0&0&0&0&0\\ 0&0&0&0&0&1&0&0\\ \hline 0&0&0&1&0&0&0&0\\ 0&0&0&0&0&0&0&1. \end{array}\right)$$

You can easily check that

$$\boldsymbol{P}_{1,3}\boldsymbol{P}_{1,3} = \boldsymbol{I},$$

which yields

$$\boldsymbol{P}_{1,3}^{-1} = \boldsymbol{P}_{1,3}.$$

Of course this result does not come as a surprise! After all, applying the same permutation twice must lead back to the original state.

Transposition and Permutation

As just explained, a transposition matrix[2] $\boldsymbol{T}_{ij}$ just swaps the ith and the jth state vector. In an N-particles system, there are $N(N-1)/\,2$ different transpositions, and it must hold

$$\boldsymbol{T}_{ii} = \boldsymbol{I}, \quad \boldsymbol{T}_{ij} = \boldsymbol{T}_{ji}.$$

Each of the N different permutations can be obtained as a product of transpositions. This can be done in various ways. In any case, however, the sign of a permutation is always defined as +1 if an even number of transpositions is needed, and −1 for an odd number of transpositions.

Considering (10.10), (10.11) can be rewritten with permutation matrices as

$$\begin{aligned} \boldsymbol{\xi}^{(N)\dagger}\boldsymbol{A}_N\boldsymbol{\xi}^{(N)} &= (\boldsymbol{P}_{ij}\boldsymbol{\xi}^{(N)})^\dagger \boldsymbol{A}_N(\boldsymbol{P}_{ij}\boldsymbol{\xi}^{(N)}) = \\ &= \boldsymbol{\xi}^{(N)\dagger}\boldsymbol{P}_{ij}^\dagger \boldsymbol{A}_N \boldsymbol{P}_{ij}\boldsymbol{\xi}^{(N)}, \end{aligned} \tag{10.13}$$

which yields the matrix identity

$$\boldsymbol{A}_N = \boldsymbol{P}_{ij}^\dagger \boldsymbol{A}_N \boldsymbol{P}_{ij}. \tag{10.14}$$

---

[2] A transposition is an exchange of two elements in an ordered list with all other elements staying the same, i.e., a permutation of two elements. For example, the swapping of 2 and 5 in the list 123456 yields a transposition 153426 of the list. The *permutation symbol* $\epsilon_{ijk\ldots}$ is defined as $(-1)^n$, where n is the number of transpositions that must be applied to build up the permutation.

Furthermore, the scalar product of the permuted state vector must satisfy the condition

$$(\boldsymbol{P}_{ij}\boldsymbol{\xi}^{(N)})^{\dagger}(\boldsymbol{P}_{ij}\boldsymbol{\xi}^{(N)}) = \boldsymbol{\xi}^{(N)\dagger}\boldsymbol{\xi}^{(N)}, \tag{10.15}$$

i.e.,

$$\boldsymbol{I}_N = \boldsymbol{P}_{ij}^{\dagger}\boldsymbol{P}_{ij}, \tag{10.16}$$

or

$$\boldsymbol{P}_{ij}^{\dagger} = \boldsymbol{P}_{ij}^{-1}. \tag{10.17}$$

In other words, the permutation matrix $\boldsymbol{P}_{ij}$ is unitary. From (10.14), we then get

$$\boldsymbol{P}_{ij}\boldsymbol{A}_N = \boldsymbol{A}_N\boldsymbol{P}_{ij},$$

i.e., the $\boldsymbol{P}_{ij}$ commute with $\boldsymbol{A}_N$:

$$\underline{\underline{[\boldsymbol{A}_N, \boldsymbol{P}_{ij}] = \boldsymbol{0}.}} \tag{10.18}$$

What are the eigenvalues of $\boldsymbol{P}_{ij}$? Since $\boldsymbol{P}_{ij}$ commutes with $\boldsymbol{A}_N$, the two matrices (operators) have the same eigenvectors. Naming the eigenvectors $\boldsymbol{e}_{ij}$ and the eigenvalues λ_{ij}, we get

$$\boldsymbol{P}_{ij}\boldsymbol{e}_{ij} = \lambda_{ij}\boldsymbol{e}_{ij}. \tag{10.19}$$

Since $\boldsymbol{P}_{ij}^2 = \boldsymbol{I}$, we have $\lambda_{ij}^2 = 1$. Therefore, the eigenvalues can only be $+1$ or -1.

10.2.1 Interchanging Two Particles

We consider the example of two spin-1/2 particles, e.g., two electrons. The spin state of an electron is described by a two-dimensional vector in $\mathbb{C}^2$. The spin itself is three-dimensional and has three components in x_1-, x_2- and x_3-direction. This means that the information about a three-dimensional real vector is included in a two-dimensional complex state vector. What are the eigenvectors? Using the Pauli matrices $\boldsymbol{\sigma}_i$, each of the two particles can be represented according to (8.24) as

$$\boldsymbol{S}_i = \frac{\hbar}{2}\boldsymbol{\sigma}_i,$$

hence

$$\boldsymbol{S}_1 = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \boldsymbol{S}_2 = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \text{and} \quad \boldsymbol{S}_3 = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

All three matrices have the eigenvalues $+\frac{\hbar}{2}$ and $-\frac{\hbar}{2}$. The eigenvectors, which are called *spinor*s in this case, are given by

$$\boldsymbol{S}_1 : \quad \boldsymbol{e}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm 1 \end{pmatrix}; \quad \boldsymbol{S}_2 : \quad \boldsymbol{e}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm i \end{pmatrix};$$

$$\boldsymbol{S}_3 : \quad \boldsymbol{e}_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \boldsymbol{e}_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The spins of the two electrons are described by the matrix vectors $\mathfrak{S}'$ and $\mathfrak{S}''$, where

$$\mathfrak{S}' = \begin{pmatrix} \boldsymbol{S}_1' \\ \boldsymbol{S}_2' \\ \boldsymbol{S}_3' \end{pmatrix} \quad \text{and} \quad \mathfrak{S}'' = \begin{pmatrix} \boldsymbol{S}_1'' \\ \boldsymbol{S}_2'' \\ \boldsymbol{S}_3'' \end{pmatrix}.$$

For the two spin angular momentum, we have the usual relations

$$[\boldsymbol{S}_1', \boldsymbol{S}_2'] = i\hbar \boldsymbol{S}_3' \quad \text{and} \quad [\boldsymbol{S}_1'', \boldsymbol{S}_2''] = i\hbar \boldsymbol{S}_3''.$$

The two sets of matrices $\boldsymbol{S}_i'$ and $\boldsymbol{S}_j''$, however, commute with each other, since the two particles are assumed to be independent of each other. This implies

$$[\boldsymbol{S_i}', \boldsymbol{S_j}''] = \boldsymbol{0}.$$

According to (10.4) and (10.5), we define

$$\boldsymbol{S}_3 = \boldsymbol{S}_3' \oplus \boldsymbol{S}_3'' = (\boldsymbol{S}_3' \otimes \boldsymbol{I}_2) + (\boldsymbol{I}_2 \otimes \boldsymbol{S}_3''). \tag{10.20}$$

With

$$\boldsymbol{S}_3' \otimes \boldsymbol{I}_2 = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

and

$$\boldsymbol{I}_2 \otimes \boldsymbol{S}_3'' = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1, \end{pmatrix}$$

we obtain

$$\boldsymbol{S}_3 = \boldsymbol{S}_3' \oplus \boldsymbol{S}_3'' = \hbar \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1. \end{pmatrix}$$

The eigenvectors of $\boldsymbol{S}_3$ are

$$\boldsymbol{e}_i' \otimes \boldsymbol{e}_j'',$$

where $\boldsymbol{S}'_3$ has the eigenvectors $\boldsymbol{e}'_i$, and $\boldsymbol{S}'_3$ has the eigenvectors $\boldsymbol{e}''_j$. The eigenvalues are then the sum of the corresponding eigenvalues, i.e.

$$\lambda'_i + \lambda''_j.$$

With (6.28) and (6.29), we get for the spinors $\boldsymbol{e}'_\pm$ of the first electron

$$\boldsymbol{S}'^2 \boldsymbol{e}'_\pm = \frac{1}{2}\left(\frac{1}{2}+1\right)\hbar^2 \boldsymbol{e}'_\pm = \frac{3}{4}\hbar^2 \boldsymbol{e}'_\pm \tag{10.21}$$

and

$$\boldsymbol{S}'_3 \boldsymbol{e}'_\pm = \pm\frac{\hbar}{2}\boldsymbol{e}'_\pm. \tag{10.22}$$

The same holds for the spinors $\boldsymbol{e}''_\pm$ of the second electron. The matrix $\boldsymbol{S}_3$ of the two-particle system then has the four eigenvectors

$$\boldsymbol{e}'_+ \otimes \boldsymbol{e}''_+ = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \quad \boldsymbol{e}'_+ \otimes \boldsymbol{e}''_- = \begin{pmatrix}0\\1\\0\\0\end{pmatrix},$$

$$\boldsymbol{e}'_- \otimes \boldsymbol{e}''_+ = \begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \quad \boldsymbol{e}'_- \otimes \boldsymbol{e}''_- = \begin{pmatrix}0\\0\\0\\1\end{pmatrix}, \tag{10.23}$$

to which we find the four eigenvalues $+\hbar$, 0, 0 and $-\hbar$, e.g. by having a look at the main diagonal of the diagonal matrix $\boldsymbol{S}_3$.

What subspaces of $\mathcal{H}_{1,2}$ do the eigenvectors span? Let the two electrons be in the states $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$, respectively, where each $\boldsymbol{\xi}_i$ completely describes the particle's coordinates: $\boldsymbol{\xi}_i = \boldsymbol{x}_i \otimes \boldsymbol{s}_i$ ($\boldsymbol{x}_i$ is the spatial coordinate and $\boldsymbol{s}_i$ the spin). The permutation matrix $\boldsymbol{P}_{1,2} = \boldsymbol{T}_{12}$ is again defined such that the states $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ of the two particles are exchanged:

$$\boldsymbol{P}_{1,2}(\boldsymbol{\xi}_1 \otimes \boldsymbol{\xi}_2) = \boldsymbol{\xi}_2 \otimes \boldsymbol{\xi}_1. \tag{10.24}$$

Of course, the permutation matrix $\boldsymbol{P}_{1,2}$ has only the two eigenvalues $+1$ and -1 also in this case. The eigenstates with eigenvalue $+1$ are called *symmetric*, and the eigenstates with eigenvalue -1 are called *antisymmetric* states. The symmetric and the antisymmetric states each form a subspace of $\mathcal{H}_{1,2} = \mathcal{H}_1 \otimes \mathcal{H}_2$, since each linear combination of symmetric or antisymmetric states is again symmetric or antisymmetric. Let $\mathcal{H}_{1,2}^{(+)}$ be the symmetric subspace and $\mathcal{H}_{1,2}^{(-)}$ the antisymmetric subspace. It can be easily verified that a state of the form

$$(\boldsymbol{e}_1 \otimes \boldsymbol{e}_2)^{(+)} \overset{\text{def}}{=} \alpha^+(\boldsymbol{e}_1 \otimes \boldsymbol{e}_2 + \boldsymbol{e}_2 \otimes \boldsymbol{e}_1) \tag{10.25}$$

is always symmetrical (with the normalization constant α^+). However, a compound state like

$$(\boldsymbol{e}_1 \otimes \boldsymbol{e}_2)^{(-)} \overset{\text{def}}{=} \alpha^-(\boldsymbol{e}_1 \otimes \boldsymbol{e}_2 - \boldsymbol{e}_2 \otimes \boldsymbol{e}_1) \tag{10.26}$$

is always antisymmetric, since in such case we get

$$\boldsymbol{P}_{12}(\boldsymbol{e}_1 \otimes \boldsymbol{e}_2)^{(-)} = -(\boldsymbol{e}_1 \otimes \boldsymbol{e}_2)^{(-)}.$$

If we want to generate members of the *symmetric subspace* $\mathcal{H}_{1,2}^{(+)}$ from members of the complete space $\mathcal{H}_{1,2} = \mathcal{H}_1 \otimes \mathcal{H}_2$, we can apply the operator

$$\boldsymbol{P_S} \overset{\text{def}}{=} \boldsymbol{I} + \boldsymbol{P}_{12}$$

on $\boldsymbol{\xi}_1 \otimes \boldsymbol{\xi}_2$. Accordingly, the elements of the *asymmetric subspace* $\mathcal{H}_{1,2}^{(-)}$ are obtained by applying the operator

$$\boldsymbol{P_A} \overset{\text{def}}{=} \boldsymbol{I} - \boldsymbol{P}_{12}.$$

As a proof, let us apply $\boldsymbol{P_S}$ or $\boldsymbol{P_A}$ on an arbitrary vector $\alpha(\boldsymbol{e}_1 \otimes \boldsymbol{e}_2) + \beta(\boldsymbol{e}_2 \otimes \boldsymbol{e}_1)$. The resulting vector is

$$\boldsymbol{P}_S(\alpha(\boldsymbol{e}_1 \otimes \boldsymbol{e}_2) + \beta(\boldsymbol{e}_2 \otimes \boldsymbol{e}_1)) = \frac{\alpha+\beta}{2}(\boldsymbol{e}_1 \otimes \boldsymbol{e}_2 + \boldsymbol{e}_2 \otimes \boldsymbol{e}_1),$$

or

$$\boldsymbol{P}_A(\alpha(\boldsymbol{e}_1 \otimes \boldsymbol{e}_2) + \beta(\boldsymbol{e}_2 \otimes \boldsymbol{e}_1)) = \frac{\alpha+\beta}{2}(\boldsymbol{e}_1 \otimes \boldsymbol{e}_2 - \boldsymbol{e}_2 \otimes \boldsymbol{e}_1).$$

Note that the following also applies:

$$\begin{aligned} &\frac{1}{2}\left(\frac{1}{\alpha^+}(\boldsymbol{e}_1 \otimes \boldsymbol{e}_2)^{(+)} + \frac{1}{\alpha^-}(\boldsymbol{e}_1 \otimes \boldsymbol{e}_2)^{(-)}\right) = \\ &= \frac{1}{2}\left[(\boldsymbol{e}_1 \otimes \boldsymbol{e}_2 + \boldsymbol{e}_2 \otimes \boldsymbol{e}_1) + (\boldsymbol{e}_1 \otimes \boldsymbol{e}_2 - \boldsymbol{e}_2 \otimes \boldsymbol{e}_1)\right] = \boldsymbol{e}_1 \otimes \boldsymbol{e}_2. \end{aligned} \tag{10.27}$$

In other words, each state of the two-body system can be written as a linear combination of a symmetric and an antisymmetric state. Therefore, the two Hilbert spaces $\mathcal{H}_{1,2}^{(+)}$ and $\mathcal{H}_{1,2}^{(-)}$ span the entire Hilbert space $\mathcal{H}_{1,2}$.

Moreover, symmetric and antisymmetric states are orthogonal to each other, since their scalar product yields $\boldsymbol{\xi}^{(+)} \in \mathcal{H}_{1,2}^{(+)}$ with $\boldsymbol{\xi}^{(-)} \in \mathcal{H}_{1,2}^{(-)}$

$$\boldsymbol{\xi}^{(+)^\dagger}\boldsymbol{\xi}^{(-)} = \boldsymbol{\xi}^{(+)^\dagger}\underbrace{\boldsymbol{P}_{12}^\dagger\boldsymbol{P}_{12}}_{\boldsymbol{I}}\boldsymbol{\xi}^{(-)} = \boldsymbol{\xi}^{(+)^\dagger}(+1)(-1)\boldsymbol{\xi}^{(-)} = -\boldsymbol{\xi}^{(+)^\dagger}\boldsymbol{\xi}^{(-)}, \tag{10.28}$$

what can only be true if $\boldsymbol{\xi}^{(+)^\dagger}\boldsymbol{\xi}^{(-)} = 0$, i.e., $\boldsymbol{\xi}^{(+)}$ and $\boldsymbol{\xi}^{(-)}$ are orthogonal to each other.

Fermions and Bosons

As we just discussed, a state with exchanged particles is physically indistinguishable from the initial state in a quantum mechanical two-body system of identical particles. After all, we can only measure the expectation value, e.g. $\langle \boldsymbol{X} \rangle = \boldsymbol{\xi}^\dagger \boldsymbol{X} \boldsymbol{\xi}$ for the observable $\boldsymbol{X}$. If under the permutation of $\boldsymbol{\xi}$ the value changes to $-\boldsymbol{\xi}$, then we still have

$$(-\boldsymbol{\xi}^\dagger)\boldsymbol{X}(-\boldsymbol{\xi}) = \boldsymbol{\xi}^\dagger \boldsymbol{X} \boldsymbol{\xi} = \langle \boldsymbol{X} \rangle \, .$$

Now experience shows that this factor ($+1$ or -1) due to the permutation is always either $+1$ or -1 for particles of a given kind. All particles with *half-integer spin* ($\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \ldots$), also called *fermion*s,[3] always have a factor -1 under exchange. Meanwhile, all particles with *integer spin* ($0, 1, 2, \ldots$), also called *bosons*,[4] always have a factor $+1$. Mathematically speaking, we get for the states of fermions

$$\boldsymbol{\xi}_{Fermion} \in \mathcal{H}^{(-)}_{1,2},$$

and for bosons

$$\boldsymbol{\xi}_{Boson} \in \mathcal{H}^{(+)}_{1,2}.$$

If we want to describe two indistinguishable particles, it is therefore not required to use the Hilbert space $\mathcal{H}_{1.2}$. Instead, only one of the subspaces $\mathcal{H}^{(-)}_{1,2}$ or $\mathcal{H}^{(+)}_{1,2}$ is enough, depending on whether the particles are fermions or bosons. The vector basis is then given according to (10.25) or (10.26):

$$(\boldsymbol{e}_1 \otimes \boldsymbol{e}_2)^{(\mp)} = \frac{1}{\sqrt{2}} (\boldsymbol{e}_1 \otimes \boldsymbol{e}_2 \mp \boldsymbol{e}_2 \otimes \boldsymbol{e}_1).$$

The atomic building blocks are all made out of fermions:

*Lepton*s: electron e, electron-neutrino ν_e, myon-neutrino ν_μ, tauon-neutrino ν_τ, myon μ, tauon τ;
*Baryon*s: proton, neutron, Λ-baryon, Σ-baryon, Ω-baryon;
Quarks: up, charm, top (all $s = +\frac{1}{2}$), down, strange, bottom (all $s = -\frac{1}{2}$);
The *Helium-isotope*s ^3He ($s = \frac{1}{2}$).

The following particles are bosons:

*Meson*s: pion, kaon, ρ-meson, ω-meson;
Photons ($s = 1$);
*Gauge boson*s ($s = 1$);
*Gluon*s ($s = 1$);
The *Helium-isotope*s ^4He ($s = 0$).

---

[3] Enrico Fermi, 1901–1954, Italian physicist, Nobel Prize 1938.

[4] Satyendra Nath Bose, 1894–1974, Indian physicist.

10.2.2 Interchanging Three Identical Particles

Before we get to the general case with N particles, let us examine the case of 3 identical particles as an intermediate stage. There are three transpositions $\boldsymbol{P}_{12}$, $\boldsymbol{P}_{23}$ and $\boldsymbol{P}_{13}$. The symmetric and antisymmetric linear combinations of the $N! = 3! = 6$ states $\boldsymbol{e}_1^{(1)} \otimes \boldsymbol{e}_2^{(2)} \otimes \boldsymbol{e}_3^{(3)}$, $\boldsymbol{e}_1^{(1)} \otimes \boldsymbol{e}_3^{(2)} \otimes \boldsymbol{e}_2^{(3)}$, $\boldsymbol{e}_2^{(1)} \otimes \boldsymbol{e}_1^{(2)} \otimes \boldsymbol{e}_3^{(3)}$, $\boldsymbol{e}_2^{(1)} \otimes \boldsymbol{e}_3^{(2)} \otimes \boldsymbol{e}_1^{(3)}$, $\boldsymbol{e}_3^{(1)} \otimes \boldsymbol{e}_1^{(2)} \otimes \boldsymbol{e}_2^{(3)}$ and $\boldsymbol{e}_3^{(1)} \otimes \boldsymbol{e}_2^{(2)} \otimes \boldsymbol{e}_1^{(3)}$ are given by

$$(\boldsymbol{e}_1 \otimes \boldsymbol{e}_2 \otimes \boldsymbol{e}_3)^{(+)} \overset{\text{def}}{=} \frac{1}{\sqrt{6}}(\boldsymbol{e}_1^{(1)} \otimes \boldsymbol{e}_2^{(2)} \otimes \boldsymbol{e}_3^{(3)} + \boldsymbol{e}_1^{(1)} \otimes \boldsymbol{e}_3^{(2)} \otimes \boldsymbol{e}_2^{(3)} + \boldsymbol{e}_2^{(1)} \otimes \boldsymbol{e}_1^{(2)} \otimes \boldsymbol{e}_3^{(3)} + \\ + \boldsymbol{e}_2^{(1)} \otimes \boldsymbol{e}_3^{(2)} \otimes \boldsymbol{e}_1^{(3)} + \boldsymbol{e}_3^{(1)} \otimes \boldsymbol{e}_1^{(2)} \otimes \boldsymbol{e}_2^{(3)} + \boldsymbol{e}_3^{(1)} \otimes \boldsymbol{e}_2^{(2)} \otimes \boldsymbol{e}_1^{(3)}) \tag{10.29}$$

and

$$(\boldsymbol{e}_1 \otimes \boldsymbol{e}_2 \otimes \boldsymbol{e}_3)^{(-)} \overset{\text{def}}{=} \frac{1}{\sqrt{6}}(\boldsymbol{e}_1^{(1)} \otimes \boldsymbol{e}_2^{(2)} \otimes \boldsymbol{e}_3^{(3)} - \boldsymbol{e}_1^{(1)} \otimes \boldsymbol{e}_3^{(2)} \otimes \boldsymbol{e}_2^{(3)} - \boldsymbol{e}_2^{(1)} \otimes \boldsymbol{e}_1^{(2)} \otimes \boldsymbol{e}_3^{(3)} + \\ + \boldsymbol{e}_2^{(1)} \otimes \boldsymbol{e}_3^{(2)} \otimes \boldsymbol{e}_1^{(3)} + \boldsymbol{e}_3^{(1)} \otimes \boldsymbol{e}_1^{(2)} \otimes \boldsymbol{e}_2^{(3)} - \boldsymbol{e}_3^{(1)} \otimes \boldsymbol{e}_2^{(2)} \otimes \boldsymbol{e}_1^{(3)}), \tag{10.30}$$

where $\boldsymbol{e}_i^{(j)}$ is the ith eigenvector of the jth particle. Both the symmetric vector $(\boldsymbol{e}_1 \otimes \boldsymbol{e}_2 \otimes \boldsymbol{e}_3)^{(+)}$ and the antisymmetric vector$(\boldsymbol{e}_1 \otimes \boldsymbol{e}_2 \otimes \boldsymbol{e}_3)^{(-)}$ span a *one-dimensional* subspace (also called a *ray*) of the Hilbert space $\mathcal{H}$. The shape of the antisymmetric vector $(\boldsymbol{e}_1 \otimes \boldsymbol{e}_2 \otimes \boldsymbol{e}_3)^{(-)}$ reminds us of the shape of a 3 × 3 determinant. Remember that for an ordinary determinant we get

$$\det \begin{vmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{vmatrix} =$$

$$= e_{11}e_{22}e_{33} - e_{11}e_{23}e_{32} - e_{12}e_{21}e_{33} + e_{12}e_{23}e_{31} + e_{13}e_{21}e_{32} - e_{13}e_{22}e_{31}.$$

If we replace the entries of the determinant by vectors and the normal multiplication by the Kronecker product (indicated by the index ⊗ at the right hand side of the determinant), then we find

$$\det \begin{vmatrix} \boldsymbol{e}_1^{(1)} & \boldsymbol{e}_2^{(1)} & \boldsymbol{e}_3^{(1)} \\ \boldsymbol{e}_1^{(2)} & \boldsymbol{e}_2^{(2)} & \boldsymbol{e}_3^{(2)} \\ \boldsymbol{e}_1^{(3)} & \boldsymbol{e}_2^{(3)} & \boldsymbol{e}_3^{(3)} \end{vmatrix}_{\otimes} \overset{\text{def}}{=}$$

$$= \boldsymbol{e}_1^{(1)} \otimes \boldsymbol{e}_2^{(2)} \otimes \boldsymbol{e}_3^{(3)} - \boldsymbol{e}_1^{(1)} \otimes \boldsymbol{e}_3^{(2)} \otimes \boldsymbol{e}_2^{(3)} - \boldsymbol{e}_2^{(1)} \otimes \boldsymbol{e}_1^{(2)} \otimes \boldsymbol{e}_3^{(3)} +$$

$$+ \boldsymbol{e}_2^{(1)} \otimes \boldsymbol{e}_3^{(2)} \otimes \boldsymbol{e}_1^{(3)} + \boldsymbol{e}_3^{(1)} \otimes \boldsymbol{e}_1^{(2)} \otimes \boldsymbol{e}_2^{(3)} - \boldsymbol{e}_3^{(1)} \otimes \boldsymbol{e}_2^{(2)} \otimes \boldsymbol{e}_1^{(3)},$$

where $\boldsymbol{e}_i^{(j)}$ is the ith eigenvector of the jth particle. Summing up, we get

$$(\boldsymbol{e}_1 \otimes \boldsymbol{e}_2 \otimes \boldsymbol{e}_3)^{(-)} = \frac{1}{\sqrt{6}} \cdot \det \begin{vmatrix} \boldsymbol{e}_1^{(1)} & \boldsymbol{e}_2^{(1)} & \boldsymbol{e}_3^{(1)} \\ \boldsymbol{e}_1^{(2)} & \boldsymbol{e}_2^{(2)} & \boldsymbol{e}_3^{(2)} \\ \boldsymbol{e}_1^{(3)} & \boldsymbol{e}_2^{(3)} & \boldsymbol{e}_3^{(3)} \end{vmatrix}_{\otimes} . \tag{10.31}$$

Each particle is assigned to one row. A transposition of two particles thus corresponds to interchanging two rows of the determinant. From basic linear algebra we know that such an interchange changes the sign of the determinant. Therefore the particle is indeed a fermion! Also, the determinant is equal to zero if two lines are equal. This fact brings us to a crucial theorem in quantum theory.

Pauli Principle Two identical fermions cannot occupy the same quantum state simultaneously.

There is an alternative version of Pauli's principle.

Theorem *The state of an atom is fully described by four quantum numbers* n, ℓ, m_J *and* m_S, *and it can be occupied by at most one electron.*

This important principle essentially determines the structure of the periodic table of the elements. Note, however, that the Pauli principle *only applies to fermions, not to bosons* (especially not for atomic nuclei)!

10.2.3 Interchanging N Identical Particles

For a system with N identical particles (think of an atom with N electrons on different shells), the state vector of each state $\boldsymbol{\xi}_k^{(i)}$ (i.e., the kth eigenvector of the ith particle) is composed of the vectors

$$\boldsymbol{\xi}(1, \ldots, N) \overset{\text{def}}{=} \boldsymbol{\xi}_{i_1}^{(1)} \otimes \boldsymbol{\xi}_{i_2}^{(2)} \otimes \cdots \otimes \boldsymbol{\xi}_{i_N}^{(N)}, \tag{10.32}$$

where $i_1, i_2, \ldots, i_N$ are permutations of the numbers $1, 2, \ldots, N$. Remember once again that each of the state vectors $\boldsymbol{\xi}^{(i)}$ of the single particle depends on both the location $\boldsymbol{x}^{(i)}$ and the spin $s^{(i)}$ of this particle. With the *transposition operator* $\boldsymbol{P}_{ij}$, which swaps two particles i and j, we get

$$\boldsymbol{P}_{ij}\boldsymbol{\xi}(1, \ldots, i, \ldots, j, \ldots, N) = \boldsymbol{\xi}(1, \ldots, j, \ldots, i, \ldots, N),$$

and obviously also

$$\boldsymbol{P}_{ij}^2 \boldsymbol{\xi}(1, \ldots, i, \ldots, j, \ldots, N) = \boldsymbol{\xi}(1, \ldots, i, \ldots, j, \ldots, N).$$

Due to $\boldsymbol{P}_{ij}^2 = \boldsymbol{I}$, the eigenvalues are bound to be $+1$. The transposition operator $\boldsymbol{P}_{ij}$ therefore can only have the eigenvalues $+1$ and -1.

If a state vector $\boldsymbol{\xi}$ and a permuted state vector $\boldsymbol{\xi}'$ are given, then the expectation values of any observable $\boldsymbol{A}$ must be the same:

$$\langle \boldsymbol{A} \rangle = \boldsymbol{\xi}^\dagger \boldsymbol{A}\, \boldsymbol{\xi} = \boldsymbol{\xi}'^\dagger \boldsymbol{A}\, \boldsymbol{\xi}'.$$

We now write formally

$$\boldsymbol{\xi}' = \boldsymbol{P}\boldsymbol{\xi},$$

where the permutation operator is given by

$$\boldsymbol{P} = \begin{pmatrix} 1 & 2 & \cdots & N \\ i_1 & i_2 & \cdots & i_N, \end{pmatrix}$$

i.e., the index i_1 occurs in the first place, the index i_2 in the second place, etc. Remember that all permutations can be assembled from transpositions $\boldsymbol{T}_{ij}$. A permutation can be either composed of an even, or of an odd number of transpositions. Accordingly, these permutations are called *even permutations* or *odd permutations*. For example, for three particles we find

$$\text{even}: \quad \boldsymbol{P} = \begin{pmatrix} 1\,2\,3 \\ 1\,2\,3 \end{pmatrix}, \quad \boldsymbol{P} = \begin{pmatrix} 1\,2\,3 \\ 2\,3\,1 \end{pmatrix}, \quad \boldsymbol{P} = \begin{pmatrix} 1\,2\,3 \\ 3\,1\,2 \end{pmatrix};$$

$$\text{odd}: \quad \boldsymbol{P} = \begin{pmatrix} 1\,2\,3 \\ 2\,1\,3 \end{pmatrix}, \quad \boldsymbol{P} = \begin{pmatrix} 1\,2\,3 \\ 1\,3\,2 \end{pmatrix}, \quad \boldsymbol{P} = \begin{pmatrix} 1\,2\,3 \\ 3\,2\,1 \end{pmatrix}.$$

With these permutations, we can reshape the above requirement for observables as

$$(\boldsymbol{P}\boldsymbol{\xi})^\dagger \boldsymbol{A}\boldsymbol{P}\boldsymbol{\xi} = (\boldsymbol{\xi})^\dagger \boldsymbol{P}^\dagger \boldsymbol{A}\boldsymbol{P}\boldsymbol{\xi} \overset{!}{=} (\boldsymbol{\xi})^\dagger \boldsymbol{A}\boldsymbol{\xi}, \quad \text{for all } \boldsymbol{P}, \boldsymbol{\xi}.$$

It follows that $\boldsymbol{P}^\dagger \boldsymbol{A}\boldsymbol{P} = \boldsymbol{A}$, or with $\boldsymbol{P}^\dagger = \boldsymbol{P}^{-1}$

$$\boldsymbol{A}\boldsymbol{P} = \boldsymbol{P}\boldsymbol{A} \text{ for all } \boldsymbol{P}.$$

In other words: In a system of identical particles, only those observables are allowed which commute with all permutations. Such observables are, e.g., the total momentum and the total angular momentum.

Next, we claim that the *symmetric state* $\boldsymbol{\xi}(1, \ldots, N)^{(+)}$ of a system can be calculated as

$$\boxed{\boldsymbol{\xi}(1,\ldots,N)^{(+)} = \frac{1}{\sqrt{N!}}\sum_{\beta}\boldsymbol{P}_{\beta}(\boldsymbol{\xi}_{i_1}^{(1)}\otimes\boldsymbol{\xi}_{i_2}^{(2)}\otimes\cdots\otimes\boldsymbol{\xi}_{i_N}^{(N)})}. \tag{10.33}$$

$\boldsymbol{P}_{\beta}$ is the permutation operator which swaps the $\boldsymbol{\xi}^{(i)}$ $(i = 1,\ldots,N)$, and β numbers the $N!$ permutations of $\boldsymbol{\xi}^{(i)}$. The normalization factor $1/\sqrt{N!}$ normalizes the sum to 1 and ensures

$$\boldsymbol{\xi}^{\dagger}(1,\ldots,N)^{(+)}\boldsymbol{\xi}(1,\ldots,N)^{(+)} = 1.$$

We shall now prove that there is only one symmetric state $\boldsymbol{\xi}(1,\ldots,N)^{(+)}$, i.e., the application of an arbitrary permutation operator $\boldsymbol{P}_{\alpha}$ to $\boldsymbol{\xi}(1,\ldots,N)^{(+)}$ returns $\boldsymbol{\xi}(1,\ldots,N)^{(+)}$. This is indeed the case, since

$$\boldsymbol{P}_{\alpha}\boldsymbol{\xi}(1,\ldots,N)^{(+)} = \frac{1}{\sqrt{N!}}\sum_{\beta}\underbrace{\boldsymbol{P}_{\alpha}\boldsymbol{P}_{\beta}}_{\boldsymbol{P}_{\gamma}}(\boldsymbol{\xi}_{i_1}^{(1)}\otimes\boldsymbol{\xi}_{i_2}^{(2)}\otimes\cdots\otimes\boldsymbol{\xi}_{i_N}^{(N)}) =$$

$$= \frac{1}{\sqrt{N!}}\sum_{\gamma}\boldsymbol{P}_{\gamma}(\boldsymbol{\xi}_{i_1}^{(1)}\otimes\boldsymbol{\xi}_{i_2}^{(2)}\otimes\cdots\otimes\boldsymbol{\xi}_{i_N}^{(N)}) = \boldsymbol{\xi}(1,\ldots,N)^{(+)}.$$

Furthermore, we now claim that the *antisymmetric state* of a system $\boldsymbol{\xi}(1,\ldots,N)^{(-)}$ can be calculated as

$$\boxed{\boldsymbol{\xi}(1,\ldots,N)^{(-)} = \frac{1}{\sqrt{N!}}\sum_{\beta}(-1)^{\pi_{\beta}}\boldsymbol{P}_{\beta}(\boldsymbol{\xi}_{i_1}^{(1)}\otimes\boldsymbol{\xi}_{i_2}^{(2)}\otimes\cdots\otimes\boldsymbol{\xi}_{i_N}^{(N)})}. \tag{10.34}$$

π_{β} denotes the minimum number of transpositions from which $\boldsymbol{P}_{\beta}$ is constructed. In other words,

$$(-1)^{\pi_{\beta}} = \begin{cases} +1, & \text{if } \boldsymbol{P}_{\beta} \text{ is even} \\ -1, & \text{if } \boldsymbol{P}_{\beta} \text{ is odd.} \end{cases}$$

This formula strongly reminds us of the Leibniz formula (see Appendix F) for $n \times n$ determinants:

$$\det \boldsymbol{A} = \sum_{\beta}\left((-1)^{\pi_{\beta}}\prod_{i=1}^{n}A_{i,P_i}\right).$$

Therefore, the antisymmetric state (10.34) can also be written with the help of the *Slater determinant*[5] as

---

[5] John Clarke Slater, 1900–1976, American physicist.

$$\boldsymbol{\xi}(1,\ldots,N)^{(-)} = \frac{1}{\sqrt{N!}} \cdot \det \begin{vmatrix} \boldsymbol{\xi}_1^{(1)} & \boldsymbol{\xi}_2^{(1)} & \cdots & \boldsymbol{\xi}_N^{(1)} \\ \boldsymbol{\xi}_1^{(2)} & \boldsymbol{\xi}_2^{(2)} & \cdots & \boldsymbol{\xi}_N^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\xi}_1^{(N)} & \boldsymbol{\xi}_2^{(N)} & \cdots & \boldsymbol{\xi}_N^{(N)} \end{vmatrix}_{\otimes} . \qquad (10.35)$$

Remember that the symbol $\otimes$ on the right side of the determinant indicates that the Kronecker product is to be used. The determinant's shape immediately reveals its antisymmetry under transpositions. After all, an exchange of two particles corresponds to exchanging two rows of the determinant! Furthermore, the Pauli principle follows immediately, since two equal lines in the determinant render the determinant zero.

The Pauli principle can also be derived by considering an antisymmetric N-particle state in which two single-particle states $\boldsymbol{\xi}^{(i)}$ and $\boldsymbol{\xi}^{(j)}$ are equal. Then

$$\boldsymbol{\xi}_{i_1}^{(1)} \otimes \cdots \otimes \boldsymbol{\xi}_{i_i}^{(i)} \otimes \cdots \otimes \boldsymbol{\xi}_{i_j}^{(j)} \otimes \cdots \otimes \boldsymbol{\xi}_{i_N}^{(N)} = -\boldsymbol{\xi}_{i_1}^{(1)} \otimes \cdots \otimes \boldsymbol{\xi}_{i_j}^{(j)} \otimes \cdots \otimes \boldsymbol{\xi}_{i_i}^{(i)} \otimes \cdots \otimes \boldsymbol{\xi}_{i_N}^{(N)} .$$

It follows immediately that

$$\boldsymbol{\xi}_{i_1}^{(1)} \otimes \cdots \otimes \boldsymbol{\xi}_{i_i}^{(i)} \otimes \cdots \otimes \boldsymbol{\xi}_{i_j}^{(j)} \otimes \cdots \otimes \boldsymbol{\xi}_{i_N}^{(N)} = \boldsymbol{0}, \quad \text{if} \quad \boldsymbol{\xi}^{(i)} = \boldsymbol{\xi}^{(j)} .$$

Summing up, two identical fermions may never have the same single-particle quantum numbers simultaneously. When filling up available states with fermions, the Pauli principle must always be taken into account! This is how periodic table of the elements arises.

10.3 Problems

10.1 Entangled States: Arrange the four components c_{ij} in a 2×2 matrix and derive the condition (10.8) from it.

10.2 Commutator: In (10.4), we defined the matrices

$$\boldsymbol{A}_{1,2} = \boldsymbol{A} \otimes \boldsymbol{I}_B, \quad \text{and} \quad \boldsymbol{B}_{1,2} = \boldsymbol{I}_A \otimes \boldsymbol{B}.$$

What is the commutator $[\boldsymbol{A}_{1,2}, \boldsymbol{B}_{1,2}]$ of these two matrices?

10.3 Expectation Value: In (10.4), we defined the matrices

$$\boldsymbol{A}_{1,2} = \boldsymbol{A} \otimes \boldsymbol{I}_B, \quad \text{and} \quad \boldsymbol{B}_{1,2} = \boldsymbol{I}_A \otimes \boldsymbol{B}.$$

What are the expectation values $\langle \boldsymbol{A}_{1,2} \rangle$ and $\langle \boldsymbol{B}_{1,2} \rangle$ of these two observables?

10.4 Permutation Matrix: Which permutation matrix $\boldsymbol{P}$ swaps two spin states? What is $\boldsymbol{P}^{-1}$ and $\boldsymbol{P}^{\dagger}$? What are the eigenvalues of $\boldsymbol{P}$?

Chapter 11
Equivalence of Matrix and Wave Mechanics

Abstract Schrödinger's wave mechanics is briefly introduced, and its equivalence with the matrix formalism of Heisenberg (and others) is demonstrated.

11.1 The De Broglie Wavelength

Schrödinger's wave mechanics is based on the matter-wave theory of de Broglie. For his hypothesis of the wave nature of material particles, de Broglie was inspired by the classical analogy between the mechanical Maupertuis-Euler principle of least action and Fermat's principle on the shortest light path in optics. According to de Broglie's hypothesis, each free particle induces a *wave field* that is linked in such a way that a particle with momentum $\boldsymbol{p}$ and energy E corresponds to a plane wave propagating in the direction of $\boldsymbol{p}$. The wavelength of the particle satisfies the same relationship as for photons, namely

$$p = \frac{h}{\lambda} = \hbar k,$$

that is, the condition

$$\boldsymbol{k} = \boldsymbol{p}/\hbar$$

for the wave vector, and for the frequency

$$\omega = E/\hbar, \tag{11.1}$$

where E is the total energy of the particle and $\hbar$ is the reduced Planck quantum of action $\hbar = h/2\pi$.

In general, one can represent a wave as

$$\varphi = A\cos\left[2\pi\left(\frac{r}{\lambda} - \frac{t}{T}\right)\right], \tag{11.2}$$

G. Ludyk, *Quantum Mechanics in Matrix Form*,
https://doi.org/10.1007/978-3-319-26366-3_11

where r is the distance from the origin, λ denotes the wavelength, and T is the oscillation period. For a wave in three-dimensional space, we need to write

$$\varphi = A\cos\left(\boldsymbol{k}^{\mathsf{T}}\boldsymbol{x} - \frac{2\pi\, t}{T}\right), \tag{11.3}$$

where the wave vector $\boldsymbol{k}$ is defined such that it represents the direction of the wave and has the length $|\boldsymbol{k}| = 2\pi/\lambda$. From a mathematical point of view, a harmonic plane wave in three-dimensional space can be written as

$$A(\mathbf{x}, t) = \operatorname{Re}\left(A_0 e^{i\left(\mathbf{k}^{\mathsf{T}}\mathbf{x} - \omega t\right)}\right) = \operatorname{Re}(A_0)\cos(\mathbf{k}^{\mathsf{T}}\mathbf{x} - \omega t) - \operatorname{Im}(A_0)\sin(\mathbf{k}^{\mathsf{T}}\mathbf{x} - \omega t), \tag{11.4}$$

with the complex amplitude A_0.

11.2 Operators in the Schrödinger Formalism

The *differentiation* $\frac{\partial}{\partial q}$ is an example of a *linear operator* because it has the property

$$\frac{\partial}{\partial q}\left[f(q) + g(q)\right] = \frac{\partial}{\partial q} f(q) + \frac{\partial}{\partial q} g(q). \tag{11.5}$$

The same property is given for the *multiplication* with q:

$$q\left[f(q) + g(q)\right] = q\, f(q) + q\, g(q). \tag{11.6}$$

Let R be a general operator acting on the function $f(q)$ with the resulting function $R\, f(q)$. If another operator T is applied to $R\, f(q)$, the result is yet another function $T\, R\, f(q)$. The composed operator $T\, R$ is called the *operator product*. By the way, adding operators follows the simple rule

$$(T + R) f(q) = T\, f(q) + R\, f(q).$$

For such operators, there exists an eigenvalue equation as well. However, we now find *eigenfunctions* corresponding to a given eigenvalue (instead of eigenvectors as in the case of matrix operators). For example, the operator $D = \frac{\partial^2}{\partial q^2}$ applied to the function $f(q) = \cos(\omega q)$ yields

$$Df(q) = \frac{\partial^2}{\partial q^2}\cos(\omega q) = -\omega^2\cos(\omega q) = -\omega^2 f(q).$$

The function $f(q) = \cos(\omega q)$ is thus an eigenfunction to D and $-\omega^2$ is the corresponding eigenvalue.

Operators that are defined on a space of functions can thus be used similarly to matrices. As in the matrix case, however, the commutative law $RT = TR$ generally does not apply! For example, we find that

$$\left(\frac{\partial}{\partial q}\, q - q\, \frac{\partial}{\partial q}\right) f(q) = \underbrace{\frac{\partial q}{\partial q}}_{1}\, f + q\frac{\partial f}{\partial q} - q\frac{\partial f}{\partial q} = f(q). \tag{11.7}$$

As an operator equation, this reads

$$\frac{\partial}{\partial q}\, q - q\, \frac{\partial}{\partial q} = 1, \tag{11.8}$$

that is, an equation that is similar to the commutation relation for position and momentum in matrix theory.

If we denote the operator "multiplication by q" with Q, if P is the operation $\frac{h}{2\pi i}\frac{\partial}{\partial q}$, and if the operator "multiplication with 1" is denoted by I, then (11.8) reads

$$P\,Q - Q\,P = \frac{h}{2\pi i}\, I, \tag{11.9}$$

which corresponds exactly to the commutation relation (3.15)

$$\boldsymbol{PX} - \boldsymbol{XP} = \frac{h}{2\,\pi\, i}\boldsymbol{I}$$

from matrix mechanics. Formally speaking, there is an "isomorphism" between the operators P, Q and the matrices $\boldsymbol{P}$, $\boldsymbol{X}$. For the harmonic oscillator, for example, we get the Hamilton operator (here again with x instead of q)

$$H = \frac{P^2}{2m} + \frac{m\omega^2 X^2}{2} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + \frac{m\omega^2 x^2}{2}, \tag{11.10}$$

where m is the particle mass, ω is the angular eigenfrequency of the oscillator, X is the position operator, and $P = -i\hbar\frac{\partial}{\partial x}$ is the momentum operator.

11.3 Schrödinger's Wave Mechanics

The starting points for the Schrödinger equation are Louis de Broglie's idea of matter-waves and the Hamilton-Jacobi theory of classical mechanics. The so-called *wave function* ψ can be interpreted as the solution of a linear partial differential equation.

What does the differential equation look like? According to de Broglie, a free particle is always associated with a plane wave. Therefore, it seems plausible to assume the wave function ψ for a particle to be

$$\psi = Ae^{i(\mathbf{k}^\intercal \mathbf{x} - \omega t)}, \tag{11.11}$$

according to (11.4). Equation (11.1) states that the total energy E of the particle is proportional to the frequency ω, namely

$$E = \hbar\omega.$$

The total energy E of the particle consists of kinetic energy E_{kin} and potential energy V:

$$E = E_{kin} + V.$$

A particle of mass m and momentum $\boldsymbol{p}$ has the kinetic energy

$$E_{kin} = \frac{1}{2}\frac{p^2}{m}, \tag{11.12}$$

where $p^2 = \boldsymbol{p}^\intercal \boldsymbol{p}$. In general, the potential energy of a particle is a function of the position $\boldsymbol{x}$ of the particle; that is, $V = V(\boldsymbol{x})$. Therefore we get

$$E = E_{kin} + V(\boldsymbol{x}) = \frac{p^2}{2m} + V(\boldsymbol{x}). \tag{11.13}$$

Assuming that the particle is represented by the wave function (11.11), we can calculate the derivative of this wave function with respect to time t:

$$\frac{\partial}{\partial t}\psi(t,\boldsymbol{x}) = -i\omega Ae^{i(\mathbf{k}^\intercal \mathbf{x} - \omega t)} = -i\omega\,\psi(t,\boldsymbol{x}) = -i\left(\frac{E}{\hbar}\right)\psi(t,\boldsymbol{x}),$$

or

$$\underline{\underline{i\hbar\frac{\partial}{\partial t}\psi(t,\boldsymbol{x}) = E\,\psi(t,\boldsymbol{x}).}} \tag{11.14}$$

The derivative of a plane wave (which in a first step depends only on x) with respect to the space variable x is

$$\frac{\partial}{\partial x}\psi(t,x) = ikAe^{i(kx - omegat)} = \left(\frac{ip}{\hbar}\right)\psi(t,x),$$

therefore

$$\frac{\partial^2}{\partial x^2}\psi(t,x) = -\left(\frac{p}{\hbar}\right)^2\psi(t,x).$$

Dividing by $2m$ and rearranging finally yields

$$\underline{\underline{-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(t,x)}} = \frac{p^2}{2m}\psi(t,x) = \underline{\underline{E_{kin}\psi(t,x)}}. \tag{11.15}$$

If we now insert $E = E_{kin} + V(x)$ into (11.14), we find

$$i\hbar\frac{\partial}{\partial t}\psi(t,x) = (E_{kin} + V(x))\,\psi(t,x) = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(t,x) + V(x)\psi(t,x). \tag{11.16}$$

In general, however, the plane wave depends not only on the x-direction, but also on the y- and z-directions, that is, on the full position vector $\boldsymbol{x}$ of the particle. If we differentiate the wave function (11.11) with respect to the vector components x_ν ($x_1 = x, x_2 = y, x_3 = z$), we obtain

$$\frac{\partial}{\partial x_\nu}\psi(t,\boldsymbol{x}) = ik_\nu A e^{i(\mathbf{k}^\mathsf{T}\mathbf{x}-\omega t)} = \left(\frac{ip_\nu}{\hbar}\right)\psi(t,\boldsymbol{x})$$

and

$$\frac{\partial^2}{\partial x_\nu^2}\psi(t,\boldsymbol{x}) = ik_\nu^2 A e^{i(\mathbf{k}^\mathsf{T}\mathbf{x}-\omega t)} = \left(\frac{ip_\nu}{\hbar}\right)^2\psi(t,\boldsymbol{x}). \tag{11.17}$$

Dividing this equation by $2m$ and rearranging yields

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x_\nu^2}\psi(t,\boldsymbol{x}) = \frac{p_\nu^2}{2m}\psi(t,\boldsymbol{x}). \tag{11.18}$$

We can add these three equations for $\nu = 1$, $\nu = 2$, and $\nu = 3$ and use the definition

$$\nabla^2 \overset{\text{def}}{=} \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2},$$

and we finally obtain

$$-\frac{\hbar^2}{2m}\nabla^2\psi(t,\boldsymbol{x}) = \frac{p^2}{2m}\psi(t,\boldsymbol{x}) = E_{kin}\psi(t,\boldsymbol{x}). \tag{11.19}$$

Putting $E = E_{kin} + V(t,\boldsymbol{x})$ into (11.14) similar to the one-dimensional case, one obtains

$$i\hbar\frac{\partial}{\partial t}\psi(t,\boldsymbol{x}) = (E_{kin} + V(t,\boldsymbol{x})\psi(t,\boldsymbol{x}) = \left(-\frac{\hbar^2}{2m}\nabla^2 + V(\boldsymbol{x})\right)\psi(t,\boldsymbol{x}). \tag{11.20}$$

If we now introduce the *Hamilton operator*

$$H \overset{\text{def}}{=} -\frac{\hbar^2}{2m}\nabla^2 + V(\boldsymbol{x}), \tag{11.21}$$

we obtain *Schrödinger's wave equation*

$$\boxed{i\hbar\frac{\partial}{\partial t}\psi(t,\boldsymbol{x}) = H\psi(t,\boldsymbol{x})}. \tag{11.22}$$

Note that we "derived" this equation only for the plane wave. In other words, if a particle is assumed to have a wave function (11.11), it is bound to satisfy the above equation. With this plausibility check in mind, Schrödinger postulated that this equation is valid for *all* quantum mechanical wave functions.

In the Schrödinger equation, we find both the wave function and the Hamilton operator. In Heisenberg's picture, however, the equations of motion only considered the operators in question. Still, we now show that the two formulations are mathematically equivalent.

11.4 Equivalence of Heisenberg and Schrödinger Pictures

The Schrödinger equation (11.22) is a linear partial differential equation. Solutions can be found by separating the variables, that is, by the ansatz

$$\psi(t,\boldsymbol{x}) = f(t)\cdot\varphi(\boldsymbol{x}). \tag{11.23}$$

Substituting (11.23) in (11.22), one obtains

$$i\hbar\varphi(\boldsymbol{x})\frac{\mathrm{d}f(t)}{\mathrm{d}t} = \left[-\frac{\hbar^2}{2m}\nabla^2\varphi(\boldsymbol{x}) + V(\boldsymbol{x})\varphi(\boldsymbol{x})\right]f(t). \tag{11.24}$$

Dividing by $f(t)\cdot\varphi(\boldsymbol{x})$ yields

$$\frac{i\hbar\,\mathrm{d}f(t)/\mathrm{d}t}{f(t)} = \frac{-(\hbar^2/2m)(\nabla^2\varphi(\boldsymbol{x})) + V(\boldsymbol{x})\varphi(\boldsymbol{x})}{\varphi(\boldsymbol{x})}. \tag{11.25}$$

Because the left side does not depend on $\boldsymbol{x}$ and the right side does not depend on t, both sides must be equal to the same constant! We can therefore set

$$\boxed{i\hbar\frac{\mathrm{d}f(t)}{\mathrm{d}t} = E\,f(t)} \tag{11.26}$$

and

$$\boxed{-\frac{\hbar^2}{2m}\nabla^2\varphi(\boldsymbol{x}) + V(\boldsymbol{x})\varphi(\boldsymbol{x}) = H\,\varphi(\boldsymbol{x}) = E\,\varphi(\boldsymbol{x})}. \tag{11.27}$$

Equation (11.26) describes the temporal development of $\psi(t, \boldsymbol{x})$. Equation (11.27) is an eigenvalue equation, which is generally referred to as a *time-independent* or stationary Schrödinger equation. It is a linear, second-order partial differential equation of the variable $\boldsymbol{x}$. In contrast, (11.26) is an ordinary linear differential equation of first order, which can be solved by the ansatz

$$f(t) = c\exp(-iE\,t/\hbar).$$

In the eigenvalue equation for H,

$$H\,\varphi_\nu = h_\nu\varphi_\nu,$$

the h_ν are the *eigenvalues* of the operator H, and the functions φ_ν represent the associated *eigenfunctions*. In other words, if we apply the operator H to an eigenfunction, we obtain the same function multiplied by the corresponding eigenvalue. If we consider, for example, the one-dimensional free motion of a particle on the interval [0, 1] of the x-axis, then the Schrödinger equation contains the Hamilton operator

$$H = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2},$$

and the eigenfunctions are

$$\varphi_\nu(x) = \sqrt{2}\sin(\nu\pi x). \tag{11.28}$$

Applying H to those eigenfunctions yields

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\sqrt{2}\sin(\nu\pi x) = \frac{\hbar^2(\nu\pi)^2}{2m}\sqrt{2}\sin(\nu\pi x),$$

thus the eigenvalues are

$$h_\nu = \frac{\hbar^2(\nu\pi)^2}{2\,m}, \nu = 1, 2, 3, \ldots.$$

For two different eigenvalues h_i and h_j, the corresponding eigenfunctions φ_i and φ_j have the property

$$\int_0^1 \varphi_i^*(x)\varphi_j(x)dx = 0 \quad \text{for} \quad i \neq j; \tag{11.29}$$

that is, they are *orthogonal* to each other. Also, the eigenfunctions $\varphi_i(x)$ are *normalized*, because

$$\int_0^1 \varphi_i^*(x)\varphi_i(x)dx = \int_0^1 |\varphi_i(x)|^2 dx = 1. \tag{11.30}$$

The general property for a set of normalized eigenfunctions is therefore

$$\int_0^1 \varphi_i^*(x)\varphi_j(x)dx = \delta_{ij}. \tag{11.31}$$

Now remember from mathematics that every continuous function $f(x)$ that is also continuously differentiable can be represented by a series with these eigenfunctions:

$$f(x) = \sum_{\nu=1}^{\infty} a_\nu \sqrt{2}\sin(\nu\pi x), \tag{11.32}$$

where the coefficients a_ν of the series expansion can be calculated according to

$$a_\nu = \int_{-\infty}^{+\infty} \varphi_\nu^*(x) f(x)\mathrm{d}x. \tag{11.33}$$

We now have all the necessary tools to prove the equivalence of Schrödinger's and Heisenberg's formalism. The proof is carried out in close analogy with the proof of Pasqual Jordan in [14]. First, we claim that the components $(\boldsymbol{X})_{\nu\mu}$ and $(\boldsymbol{P})_{\nu\mu}$ of the matrices $\boldsymbol{X}$ and $\boldsymbol{P}$ of Heisenberg's theory can be calculated from the normalized eigenfunctions φ_ν as follows.

$$(\boldsymbol{X})_{\nu\mu} = \int_{-\infty}^{+\infty} \varphi_\nu^*(x) \cdot x \cdot \varphi_\mu(x)\mathrm{d}x, \tag{11.34}$$

and

$$(\boldsymbol{P})_{\nu\mu} = \int_{-\infty}^{+\infty} \varphi_\nu^*(x) \cdot \frac{\hbar}{i} \frac{\mathrm{d}\varphi_\mu(x)}{\mathrm{d}x} \cdot \mathrm{d}x. \tag{11.35}$$

In other words, if a quantum problem is completely solved in the Schrödinger picture, we can also construct the corresponding matrix solution according to Heisenberg's matrix mechanics. The proof is divided into three steps.

1. First we show that the matrices as defined in (11.34) and (11.35) in fact satisfy the commutation relation (3.15)

$$\boldsymbol{PX} - \boldsymbol{XP} = \frac{\hbar}{i}\boldsymbol{I}. \tag{11.36}$$

For the elements $(\boldsymbol{PX})_{\nu\mu}$ of the matrix product $\boldsymbol{PX}$, we get

$$(\boldsymbol{PX})_{\nu\mu} = \sum_{\kappa} (\boldsymbol{P})_{\nu\kappa} (\boldsymbol{X})_{\kappa\mu} = \frac{\hbar}{i} \int_{-\infty}^{+\infty} \varphi_{\nu}^{*}(x) \frac{\mathrm{d}}{\mathrm{d}x} \sum_{\kappa} \varphi_{\kappa}(x) (\boldsymbol{X})_{\kappa\mu} \mathrm{d}x. \tag{11.37}$$

Now the function

$$f(x) \overset{\text{def}}{=} x \cdot \varphi_{\mu}(x)$$

can also be written as a series according to (11.32):

$$f(x) = x \cdot \varphi_{\mu}(x) = \sum_{\nu=1}^{\infty} a_{\nu} \varphi_{\nu}(x). \tag{11.38}$$

For the a_{ν}, (11.33) yields

$$a_{\nu} = \int_{-\infty}^{+\infty} \varphi_{\nu}^{*}(x) f(x) \mathrm{d}x = \int_{-\infty}^{+\infty} \varphi_{\nu}^{*}(x) x \cdot \varphi_{\mu}(x) \mathrm{d}x.$$

Putting this in (11.38) provides

$$x \cdot \varphi_{\mu}(x) = \sum_{\nu=1}^{\infty} \varphi_{\nu}(x) \int_{-\infty}^{+\infty} \varphi_{\nu}^{*}(y) y \cdot \varphi_{\mu}(y) \mathrm{d}y, \tag{11.39}$$

and with (11.34), we finally get

$$x \cdot \varphi_{\mu}(x) = \sum_{\nu=1}^{\infty} \varphi_{\nu}(x) (\boldsymbol{X})_{\nu\mu}. \tag{11.40}$$

Thus we find that

$$(\boldsymbol{PX})_{\nu\mu} = \frac{\hbar}{i} \int_{-\infty}^{+\infty} \varphi_{\nu}^{*}(x) \frac{\mathrm{d}}{\mathrm{d}x} x \cdot \varphi_{\mu}(x) \mathrm{d}x. \tag{11.41}$$

For the elements $(\boldsymbol{XP})_{\nu\mu}$ of the matrix product $\boldsymbol{XP}$, we similarly get with the help of (11.34)

$$(\boldsymbol{XP})_{\nu\mu} = \sum_{\kappa} (\boldsymbol{X})_{\nu\kappa} (\boldsymbol{P})_{\kappa\mu} = \frac{\hbar}{i} \int_{-\infty}^{+\infty} \varphi_{\nu}^{*}(x) x \sum_{\kappa} \varphi_{\kappa}(x) (\boldsymbol{P})_{\kappa\mu} \mathrm{d}x. \tag{11.42}$$

This requires a series expansion of the function

$$g(x) \overset{\text{def}}{=} \frac{\hbar}{i} \frac{\mathrm{d}}{\mathrm{d}x} \varphi(x) = \frac{\hbar}{i} \sum_{\nu} b_{\nu} \varphi_{\nu}(x),$$

with

$$b_\nu = \int_{-\infty}^{+\infty} \varphi_\nu^*(x) \frac{\mathrm{d}}{\mathrm{d}x} \varphi(x) \mathrm{d}x.$$

This leads to

$$g(x) = \frac{\hbar}{i} \frac{\mathrm{d}}{\mathrm{d}x} \varphi(x) = \frac{\hbar}{i} \sum_\nu \varphi_\nu(x) \int_{-\infty}^{+\infty} \varphi_\nu^*(y) \frac{\mathrm{d}}{\mathrm{d}y} \varphi(y) \mathrm{d}y = \frac{\hbar}{i} \sum_\nu \varphi_\nu(x) (\boldsymbol{P})_{\nu\mu}. \tag{11.43}$$

Inserting (11.43) into (11.42) finally gives

$$(\boldsymbol{X}\boldsymbol{P})_{\nu\mu} = \frac{\hbar}{i} \int_{-\infty}^{+\infty} \varphi_\nu^*(x) x \cdot \frac{\mathrm{d}}{\mathrm{d}x} \varphi_\mu(x) \mathrm{d}x. \tag{11.44}$$

Now we can combine the results (11.41) and (11.44), and we find

$$(\boldsymbol{P}\boldsymbol{X} - \boldsymbol{X}\boldsymbol{P})_{\nu\mu} = \frac{\hbar}{i} \int_{-\infty}^{+\infty} \varphi_\nu^*(x) \left[\frac{\mathrm{d}}{\mathrm{d}x} x - x \cdot \frac{\mathrm{d}}{\mathrm{d}x} \right] \varphi_\mu(x) \mathrm{d}x =$$

$$= \frac{\hbar}{i} \int_{-\infty}^{+\infty} \varphi_\nu^*(x) \varphi_\mu(x) \mathrm{d}x = \frac{\hbar}{i} \delta_{\nu\mu},$$

thus we get the commutation relation (11.36) indeed.

2. Now we show that the matrices $\boldsymbol{X}$ and $\boldsymbol{P}$ as constructed in (11.34) and (11.35) lead to the following representations of the matrices $\boldsymbol{U}(\boldsymbol{X})$ and $\boldsymbol{P}^2$ that occur in the Hamilton matrix $\boldsymbol{H}$.

$$\boldsymbol{U}(\boldsymbol{X})_{\nu\mu} = \int_{-\infty}^{+\infty} \varphi_\nu^*(x) \cdot V(x) \cdot \varphi_\mu(x) \mathrm{d}x \tag{11.45}$$

and

$$(\boldsymbol{P}^2)_{\nu\mu} = \int_{-\infty}^{+\infty} \varphi_\nu^*(x) \cdot -\hbar^2 \frac{\mathrm{d}^2 \varphi_\mu(x)}{\mathrm{d}x^2} \mathrm{d}x. \tag{11.46}$$

Note that $V(x)$ is the operator that corresponds to the matrix $\boldsymbol{V}(\boldsymbol{x})$. Equation (11.46) can be proved with the help of (11.35) and (11.43):

$$(\boldsymbol{P}^2)_{\nu\mu} = \sum_\kappa (\boldsymbol{P})_{\nu\kappa} (\boldsymbol{P})_{\kappa\mu} = \int_{-\infty}^{+\infty} \varphi_\nu^*(x) \frac{\hbar}{i} \frac{\mathrm{d}}{\mathrm{d}x} \sum_\kappa \varphi_\kappa(x) (\boldsymbol{P})_{\kappa\mu} =$$

$$= \int_{-\infty}^{+\infty} \varphi_\nu^*(x) \left(\frac{\hbar}{i} \right)^2 \frac{\mathrm{d}^2 \varphi_\mu(x)}{\mathrm{d}x^2} \mathrm{d}x. \tag{11.47}$$

Equation (11.45) can be proved by first showing that the assumption is correct for a power $\boldsymbol{V}(\boldsymbol{X}) = \boldsymbol{X}^r$. Because the potential can be written as a series $\boldsymbol{V}(\boldsymbol{X}) = \sum_r c_r \boldsymbol{X}^r$, the assumption then holds for an arbitrary $\boldsymbol{V}(\boldsymbol{X})$. We prove the simplified claim inductively. For this, let us assume that the assertion is true for $\boldsymbol{V}(\boldsymbol{X}) = \boldsymbol{X}^r$. Then according to (11.34) and (11.40), we indeed get

$$(\boldsymbol{X}^{r+1})_{\nu\mu} = \sum_\kappa (\boldsymbol{X}^r)_{\nu\kappa}(\boldsymbol{X})_{\kappa\mu} =$$

$$= \int_{-\infty}^{+\infty} \varphi_\nu^*(x) x^r \sum_\kappa \varphi_\kappa(x)(\boldsymbol{X})_{\kappa\mu} \cdot \mathrm{d}x = \int_{-\infty}^{+\infty} \varphi_\nu^*(x) x^r \cdot x \varphi_\mu(x) \mathrm{d}x. \tag{11.48}$$

3. It remains to be shown that the matrices $\boldsymbol{X}$ and $\boldsymbol{P}$ as constructed in (11.34) and (11.35) meet the commutation relation (11.36) and have the property

$$\boldsymbol{H}(\boldsymbol{P}, \boldsymbol{X}) = \frac{1}{2m}\boldsymbol{P}^2 + \boldsymbol{V}(\boldsymbol{X}) = \text{ diagonal matrix.} \tag{11.49}$$

With (11.45), (11.46), and (11.27) for $\varphi_\mu(x)$, we find

$$(\boldsymbol{H})_{\nu\mu} = \int_{-\infty}^{+\infty} \varphi_\nu^*(x) \left[\frac{-\hbar}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2}\varphi_\mu(x) + V(x)\varphi_\mu(x)\right] \mathrm{d}x =$$

$$= \int_{-\infty}^{+\infty} \varphi_\nu^*(x) E_\mu \varphi_\mu(x) \mathrm{d}x = E_\mu \delta_{\nu\mu}. \tag{11.50}$$

Indeed, the eigenvalues of the Hamilton matrix $\boldsymbol{H}(\boldsymbol{P}, \boldsymbol{X})$ equal the eigenvalues E_μ of the Schrödinger equation.

11.5 Example: The Harmonic Oscillator

In accordance with (11.10) and (11.27), the stationary Schrödinger equation for the harmonic oscillator is

$$\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + \frac{m\omega^2 x^2}{2}\right)\varphi_\nu(x) = E_\nu \varphi_\nu(x). \tag{11.51}$$

With

$$q \overset{\text{def}}{=} x\sqrt{m\omega/\hbar}, \tag{11.52}$$

we get

$$\frac{\mathrm{d}^2\varphi_\nu(q)}{\mathrm{d}q^2} + \left(\frac{2E_\nu}{\hbar\omega} - q^2\right)\varphi_\nu(q) = 0. \tag{11.53}$$

This nonlinear ordinary differential equation of second order is solved by the *Hermite polynomials*

$$H_\nu \overset{\text{def}}{=} (-1)^\nu e^{q^2} \frac{\mathrm{d}^\nu}{\mathrm{d}q^\nu} e^{-q^2}, \quad \nu = 0, 1, 2, \ldots \tag{11.54}$$

We therefore obtain the normalized eigenfunctions of (11.53) as

$$\varphi_\nu(q) = (2^\nu \nu! \sqrt{\pi})^{-1/2}\, e^{-q^2/2} H_\nu(q), \quad \nu = 0, 1, 2, \ldots \tag{11.55}$$

The eigenvalues are given by

$$\frac{2E_\nu}{\hbar\omega} = 2\nu + 1. \tag{11.56}$$

Compare this result to the energy values

$$E_\nu = \hbar\,\omega\left(\nu + \frac{1}{2}\right) \tag{11.57}$$

that we found in Chap. 5 with the help of quantum matrix mechanics.

11.6 Problems

11.1 **Schrödinger's Cat**: A cat is locked in a steel chamber, along with a machine that the cat cannot manipulate by any means. The machine includes a Geiger counter tube that contains a tiny amount of radioactive substance. The amount is so small that in the course of an hour perhaps one of the atoms decays, but also (with equal probability) perhaps none decays. If an atom decays, the counter tube discharges and releases a hammer that smashes a small bottle with cyanide. If we leave the entire system to itself for an hour, we can predict that the cat is still alive if no atom has decayed in the meantime. However, one single atomic decay would have poisoned it. How do you interpret this experiment from a quantum mechanical perspective?

11.2 **Commutators**: What are the commutators $[p, x]$, $[p, x^n]$, and $[p^n, x]$ for $p = \frac{\hbar}{i}\frac{\partial}{\partial x}$?

11.3 **Differentiation of Operators**: Derive the differentiation rules for functions of operators.

Chapter 12
Relativistic Quantum Mechanics

Abstract The basic concepts of the special theory of relativity are provided, followed by Dirac's application of these concepts to quantum mechanics.

12.1 Special Relativity

12.1.1 Four-Dimensional Spacetime

Dirac writes in his book [8]: "Let us now try to make the theory invariant under Lorentz transformations, so that it conforms to the principle of special relativity. This is necessary in order for the theory to apply to particles at high speed." In this section, we present the most important results of the special theory of relativity, as shown, for example, in my book [15]. Note that the components of the vectors that occur here are numbers, not matrices.

The theory of special relativity is based on two postulates:

- The laws of physics are the same for all inertial frames.
- The speed of light c is constant in all inertial frames.

An *inertial frame* is a frame of reference in which Newton's laws apply. In special relativity, one speaks of an *event* if something happens at a time t at some spatial location

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

With the help of the speed of light c, the time t is transformed into a space coordinate as ct. Space and time therefore form the four-dimensional *spacetime*. We denote an event in spacetime by a four-dimensional vector

$$\vec{\boldsymbol{x}} = \begin{pmatrix} ct \\ x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

G. Ludyk, *Quantum Mechanics in Matrix Form*,
https://doi.org/10.1007/978-3-319-26366-3_12

One consequence of the second postulate is the *invariance of spacetime intervals*. If a flash of light is emitted at the origin at $t = 0$, the spherical wavefront of the light at some later time Δt is described by

$$c^2 \Delta t^2 = \Delta x_1^2 + \Delta x_2^2 + \Delta x_3^2 = \Delta \boldsymbol{x}^\intercal \Delta \boldsymbol{x},$$

or

$$c^2 \Delta t^2 - \Delta \boldsymbol{x}^\intercal \Delta \boldsymbol{x} = 0. \tag{12.1}$$

With the *Minkowski matrix*

$$\boldsymbol{M} \overset{\text{def}}{=} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

and the vector

$$\Delta \vec{\boldsymbol{x}} = \begin{pmatrix} c\Delta t \\ \Delta x_1 \\ \Delta x_2 \\ \Delta x_3 \end{pmatrix},$$

this can be rewritten as a quadratic form

$$\Delta \vec{\boldsymbol{x}}^\intercal \boldsymbol{M} \Delta \vec{\boldsymbol{x}} = 0.$$

Due to the second postulate, this equation must also hold for an observer in another frame with the coordinates $\vec{\boldsymbol{x}}'$:

$$\Delta \vec{\boldsymbol{x}}'^\intercal \boldsymbol{M} \Delta \vec{\boldsymbol{x}}' = 0.$$

12.1.2 Lorentz Transformation

If an inertial reference frame $\vec{\boldsymbol{x}}'$ moves with respect to another inertial reference system $\vec{\boldsymbol{x}}$ with a constant speed $\boldsymbol{v}$, then the *Lorentz transformation* connects the two frames (see [8]):

$$\boldsymbol{L}(\boldsymbol{v}) \overset{\text{def}}{=} \left(\begin{array}{c|c} \gamma & -\frac{\gamma}{c}\, \boldsymbol{v}^\intercal \\ \hline -\frac{\gamma}{c}\, \boldsymbol{v} & \boldsymbol{I} + (\gamma - 1) \frac{\boldsymbol{v}\,\boldsymbol{v}^\intercal}{v^2} \end{array} \right), \tag{12.2}$$

where the factor γ is given by

$$\gamma = (1 - v^2/c^2)^{-1/2}.$$

The transformation is then given by

$$\vec{\boldsymbol{x}}' = \boldsymbol{L}(\boldsymbol{v})\vec{\boldsymbol{x}}.$$

Componentwise, this reads

$$c\,t' = \gamma\,c\,t - \frac{\gamma}{c}\,\boldsymbol{v}^{\mathsf{T}}\boldsymbol{x},$$

and

$$\boldsymbol{x}' = \boldsymbol{x} + (\gamma - 1)\frac{\boldsymbol{v}^{\mathsf{T}}\boldsymbol{x}}{v^2}\,\boldsymbol{v} - \gamma\,\boldsymbol{v}\,t. \tag{12.3}$$

Note that these relations assume the coordinate axes of both systems to be parallel.

12.1.3 Velocity and Its Lorentz Transformation

If we define the four-dimensional velocity vector

$$\vec{\boldsymbol{u}} \overset{\text{def}}{=} \gamma_u \begin{pmatrix} c \\ \boldsymbol{u} \end{pmatrix} \tag{12.4}$$

with

$$\gamma_u \overset{\text{def}}{=} \frac{1}{\sqrt{1 - \frac{u^2}{c^2}}},$$

the Lorentz transformed velocity vector $\vec{\boldsymbol{u}}'$ is [15]

$$\vec{\boldsymbol{u}}' = \boldsymbol{L}\,\vec{\boldsymbol{u}}, \tag{12.5}$$

with

$$\vec{\boldsymbol{u}}' = \gamma_{u'} \begin{pmatrix} c \\ \boldsymbol{u}' \end{pmatrix}.$$

Inasmuch as such a velocity vector $\vec{\boldsymbol{u}}$ looks the same in every inertial frame, it is much better suited for the formulation of physical laws. The quadratic form for the velocity is given by

$$\vec{\boldsymbol{u}}^{\mathsf{T}}\boldsymbol{M}\vec{\boldsymbol{u}} = \gamma_u^2 c^2 - \gamma_u^2 \boldsymbol{u}^{\mathsf{T}}\boldsymbol{u} = \frac{c^4}{c^2 - u^2} - \frac{c^2 u^2}{c^2 - u^2} = c^2, \tag{12.6}$$

invariant with respect to Lorentz transformations, because $\vec{\boldsymbol{u}}'^{\mathsf{T}}\boldsymbol{M}\vec{\boldsymbol{u}}' = c^2$ holds as well.

12.1.4 Momentum and Its Lorentz Transformation

Multiplying (12.5) with the rest mass m_0, we obtain

$$\begin{pmatrix} m_0\gamma_{u'}c \\ m_0\gamma_{u'}\boldsymbol{u}' \end{pmatrix} = \boldsymbol{L}\begin{pmatrix} m_0\gamma_u c \\ m_0\gamma_u\boldsymbol{u} \end{pmatrix}. \tag{12.7}$$

We define the *momentum* as usual:

$$\boldsymbol{p} \overset{\text{def}}{=} m_0\gamma_u\boldsymbol{u} = m_u\boldsymbol{u} = m_u\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}, \tag{12.8}$$

where

$$m_u \overset{\text{def}}{=} m_0\gamma_u = \frac{m_0}{\sqrt{1-\frac{u^2}{c^2}}}.$$

The four-dimensional momentum vector

$$\vec{\boldsymbol{p}} \overset{\text{def}}{=} \begin{pmatrix} m_u\, c \\ \boldsymbol{p} \end{pmatrix} = m_0\,\vec{\boldsymbol{u}} = m_0\gamma_u\begin{pmatrix} c \\ \boldsymbol{u} \end{pmatrix} \tag{12.9}$$

transforms according to (12.7) as

$$\vec{\boldsymbol{p}}' = \boldsymbol{L}\,\vec{\boldsymbol{p}}. \tag{12.10}$$

Also, the quadratic form associated with the momentum vector

$$\vec{\boldsymbol{p}}^{\top}\boldsymbol{M}\vec{\boldsymbol{p}} = m_0^2\vec{\boldsymbol{u}}^{\top}\boldsymbol{M}\vec{\boldsymbol{u}} = m_0^2c^2 \tag{12.11}$$

is invariant with respect to Lorentz transformations, because we also find

$$\vec{\boldsymbol{p}}'^{\top}\boldsymbol{M}\vec{\boldsymbol{p}}' = m_0^2c^2.$$

12.1.5 Equation of Motion and Force

The relativistic equation of motion for a particle has to be Lorentz invariant. Additionally, Newton's equation of motion has to hold in the inertial frame of the considered particle:

$$m_0\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t} = \boldsymbol{f} \in \mathbb{R}^3. \tag{12.12}$$

Let the respective inertial system be $\mathcal{X}$. Furthermore, suppose that $\mathcal{X}'$ is an inertial system that moves relative to $\mathcal{X}$ with the constant speed $\boldsymbol{u}(t_0)$. Then the particle rests momentarily at the time $t = t_0$ in $\mathcal{X}'$. Now note that the equation of motion (12.12) refers to a point of time and its neighborhood. For the neighborhood $t = t_0 \pm \mathrm{d}t$, the speed in $\mathcal{X}'$ is arbitrarily small. For speeds $v \ll c$, (12.12) holds. Hence in $\mathcal{X}'$,

$$m_0 \frac{\mathrm{d}\boldsymbol{u}'}{\mathrm{d}t'} = \boldsymbol{f}' \in \mathbb{R}^3 \tag{12.13}$$

also holds exactly. Here, m_0 is the rest mass and $\boldsymbol{f}'$ the three-dimensional force in $\mathcal{X}'$. From (12.13), we can derive the relativistic equations of motion in an arbitrary reference frame.

Expanding the vector $\boldsymbol{f}'$ in (12.13) to a four-vector and calling the result $\vec{\boldsymbol{f}}'$, we get

$$m_0 \frac{\mathrm{d}}{\mathrm{d}t'} \begin{pmatrix} c \\ \boldsymbol{u}' \end{pmatrix} = \begin{pmatrix} 0 \\ \boldsymbol{f}' \end{pmatrix} \overset{\text{def}}{=} \vec{\boldsymbol{f}}'. \tag{12.14}$$

Note that $\vec{\boldsymbol{f}}'$ is specified in the rest system $\mathcal{X}'$. In the inertial system $\mathcal{X}$ (in which the particle moves with the velocity $\boldsymbol{u}$), $\vec{\boldsymbol{f}}$ is obtained by a Lorentz transformation $\boldsymbol{L}(-\boldsymbol{u})$:

$$\vec{\boldsymbol{f}} = \boldsymbol{L}(-\boldsymbol{u}) \begin{pmatrix} 0 \\ \boldsymbol{f}' \end{pmatrix} = \begin{pmatrix} \frac{\gamma_u}{c} \boldsymbol{u}^\mathsf{T} \boldsymbol{f}' \\ \boldsymbol{A}(\boldsymbol{u}) \boldsymbol{f}' \end{pmatrix} \overset{\text{def}}{=} \begin{pmatrix} f_0 \\ \boldsymbol{f} \end{pmatrix}, \tag{12.15}$$

where

$$\boldsymbol{A}(\boldsymbol{u}) \overset{\text{def}}{=} \boldsymbol{I} + (\gamma_u - 1) \frac{\boldsymbol{u}\boldsymbol{u}^\mathsf{T}}{u^2}$$

is the known part from the Lorentz matrix. Finally, the equation

$$m_0 \gamma \frac{\mathrm{d}}{\mathrm{d}t} \begin{pmatrix} \gamma c \\ \gamma \boldsymbol{u} \end{pmatrix} = \begin{pmatrix} f_0 \\ \boldsymbol{f} \end{pmatrix};$$

that is,

$$m_0 \vec{\boldsymbol{a}} = \vec{\boldsymbol{f}}, \tag{12.16}$$

has all the properties that we need. Namely, the four-vectors $\vec{\boldsymbol{a}}$ and $\vec{\boldsymbol{f}}$ are Lorentz-invariant and the equation changes into Newton's equation of motion in the rest frame of the particle:

$$m_0 \begin{pmatrix} 0 \\ \frac{\mathrm{d}\boldsymbol{u}'}{\mathrm{d}t'} \end{pmatrix} = \begin{pmatrix} 0 \\ \boldsymbol{f}' \end{pmatrix}.$$

For the last three components of the equation of motion (12.16), we get

$$\frac{\mathrm{d}(m_u \boldsymbol{u})}{\mathrm{d}t} = \frac{1}{\gamma_u} \boldsymbol{f}, \tag{12.17}$$

with the velocity-dependent mass

$$m_u \overset{\text{def}}{=} \gamma_u m_0. \tag{12.18}$$

In the theory of relativity, the time derivative of the momentum $m_u\,\boldsymbol{u}$ is also interpreted as a force. According to (12.15) and (12.16), the components f_i of the relativistic equation of motion are thus

$$\underline{\underline{f_0 = \gamma_u \frac{\mathrm{d}}{\mathrm{d}t}(m_u\,c) = \frac{\gamma_u}{c}\boldsymbol{u}^\mathsf{T}\boldsymbol{f}'}} \tag{12.19}$$

and

$$\underline{\underline{\boldsymbol{f} = \gamma_u \frac{\mathrm{d}}{\mathrm{d}t}(m_u\,\boldsymbol{u}) = \boldsymbol{A}(\boldsymbol{u})\,\boldsymbol{f}'}}. \tag{12.20}$$

12.1.6 Energy and Rest Mass

Equation (12.19) multiplied with c/γ_u provides

$$\frac{\mathrm{d}}{\mathrm{d}t}(m_u c^2) = \boldsymbol{u}^\mathsf{T}\boldsymbol{f}, \tag{12.21}$$

where $\boldsymbol{u}^\mathsf{T}\boldsymbol{f}$ is the instantaneous power, that is, the work per time unit accomplished by the force $\boldsymbol{f}$. Therefore, the left-hand side of (12.21) must be the temporal change of energy, which requires that $m_u c^2 = \gamma_u m_0 c^2$ is an energy. For the *relativistic energy*, we obtain the formula

$$E = m_u c^2. \tag{12.22}$$

For $\boldsymbol{u} = \boldsymbol{0}$ (i.e., the particle is at rest), we find $\gamma_u = 1$ and therefore

$$\underline{\underline{E_0 = m_0 c^2}}. \tag{12.23}$$

This is Einstein's famous formula for the "rest energy".

The four-dimensional momentum vector $\vec{\boldsymbol{p}}$ can now be recognized as a combination of energy and momentum:

$$\vec{\boldsymbol{p}} = \begin{pmatrix} E/c \\ \boldsymbol{p} \end{pmatrix}. \tag{12.24}$$

The quadratic form (12.11),

$$\vec{\boldsymbol{p}}^\mathsf{T}\boldsymbol{M}\vec{\boldsymbol{p}} = m_0^2 c^2,$$

now reads

$$\vec{\boldsymbol{p}}^\mathsf{T}\boldsymbol{M}\vec{\boldsymbol{p}} = \left(E/c,\, \boldsymbol{p}^\mathsf{T}\right)\begin{pmatrix} E/c \\ -\boldsymbol{p} \end{pmatrix} = E^2/c^2 - p^2 = m_0^2 c^2,$$

that is, Einstein's *relativistic energy-momentum relationship*

$$\underline{\underline{E^2 = p^2c^2 + m_0^2c^4.}} \tag{12.25}$$

12.2 The Dirac Equation

12.2.1 The Wave Equation for a Free Particle

In classical mechanics, a moving particle has the energy

$$E = \frac{p^2}{2m}, \tag{12.26}$$

where

$$p^2 = \boldsymbol{p}^\mathsf{T}\boldsymbol{p}, \quad \boldsymbol{p} \in \mathbb{R}^3.$$

On the other hand, we saw in Sect. 11.3 that (12.26) transforms into the Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\psi(t, \boldsymbol{x}) = -\frac{\hbar^2}{2m}\nabla^2\psi(t, \boldsymbol{x}) \tag{12.27}$$

for the wave function $\psi(t, \boldsymbol{x})$. The temporal and spatial derivatives occur in different orders, therefore the Schrödinger equation cannot be invariant under Lorentz transformations. Rather, its structure changes during transition from one inertial system to another. The Schrödinger equation is a nonrelativistic approximation, for low velocities, of the fully relativistic equation of motion.

Still, let us note the two formal correspondences (substitutions)

$$E \to i\hbar\frac{\partial}{\partial t} \quad \text{and} \quad \boldsymbol{p} \to -i\hbar\boldsymbol{\nabla} \tag{12.28}$$

that transformed (12.26) into (12.27). If we now consider a relativistic quantum particle and assume the same correspondence in the energy-momentum relationship (12.25), we obtain the so-called *Klein–Gordon equation*

$$\boxed{\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right)\psi(t, \boldsymbol{x}) = \frac{m_0^2c^2}{\hbar^2}\psi(t, \boldsymbol{x})}. \tag{12.29}$$

The Hamiltonian is obtained from the relationship (12.25) as

$$E = H = \pm\sqrt{p^2c^2 + m_0^2c^4}, \tag{12.30}$$

where m_0 is the rest mass of the moving particle. However, the minus sign causes a problem. After all, what is a negative energy? At the end of this chapter we learn more about Dirac's interpretation of this negative energy. Still, the solutions of the Klein–Gordon equation can even contain negative probability densities!

Let us first turn to yet another problem, namely how to calculate the square root of an operator. Dirac made the ingenious ansatz

$$\sqrt{p^2c^2 + m_0^2c^4} = c\left(\alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \beta m_0 c\right). \tag{12.31}$$

Squaring this ansatz gives

$$\begin{aligned} p^2c^2 + m_0^2c^4 = c^2\left(\alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \beta m_0 c\right)^2 = \\ = c^2\left(\alpha_1^2 p_1^2 + \alpha_2^2 p_2^2 + \alpha_3^2 p_3^2 + \beta^2 m_0^2 c^2 + \right. \\ +\alpha_1\alpha_2 p_1 p_2 + \alpha_1\alpha_3 p_1 p_3 + \alpha_2\alpha_3 p_2 p_3 + \\ +\alpha_2\alpha_1 p_2 p_1 + \alpha_3\alpha_1 p_3 p_1 + \alpha_3\alpha_2 p_3 p_2 + \\ +\alpha_1 p_1 \beta m_0 c + \alpha_2 p_2 \beta m_0 c + \alpha_3 p_3 \beta m_0 c + \\ \left. +\beta m_0 c \alpha_1 p_1 + \beta m_0 c \alpha_2 p_2 + \beta m_0 c \alpha_3 p_3\right). \end{aligned} \tag{12.32}$$

Under which conditions for the $\alpha's$ and β is (12.32) valid? The conditions are

$$\alpha_1^2 = \alpha_2^2 = \alpha_3^2 = \beta^2 = 1, \tag{12.33}$$

$$\alpha_1\alpha_2 + \alpha_2\alpha_1 = \alpha_1\alpha_3 + \alpha_3\alpha_1 = \alpha_2\alpha_3 + \alpha_3\alpha_2 = 0, \tag{12.34}$$

$$\alpha_1\beta + \beta\alpha_1 = \alpha_2\beta + \beta\alpha_2 = \alpha_3\beta + \beta\alpha_3 = 0. \tag{12.35}$$

In other words: The α's all anti-commute with one another and with β, and their square is unity. These properties of the $\alpha's$ and β suggest that they are not just (real or complex) numbers. Now recall that the Pauli 2 × 2 matrices satisfy very similar conditions. Therefore, let us assume that the $\boldsymbol{\alpha}'s$ and $\boldsymbol{\beta}$ are $n \times n$ matrices,[1] where n is initially unknown. We obtain for the energy

$$E = \pm c\left(\boldsymbol{\alpha}_1 p_1 + \boldsymbol{\alpha}_2 p_2 + \boldsymbol{\alpha}_3 p_3 + \boldsymbol{\beta} m_0 c\right). \tag{12.36}$$

Inasmuch as the squares of all matrices are equal to the unit matrix, all eigenvalues must be equal to ±1. Also, remember that the sum of the diagonal elements of a matrix is equal to the sum of its eigenvalues. Now, however,

---

[1] Written now in boldface letters.

$$\boldsymbol{\alpha}_i = \boldsymbol{\alpha}_i \boldsymbol{\beta}^2 = \boldsymbol{\alpha}\boldsymbol{\beta}\boldsymbol{\beta} = -\boldsymbol{\beta}\boldsymbol{\alpha}\boldsymbol{\beta},$$

and therefore

$$\text{trace}\,\boldsymbol{\alpha}_i = \text{trace}(-\boldsymbol{\beta}\boldsymbol{\alpha}_i\boldsymbol{\beta}) = -\text{trace}(\boldsymbol{\alpha}_i\boldsymbol{\beta}^2) = -\text{trace}\,\boldsymbol{\alpha}_i. \tag{12.37}$$

Note that we exploited the fact that the trace operation is *cyclic*; that is,

$$\text{trace}(\boldsymbol{\alpha}\boldsymbol{\beta}\boldsymbol{\gamma}) = \text{trace}(\boldsymbol{\beta}\boldsymbol{\gamma}\boldsymbol{\alpha}).$$

Also, we used the fact that

$$\text{trace}(a\,\boldsymbol{\alpha}) = a\,\text{trace}(\boldsymbol{\alpha}),\, a \in \mathbb{C}.$$

Equation (12.37) can only be true if trace $\boldsymbol{\alpha}_i = 0$, which is possible only if the number of $+1$ eigenvalues is equal to the number of -1 eigenvalues. Hence, n must be an even number. Now how can we make use of the three Pauli matrices

$$\boldsymbol{\sigma}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \boldsymbol{\sigma}_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \text{and} \quad \boldsymbol{\sigma}_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

for finding a solution to our problem? Perhaps finding a fourth matrix that anticommutes with these three? It turns out, however, that this is impossible. Therefore let us move on to the next even number, namely $n = 4$. In other words, we are now looking for 4×4 matrices that satisfy our conditions (12.33)–(12.35). One possibility for the $\boldsymbol{\alpha}'s$ and $\boldsymbol{\beta}$ are the four matrices called *Dirac matrices:*

$$\boldsymbol{\alpha}_1 = \boldsymbol{\sigma}_1 \otimes \boldsymbol{\sigma}_1 = \begin{pmatrix} \mathbf{0} & \boldsymbol{\sigma}_1 \\ \boldsymbol{\sigma}_1 & \mathbf{0} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \tag{12.38}$$

$$\boldsymbol{\alpha}_2 = \boldsymbol{\sigma}_1 \otimes \boldsymbol{\sigma}_2 = \begin{pmatrix} \mathbf{0} & \boldsymbol{\sigma}_2 \\ \boldsymbol{\sigma}_2 & \mathbf{0} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}, \tag{12.39}$$

$$\boldsymbol{\alpha}_3 = \boldsymbol{\sigma}_1 \otimes \boldsymbol{\sigma}_3 = \begin{pmatrix} \mathbf{0} & \boldsymbol{\sigma}_3 \\ \boldsymbol{\sigma}_3 & \mathbf{0} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \tag{12.40}$$

$$\boldsymbol{\beta} = \boldsymbol{\sigma}_3 \otimes \boldsymbol{I}_2 = \begin{pmatrix} \boldsymbol{I}_2 & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{I}_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{12.41}$$

If we interpret (12.36) as the Hamiltonian and use the known relation $E = i\hbar\frac{\partial}{\partial t}\boldsymbol{\psi}(\boldsymbol{x}, t)$, we obtain the *Dirac equation*

$$\boxed{i\hbar\frac{\partial}{\partial t}\boldsymbol{\psi}(\boldsymbol{x}, t) = c\left(\boldsymbol{\alpha}_1\boldsymbol{p}_1 + \boldsymbol{\alpha}_2\boldsymbol{p}_2 + \boldsymbol{\alpha}_3\boldsymbol{p}_3 + \boldsymbol{\beta} m_0 c\right)\boldsymbol{\psi}(\boldsymbol{x}, t)}. \tag{12.42}$$

Because the matrices $\boldsymbol{\alpha}_i$ and $\boldsymbol{\beta}$ are 4 × 4 matrices, the *Dirac wave function* must be a four-column vector, also called a *Dirac spinor*:

$$\boldsymbol{\psi}(\boldsymbol{x}, t) = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix}.$$

Note that the $\boldsymbol{p}_i$ are also represented as 4 × 4 matrices

$$\boldsymbol{p}_i = \begin{pmatrix} p_i & 0 & 0 & 0 \\ 0 & p_i & 0 & 0 \\ 0 & 0 & p_i & 0 \\ 0 & 0 & 0 & p_i \end{pmatrix}.$$

12.2.2 Invariant Form of the Dirac Equation

Multiplying the Dirac equation (12.42) from the left by the matrix $\boldsymbol{\beta}$, one obtains with $\boldsymbol{\beta\beta} = \boldsymbol{I}_4$

$$i\hbar\boldsymbol{\beta}\frac{\partial}{\partial t}\boldsymbol{\psi}(\boldsymbol{x}, t) = c\left(\boldsymbol{\beta\alpha}_1\boldsymbol{p}_1 + \boldsymbol{\beta\alpha}_2\boldsymbol{p}_2 + \boldsymbol{\beta\alpha}_3\boldsymbol{p}_3 + m_0 c\boldsymbol{I}_4\right)\boldsymbol{\psi}(\boldsymbol{x}, t). \tag{12.43}$$

If we define the alternative Dirac matrices

$$\gamma_0 \overset{\text{def}}{=} \boldsymbol{\beta} \quad \text{and} \quad \gamma_j \overset{\text{def}}{=} \boldsymbol{\beta\alpha}_j \quad (j = 1, 2, 3), \tag{12.44}$$

the Dirac equation looks like

$$i\hbar\frac{\partial}{\partial t}\gamma_0\boldsymbol{\psi}(\boldsymbol{x}, t) = c\left(\gamma_1\boldsymbol{p}_1 + \gamma_2\boldsymbol{p}_2 + \gamma_3\boldsymbol{p}_3 + m_0 c\boldsymbol{I}_4\right)\boldsymbol{\psi}(\boldsymbol{x}, t). \tag{12.45}$$

If we set $p_i \overset{\text{def}}{=} -i\hbar \frac{\partial}{\partial x_i}$ and $x_0 = c\,t$, we get the equation

$$\left(i\hbar \sum_{i=0}^{3} \gamma_i \frac{\partial}{\partial x_i} - m_0 \boldsymbol{I}_4 \right) \boldsymbol{\psi}(\boldsymbol{x}, t) = \boldsymbol{0}. \tag{12.46}$$

With the Feynman[2] dash-notation

$$\not{\partial} \overset{\text{def}}{=} \sum_{I=0}^{3} \gamma_i \frac{\partial}{\partial x_i},$$

where $\not{\partial}$ is a 4×4 matrix, we finally get the compact Feynman notation of Dirac's equation:

$$\boxed{(i\not{\partial} - \frac{m_0}{\hbar} \boldsymbol{I}_4) \boldsymbol{\psi}(\boldsymbol{x}, t) = \boldsymbol{0}}. \tag{12.47}$$

A detailed proof of the invariance of the Dirac equation under Lorentz transformations can be found, for example, in Klaus Schulten's "Notes on Quantum Mechanics" (freely available for download at the website of the University of Illinois).

12.2.3 Solution of the Dirac Equation

The Dirac equation is a system of four linear differential equations. As an ansatz, we try the four wave functions

$$\boldsymbol{\psi} = \boldsymbol{\theta} \exp\left(i(\boldsymbol{p}^\intercal \boldsymbol{x}/\hbar - \omega t) \right) \tag{12.48}$$

as eigenfunctions of energy and momentum with the eigenvalues $E = \hbar\omega$, p_1, p_2, and p_3. If we insert the ansatz (12.48) in the Dirac equation (12.42), we obtain the algebraic equation

$$E\boldsymbol{\theta} = c\left(\boldsymbol{\alpha}_1 \boldsymbol{p}_1 + \boldsymbol{\alpha}_2 \boldsymbol{p}_2 + \boldsymbol{\alpha}_3 \boldsymbol{p}_3 + \boldsymbol{\beta} m_0 c \right) \boldsymbol{\theta}. \tag{12.49}$$

Due to the block structure of the $\boldsymbol{\alpha}$-matrices and of $\boldsymbol{\beta}$, it is useful to split the Dirac spinor $\boldsymbol{\psi}$ into two two-dimensional spinors $\boldsymbol{\chi}$ and $\boldsymbol{\eta}$:

$$\boldsymbol{\psi}(\boldsymbol{x}, t) = \begin{pmatrix} \boldsymbol{\chi} \\ \boldsymbol{\eta} \end{pmatrix} = \left(\begin{array}{c} \psi_1 \\ \psi_2 \\ \hline \psi_3 \\ \psi_4 \end{array} \right). \tag{12.50}$$

---

[2] Richard Phillips Feynman, 1918–1988, American physicist, Nobel Prize 1965.

In the solution (12.48), we similarly split the vector

$$\boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{\xi} \\ \boldsymbol{\zeta} \end{pmatrix} = \left(\begin{array}{c} \theta_1 \\ \theta_2 \\ \hline \theta_3 \\ \theta_4 \end{array}\right). \tag{12.51}$$

Also, we introduce the notation

$$\boldsymbol{\sigma}\cdot\bar{\boldsymbol{p}} \overset{\text{def}}{=} \boldsymbol{\sigma_1}\bar{\boldsymbol{p}}_1 + \boldsymbol{\sigma_2}\bar{\boldsymbol{p}}_2 + \boldsymbol{\sigma_3}\bar{\boldsymbol{p}}_3, \tag{12.52}$$

where

$$\bar{\boldsymbol{p}}_i \overset{\text{def}}{=} \begin{pmatrix} p_i & 0 \\ 0 & p_i \end{pmatrix}.$$

The algebraic Dirac equation (12.49) now reads

$$E\begin{pmatrix} \boldsymbol{\xi} \\ \boldsymbol{\zeta} \end{pmatrix} = c\begin{pmatrix} \boldsymbol{0} & \boldsymbol{\sigma}\cdot\bar{\boldsymbol{p}} \\ \boldsymbol{\sigma}\cdot\bar{\boldsymbol{p}} & \boldsymbol{0} \end{pmatrix}\begin{pmatrix} \boldsymbol{\xi} \\ \boldsymbol{\zeta} \end{pmatrix} + m_0c^2\begin{pmatrix} \boldsymbol{I}_2 & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{I}_2 \end{pmatrix}\begin{pmatrix} \boldsymbol{\xi} \\ \boldsymbol{\zeta} \end{pmatrix}. \tag{12.53}$$

Rearranging yields

$$(E - m_0c^2)\boldsymbol{\xi} = c(\boldsymbol{\sigma}\cdot\bar{\boldsymbol{p}})\,\boldsymbol{\zeta}, \tag{12.54}$$

$$(E + m_0c^2)\,\boldsymbol{\zeta} = c(\boldsymbol{\sigma}\cdot\bar{\boldsymbol{p}})\,\boldsymbol{\xi}. \tag{12.55}$$

If we multiply (12.54) with $(E + m_0c^2)$ and insert (12.55), we obtain

$$(E^2 - m_0^2c^4)\,\boldsymbol{\xi} = c^2(\boldsymbol{\sigma}\cdot\bar{\boldsymbol{p}})^2\boldsymbol{\xi}. \tag{12.56}$$

For the product $\boldsymbol{\sigma}\cdot\bar{\boldsymbol{p}}$, we can show that

$$(\boldsymbol{\sigma}\cdot\bar{\boldsymbol{p}})^2 = p^2, \tag{12.57}$$

which brings us back to the known condition

$$E^2 = m_0^2c^4 + c^2p^2. \tag{12.58}$$

With the definitions

$$p_+ \overset{\text{def}}{=} p_1 + ip_2 \quad \text{and} \quad p_- \overset{\text{def}}{=} p_1 - ip_2, \tag{12.59}$$

the two (12.54) and (12.55) are then

$$(E - m_0c^2)\theta_1 = c(p_3\theta_3 + p_-\theta_4),$$

$$(E - m_0c^2)\theta_2 = c(p_+\theta_3 + p_3\theta_4),$$

$$(E + m_0c^2)\theta_3 = c(p_3\theta_1 + p_-\theta_2),$$

$$(E + m_0c^2)\theta_4 = c(p_+\theta_1 + p_3\theta_2).$$

If θ_3 and θ_4 are given, these four equations determine θ_1 and θ_2 (and vice versa). For a given p with $E = +\sqrt{c^2p^2 + m_0^2c^4}$, there are therefore two independent solutions of the four equations, namely

$$\begin{pmatrix} 1 \\ 0 \\ \frac{cp_3}{E+m_0c^2} \\ \frac{cp_+}{E+m_0c^2} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ 1 \\ \frac{cp_-}{E+m_0c^2} \\ \frac{-cp_3}{E+m_0c^2} \end{pmatrix}. \tag{12.60}$$

These solutions represent the two *spin states* of an electron with the given momentum p, as is physically required. This becomes particularly clear if we consider the nonrelativistic limit $v \ll c$ of the solution. In this case, we get the two vectors (see Chap. 8)

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

12.2.4 Dirac's Interpretation of the Negative Energy

Dirac developed the idea that the states in an atom with negative energies are already occupied with electrons, like the filled electron shells in the Pauli exclusion scheme. If the states with negative energies were unoccupied, all electrons would fall "down" into these states, which would lead to a release of huge amounts of energy. Because this does not happen in reality, these states must already be occupied. Countless numbers of electrons fill the so-called "Dirac sea" of negative energies and make sure that all such states are occupied. The vacuum state is characterized by the fact that all the negative energy states are filled. Based on this idea of occupied negative energy states, Dirac developed his theory of *holes*. According to this theory, an electron that absorbs a photon can make a transition from a negative energy state to a state with positive energy, leaving behind a state with a positive charge that can move like a particle. Casually speaking, this state is a piece of *anti-matter*. In other words, Dirac predicted the positron! In 1932, Anderson[3] finally discovered the positron with a positive electric charge in cosmic rays.

12.3 Problems

12.1 Gamma Matrices: What is the form of the gamma matrices and what relationships exist between them?

12.2 Dirac Equation: Solve the Dirac equation

$$(i\not{\partial} - \frac{m_0}{\hbar}\boldsymbol{I}_4)\boldsymbol{\psi}(\boldsymbol{x}, t) = \boldsymbol{0} \tag{12.61}$$

with the ansatz

$$\boldsymbol{\psi}(\boldsymbol{x}, t) = \boldsymbol{c}\, \exp\big(i(\boldsymbol{k}^{\mathsf{T}}\boldsymbol{x} - \omega t)\big). \tag{12.62}$$

---

[3] Carl David Anderson, 1905–1991, American physicist, Nobel Prize 1936.

Appendix A
Solutions to Problems

2.1. Given the two Hermite matrices $\boldsymbol{A} = \boldsymbol{A}^\dagger$ and $\boldsymbol{B} = \boldsymbol{B}^\dagger$, we find $(\boldsymbol{AB})^\dagger = \boldsymbol{B}^\dagger \boldsymbol{A}^\dagger = \boldsymbol{BA}$. This is equal to $\boldsymbol{AB}$ only if the matrices commute; that is, $\boldsymbol{AB} = \boldsymbol{BA}$.

2.2. Let $\boldsymbol{A}\boldsymbol{e}_1 = \lambda_1 \boldsymbol{e}_1$ and $\boldsymbol{A}\boldsymbol{e}_2 = \lambda_2 \boldsymbol{e}_2$ for $\lambda_1 \neq \lambda_2$. Proof by contradiction. We assume that there are $c_1, c_2 \neq 0$ such that

$$c_1 \boldsymbol{e}_1 + c_2 \boldsymbol{e}_2 = \boldsymbol{0}. \tag{A.1}$$

Multiplying this equation by $\boldsymbol{A}$ and taking into account the above eigenvalue equations, we receive

$$c_1 \lambda_1 \boldsymbol{e}_1 + c_2 \lambda_2 \boldsymbol{e}_2 = \boldsymbol{0}. \tag{A.2}$$

If we multiply (A.1) by λ_1 and subtract the result from (A.2), we obtain

$$c_2(\lambda_2 - \lambda_1)\boldsymbol{e}_2 = \boldsymbol{0}.$$

Because $\lambda_1 \neq \lambda_2$ and $\boldsymbol{e}_2 \neq \boldsymbol{0}$, it must be $c_2 = 0$, which contradicts the assumption.

2.3. An Hermitian matrix satisfies the equation

$$\boldsymbol{A}\boldsymbol{e} = \lambda \boldsymbol{e}.$$

Because $\boldsymbol{A}^\dagger = \boldsymbol{A}$, we have

$$(\boldsymbol{A}\boldsymbol{e})^\dagger \boldsymbol{e} = \boldsymbol{e}^\dagger \underbrace{\boldsymbol{A}^\dagger \boldsymbol{e}}_{\lambda^* \boldsymbol{e}} = \boldsymbol{e}^\dagger \underbrace{\boldsymbol{A}\boldsymbol{e}}_{\lambda \boldsymbol{e}},$$

or $\lambda^* \boldsymbol{e}^\dagger \boldsymbol{e} = \lambda\, \boldsymbol{e}^\dagger \boldsymbol{e}$. Because $\boldsymbol{e}^\dagger \boldsymbol{e} > 0$, we have $\lambda^* = \lambda$; that is, the eigenvalue λ is real.

G. Ludyk, *Quantum Mechanics in Matrix Form*,
https://doi.org/10.1007/978-3-319-26366-3

2.4. For unitary matrices, we have $\boldsymbol{U}^\dagger\boldsymbol{U} = \boldsymbol{I}$. With $\boldsymbol{U}\boldsymbol{e} = \lambda\boldsymbol{e}$, we find from

$$(\boldsymbol{U}\boldsymbol{e})^\dagger(\boldsymbol{U}\boldsymbol{e}) = \boldsymbol{e}^\dagger\boldsymbol{U}^\dagger\boldsymbol{U}\boldsymbol{e} = \boldsymbol{e}^\dagger\boldsymbol{e}$$

that

$$\lambda^*\lambda\,\boldsymbol{e}^\dagger\boldsymbol{e} = \boldsymbol{e}^\dagger\boldsymbol{e},$$

thus $|\lambda| = 1$.

2.5. Take $\boldsymbol{a}_1$ and construct

$$\boldsymbol{e}_1 = \frac{\boldsymbol{a}_1}{\sqrt{\boldsymbol{a}_1^\dagger\boldsymbol{a}_1}}.$$

Then we get indeed

$$\boldsymbol{e}_1^\dagger\boldsymbol{e}_1 = \frac{\boldsymbol{a}_1^\dagger\boldsymbol{a}_1}{\sqrt{\boldsymbol{a}_1^\dagger\boldsymbol{a}_1}\sqrt{\boldsymbol{a}_1^\dagger\boldsymbol{a}_1}} = 1.$$

Now we subtract the vector $(\boldsymbol{e}_1^\dagger\boldsymbol{a}_2)\boldsymbol{e}_1$ from the vector $\boldsymbol{a}_2$ and normalize the result to get

$$\boldsymbol{e}_2 = \frac{\boldsymbol{a}_2 - (\boldsymbol{e}_1^\dagger\boldsymbol{a}_2)\boldsymbol{e}_1}{\sqrt{(\boldsymbol{a}_2 - (\boldsymbol{e}_1^\dagger\boldsymbol{a}_2)\boldsymbol{e}_1)^\dagger(\boldsymbol{a}_2 - (\boldsymbol{e}_1^\dagger\boldsymbol{a}_2)\boldsymbol{e}_1)}}.$$

We find again that

$$\boldsymbol{e}_2^\dagger\boldsymbol{e}_2 =$$

$$= \frac{\Big(\boldsymbol{a}_2 - (\boldsymbol{e}_1^\dagger\boldsymbol{a}_2)\boldsymbol{e}_1\Big)^\dagger\Big(\boldsymbol{a}_2 - (\boldsymbol{e}_1^\dagger\boldsymbol{a}_2)\boldsymbol{e}_1\Big)}{\Big(\boldsymbol{a}_2 - (\boldsymbol{e}_1^\dagger\boldsymbol{a}_2)\boldsymbol{e}_1\Big)^\dagger\Big(\boldsymbol{a}_2 - (\boldsymbol{e}_1^\dagger\boldsymbol{a}_2)\boldsymbol{e}_1\Big)} = 1.$$

Also, $\boldsymbol{e}_2$ is orthogonal to $\boldsymbol{e}_1$ because of

$$\boldsymbol{e}_1^\dagger\boldsymbol{e}_2 = \frac{\boldsymbol{a}_1^\dagger}{\sqrt{\boldsymbol{a}^\dagger\boldsymbol{a}_1}}\frac{\boldsymbol{a}_2 - (\boldsymbol{e}_1^\dagger\boldsymbol{a}_2)\boldsymbol{e}_1}{\sqrt{\Big(\boldsymbol{a}_2 - (\boldsymbol{e}_1^\dagger\boldsymbol{a}_2)\boldsymbol{e}_1\Big)^\dagger\Big(\boldsymbol{a}_2 - (\boldsymbol{e}_1^\dagger\boldsymbol{a}_2)\boldsymbol{e}_1\Big)}} =$$

$$= \frac{\boldsymbol{a}_1^\dagger\boldsymbol{a}_2 - \boldsymbol{a}_1^\dagger(\boldsymbol{e}_1^\dagger\boldsymbol{a}_2)\boldsymbol{e}_1}{\sqrt{\cdots}} = \frac{\boldsymbol{a}_1^\dagger\boldsymbol{a}_2 - \frac{\boldsymbol{a}_1^\dagger(\boldsymbol{a}_1^\dagger\boldsymbol{a}_2)\boldsymbol{a}_1}{\boldsymbol{a}_1^\dagger\boldsymbol{a}_1}}{\sqrt{\cdots}} = 0.$$

Similarly, we obtain the general formula

$$\boldsymbol{e}_j = \frac{\boldsymbol{a}_j - \sum_{i=1}^{j-1}(\boldsymbol{e}_i^\dagger\boldsymbol{a}_j)\boldsymbol{e}_i}{\sqrt{\Big(\boldsymbol{a}_j - \sum_{I=1}^{j-1}(\boldsymbol{e}_i^\dagger\boldsymbol{a}_j)\boldsymbol{e}_i\Big)^\dagger\Big(\boldsymbol{a}_j - \sum_{i=1}^{j-1}(\boldsymbol{e}_i^\dagger\boldsymbol{a}_j)\boldsymbol{e}_i\Big)}}\ .$$

2.6. (**a**)

$$\boldsymbol{T}\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix} \overset{!}{=} \begin{pmatrix}1\\0\end{pmatrix}$$

and

$$\boldsymbol{T}\begin{pmatrix}-1\\1\end{pmatrix} = \begin{pmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{pmatrix}\begin{pmatrix}-1\\1\end{pmatrix} \overset{!}{=} \begin{pmatrix}0\\1\end{pmatrix}$$

provides the four conditions

$$t_{11} + t_{12} = 1, \quad t_{21} + t_{22} = 0, \quad -t_{11} + t_{12} = 0, \quad -t_{21} + t_{22} = 1.$$

The transformation matrix is therefore given by

$$\underline{\underline{\boldsymbol{T} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}}}.$$

Because of $\boldsymbol{T}^{\dagger}\boldsymbol{T} = \boldsymbol{I}$, $\boldsymbol{T}$ is a unitary matrix.
(**b**) According to Problem 2.5, we get

$$\underline{\underline{\boldsymbol{e}_1}} = \frac{\boldsymbol{a}_1}{\sqrt{\boldsymbol{a}_1^{\dagger}\boldsymbol{a}_1}} = \underline{\underline{\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}}}$$

and

$$\underline{\underline{\boldsymbol{e}_2}} = \frac{\boldsymbol{a}_2 - (\boldsymbol{e}_1^{\dagger}\boldsymbol{a}_2)\boldsymbol{e}_1}{\sqrt{(\boldsymbol{a}_2 - (\boldsymbol{e}_1^{\dagger}\boldsymbol{a}_2)\boldsymbol{e}_1)^{\dagger}(\boldsymbol{a}_2 - (\boldsymbol{e}_1^{\dagger}\boldsymbol{a}_2)\boldsymbol{e}_1)}} =$$

$$= \frac{\begin{pmatrix}-1\\1\end{pmatrix} - 0 \cdot \boldsymbol{e}_1}{\sqrt{2}} = \underline{\underline{\begin{pmatrix} \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}}}.$$

3.1. We have

$$[\boldsymbol{A}, [\boldsymbol{B}, \boldsymbol{C}]] + [\boldsymbol{B}, [\boldsymbol{C}, \boldsymbol{A}]] + [\boldsymbol{C}, [\boldsymbol{A}, \boldsymbol{B}]] =$$

$$\begin{aligned}
&= [\boldsymbol{A}, \boldsymbol{BC} - \boldsymbol{CB}] + [\boldsymbol{B}, \boldsymbol{CA} - \boldsymbol{AC}]] + [\boldsymbol{C}, \boldsymbol{AB} - \boldsymbol{BA}] \\
&= \boldsymbol{ABC} - \boldsymbol{ACB} - \boldsymbol{BCA} + \boldsymbol{CBA} + \\
&\quad + \boldsymbol{BCA} - \boldsymbol{BAC} - \boldsymbol{CAB} + \boldsymbol{ACB} + \\
&\quad + \boldsymbol{CAB} - \boldsymbol{CBA} - \boldsymbol{ABC} + \boldsymbol{BAC} \\
&= \underline{\underline{\boldsymbol{0}}}.
\end{aligned}$$

3.2. If we multiply the relation

$$[\boldsymbol{X}, \boldsymbol{P}^n] = \boldsymbol{X}\boldsymbol{P}^n - \boldsymbol{P}^n\boldsymbol{X} = n\, i\hbar \boldsymbol{P}^{n-1}$$

with $\boldsymbol{P}$ from the left, we find

$$\boldsymbol{P}\boldsymbol{X}\boldsymbol{P}^n - \boldsymbol{P}^{n+1}\boldsymbol{X} = n\, i\hbar \boldsymbol{P}^n. \tag{A.3}$$

From

$$[\boldsymbol{X}, \boldsymbol{P}] = i\hbar \boldsymbol{I}$$

we get

$$\boldsymbol{P}\boldsymbol{X} = \boldsymbol{X}\boldsymbol{P} - i\hbar \boldsymbol{I}.$$

A substitution in (A.3) finally yields the assertion for $n+1$:

$$\boldsymbol{X}\boldsymbol{P}^{n+1} - \boldsymbol{P}^{n+1}\boldsymbol{X} = (n+1)\, i\hbar \boldsymbol{P}^n.$$

3.3. Because $(\boldsymbol{A}\boldsymbol{B})^\dagger = \boldsymbol{B}^\dagger\boldsymbol{A}^\dagger$ and $(\boldsymbol{A}+\boldsymbol{B})^\dagger = \boldsymbol{A}^\dagger + \boldsymbol{B}^\dagger$, we find

$$(\boldsymbol{A}\boldsymbol{B} + \boldsymbol{B}\boldsymbol{A})^\dagger = \boldsymbol{B}^\dagger\boldsymbol{A}^\dagger + \boldsymbol{A}^\dagger\boldsymbol{B}^\dagger = \boldsymbol{A}\boldsymbol{B} + \boldsymbol{B}\boldsymbol{A}.$$

3.4. With

$$\boldsymbol{H} = \boldsymbol{E} = \begin{pmatrix} E_1 & 0 & \cdots & \cdots \\ 0 & E_2 & 0 & \cdots \\ \vdots & \ddots & \ddots & \\ 0 & \cdots & 0 & E_N \end{pmatrix},$$

we can write

$$\boldsymbol{X}(t) = \exp(-it\boldsymbol{E}/\hbar)\boldsymbol{X}(0)\exp(it\boldsymbol{E}/\hbar) =$$

$$= \begin{pmatrix} e^{(-itE_1/\hbar)} & 0 & \cdots & 0 \\ 0 & e^{(-itE_2/\hbar)} & 0 & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & e^{(-itE_N/\hbar)} \end{pmatrix} \boldsymbol{X}(0)\exp(it\boldsymbol{E}/\hbar) =$$

$$= \begin{pmatrix} X_{11}(0) & X_{12}(0)e^{(-it(E_1-E_2)/\hbar)} & \cdots & X_{1N}(0)e^{(-it(E_1-E_N)/\hbar)} \\ \vdots & X_{22}(0) & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ X_{N1}(0)e^{(-it(E_N-E_1)/\hbar)} & \cdots & \cdots & X_{NN}(0) \end{pmatrix}.$$

This implies

$$\underline{\underline{X_{\mu\nu}(t) = e^{(it(E_\nu - E_\mu)/\hbar)} X_{\mu\nu}(0).}}$$

3.5. The matrix $\exp(i\boldsymbol{H})$ is a power series in $\boldsymbol{H}$, therefore we can easily show that

$$(\exp(i\boldsymbol{H}))^\dagger = \exp(-i\boldsymbol{H}^\dagger) = \exp(-i\boldsymbol{H}).$$

This leads to

$$(\exp(i\boldsymbol{H}))^\dagger\,(\exp(i\boldsymbol{H})) = \exp(-i\boldsymbol{H})\,(\exp(i\boldsymbol{H})) = \exp(\boldsymbol{0}) = \boldsymbol{I}.$$

In other words, $\exp(i\boldsymbol{H})$ is indeed a unitary matrix.

3.6. This question is equivalent to the question of whether $-\hbar\boldsymbol{C} = i[\boldsymbol{AB} - \boldsymbol{BA}]$ is Hermitian. We find that

$$\begin{aligned}(i[\boldsymbol{A},\boldsymbol{B}])^\dagger &= -i([\boldsymbol{AB} - \boldsymbol{BA}])^\dagger \\ &= -i((\boldsymbol{AB})^\dagger - (\boldsymbol{BA})^\dagger) \\ &= -i\boldsymbol{B}^\dagger\boldsymbol{A}^\dagger + i\boldsymbol{A}^\dagger\boldsymbol{B}^\dagger \\ &= -i\boldsymbol{BA} + i\boldsymbol{AB} \\ &= i[\boldsymbol{A},\boldsymbol{B}].\end{aligned} \tag{A.4}$$

Therefore, $-\hbar\boldsymbol{C}$ and thus $\hbar\boldsymbol{C} = \frac{1}{i}[\boldsymbol{A},\boldsymbol{B}]$ are indeed Hermitian. Note, however, that $[\boldsymbol{A},\boldsymbol{B}]$ is *anti-hermitian*!

3.7.

$$(a)\ [\boldsymbol{A},\boldsymbol{BC}] = \boldsymbol{ABC} - \boldsymbol{BCA} + \underbrace{\boldsymbol{BAC} - \boldsymbol{BAC}}_{\boldsymbol{0}} = \boldsymbol{B}[\boldsymbol{A},\boldsymbol{C}] + [\boldsymbol{A},\boldsymbol{B}]\boldsymbol{C}.$$

$$(b)\ [\boldsymbol{AB},\boldsymbol{C}] = \boldsymbol{ABC} - \boldsymbol{CAB} + \underbrace{\boldsymbol{ACB} - \boldsymbol{ACB}}_{\boldsymbol{0}} = \boldsymbol{A}[[\boldsymbol{B},\boldsymbol{C}] + [\boldsymbol{A},\boldsymbol{C}]\boldsymbol{B}.$$

3.8. A nilpotent matrix is a square matrix $\boldsymbol{N}$ such that $\boldsymbol{N}^k = \boldsymbol{0}$ for some positive integer k. In the present case, we find

$$\boldsymbol{N}^2 = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \boldsymbol{N}^3 = \boldsymbol{0},$$

which leads to

$$\exp(t\boldsymbol{N}) = \boldsymbol{I} + t\boldsymbol{N} + \frac{t^2}{2}\boldsymbol{N}^2 = \begin{pmatrix} 1 & t & t^2 \\ 0 & 1 & 2t \\ 0 & 0 & 1 \end{pmatrix}.$$

4.1. The momentum of the electron is

$$p = m_e \cdot v = (9.11 \cdot 10^{-31}\mathrm{kg})(10^3\,\mathrm{ms}^{-1}) = 9.11 \cdot 10^{-28}\,\mathrm{m\,kg\,s}^{-1}.$$

The percentage of the momentum accuracy is

$$\frac{\Delta p}{p} \cdot 100 = 0.1,$$

therefore

$$\Delta p = \frac{p \cdot 0.1}{100} = 9.11 \cdot 10^{-31}\,\mathrm{m\,kg\,s}^{-1}.$$

Heisenberg's uncertainty principle then yields

$$\Delta x \geq \frac{\hbar}{2\Delta p} = \frac{1.055 \cdot 10^{-34}\mathrm{Js}}{2 \cdot 9.11 \cdot 10^{-31}\mathrm{kg\,m\,s}^{-1}} = 0.0579 \cdot 10^{-3}\mathrm{m} = \underline{\underline{0.0579\,\mathrm{mm}}}.$$

4.2. A matrix is a projection matrix if (1) $\boldsymbol{P}$ is Hermitian and (2) $\boldsymbol{P}^2 = \boldsymbol{P}$. Let us check these conditions for the product matrix.

(1) Because $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ are projection matrices, they are Hermitian; that is, $\boldsymbol{P}_1 = \boldsymbol{P}_1^\dagger$ and $\boldsymbol{P}_2 = \boldsymbol{P}_2^\dagger$. Therefore

$$(\boldsymbol{P}_1\boldsymbol{P}_2)^\dagger = \boldsymbol{P}_2^\dagger\boldsymbol{P}_1^\dagger = \boldsymbol{P}_2 \cdot \boldsymbol{P}_1.$$

A necessary condition for the product matrix being Hermitian is therefore that $\boldsymbol{P}_1 \cdot \boldsymbol{P}_2 = \boldsymbol{P}_2 \cdot \boldsymbol{P}_1$; that is,

$$[\boldsymbol{P}_1, \boldsymbol{P}_2] = \boldsymbol{0}.$$

In other words, $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ must commute.

(2) If $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ commute, we have

$$(\boldsymbol{P}_1\boldsymbol{P}_2)^2 = \boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{P}_1\boldsymbol{P}_2 = \boldsymbol{P}_1(\boldsymbol{P}_2\boldsymbol{P}_1)\boldsymbol{P}_2 = \boldsymbol{P}_1(\boldsymbol{P}_1\boldsymbol{P}_2)\boldsymbol{P}_2 = \boldsymbol{P}_1^2\boldsymbol{P}_2^2 = \boldsymbol{P}_1\boldsymbol{P}_2.$$

We see that the second condition is automatically fulfilled if the two projection matrices commute.

4.3. **(1)** The density matrix is

$$\boldsymbol{D} = \boldsymbol{e}_{3,1}\boldsymbol{e}_{3,1}^\dagger = \begin{pmatrix} 1 \\ 0 \end{pmatrix} (1\ 0) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

and it is indeed $\boldsymbol{D}^2 = \boldsymbol{D}$.

(2) In this case, the density matrix is

$$\boldsymbol{D} = \boldsymbol{e}_{2,1}\boldsymbol{e}_{2,1}^{\dagger} = \frac{1}{2}\begin{pmatrix} 1 \\ i \end{pmatrix}(1 \ \ -i) = \begin{pmatrix} 0.5 & -0.5i \\ 0.5i & 0.5 \end{pmatrix},$$

and it is again $\boldsymbol{D}^2 = \boldsymbol{D}$.

In both cases $trace(\boldsymbol{D}) = 1$, therefore both systems are in a pure state.

4.4. Because

$$\boldsymbol{M}\boldsymbol{e}_1 = (\boldsymbol{e}_1\boldsymbol{e}_1^{\mathsf{T}} - \boldsymbol{e}_2\boldsymbol{e}_2^{\mathsf{T}})\boldsymbol{e}_1 = \boldsymbol{e}_1,$$

and

$$\boldsymbol{M}\boldsymbol{e}_2 = (\boldsymbol{e}_1\boldsymbol{e}_1^{\mathsf{T}} - \boldsymbol{e}_2\boldsymbol{e}_2^{\mathsf{T}})\boldsymbol{e}_2 = -\boldsymbol{e}_2,$$

the eigenvectors are $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$, and the eigenvalues are $+1$ and -1.

5.1. With

$$\boldsymbol{X} = \sqrt{\frac{\hbar}{2m\omega_0}} \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & \cdots \\ 1 & 0 & \sqrt{2} & 0 & 0 & \cdots \\ 0 & \sqrt{2} & 0 & \sqrt{3} & 0 & \cdots \\ 0 & 0 & \sqrt{3} & 0 & \sqrt{4} & \cdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$

and

$$\boldsymbol{P} = i\sqrt{\frac{\hbar m\omega_0}{2}} \begin{pmatrix} 0 & -1 & 0 & 0 & 0 & \cdots \\ 1 & 0 & -\sqrt{2} & 0 & 0 & \cdots \\ 0 & \sqrt{2} & 0 & -\sqrt{3} & 0 & \cdots \\ 0 & 0 & \sqrt{3} & 0 & -\sqrt{4} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

we get

$$\boldsymbol{X}\boldsymbol{P} = i\frac{\hbar}{2} \begin{pmatrix} 1 & 0 & -\sqrt{2} & 0 & 0 & \cdots \\ 0 & 1 & 0 & -\sqrt{6} & 0 & \cdots \\ \sqrt{2} & 0 & 1 & 0 & -\sqrt{12} & \cdots \\ 0 & \sqrt{6} & 0 & 1 & 0 & \cdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$

and

$$\boldsymbol{P}\boldsymbol{X} = i\frac{\hbar}{2} \begin{pmatrix} -1 & 0 & -\sqrt{2} & 0 & 0 & \cdots \\ 0 & -1 & 0 & -\sqrt{6} & 0 & \cdots \\ \sqrt{2} & 0 & -1 & 0 & -\sqrt{12} & \cdots \\ 0 & \sqrt{6} & 0 & -1 & 0 & \cdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots \end{pmatrix};$$

that is, indeed

$$\boldsymbol{X}\boldsymbol{P} - \boldsymbol{P}\boldsymbol{X} = i\hbar\boldsymbol{I}.$$

5.2. We have

$$\widetilde{\boldsymbol{X}} = \sqrt{\frac{m\omega_0}{2\hbar}}\ \boldsymbol{X}$$

and

$$\widetilde{\boldsymbol{P}} = \sqrt{\frac{1}{2m\omega_0\hbar}}\ \boldsymbol{P}.$$

$\hbar$ has the dimension of an action ($M \cdot L^2 \cdot T^{-1}$). The factor $\sqrt{\frac{m\omega_0}{2\hbar}}$ therefore has the dimension L^{-1}. Because $\boldsymbol{X}$ has the dimension of a length L, the first term of $\boldsymbol{A}$ is dimensionless. We find the same result for the second term, therefore $\boldsymbol{A}$ is dimensionless.

5.3. $\boldsymbol{A}$ is not Hermitian, inasmuch as

$$\begin{aligned}\boldsymbol{A}^\dagger &= \frac{1}{\sqrt{2\hbar}}\left(\sqrt{m\omega_0}\boldsymbol{X}^\dagger - \frac{i}{\sqrt{m\omega_0}}\boldsymbol{P}^\dagger\right)\\ &= \frac{1}{\sqrt{2\hbar}}\left(\sqrt{m\omega_0}\boldsymbol{X} - \frac{i}{\sqrt{m\omega_0}}\boldsymbol{P}\right)\\ &= \widetilde{\boldsymbol{X}} - i\widetilde{\boldsymbol{P}}\\ &\neq \boldsymbol{A}.\end{aligned}$$

Therefore, $\boldsymbol{A}$ cannot be an observable.

5.4. Because

$$\boldsymbol{N}^\dagger = (\boldsymbol{A}^\dagger\boldsymbol{A})^\dagger = \boldsymbol{A}^\dagger\boldsymbol{A} = \boldsymbol{N},$$

$\boldsymbol{N}$ is Hermitian.

5.5. We prove the formula by induction.

1. $\underline{n = 1}$: Is $[\boldsymbol{A}, \boldsymbol{A}^\dagger\boldsymbol{A}] = \boldsymbol{A}$? According to Problem 3.7,

$$[\boldsymbol{A}, \boldsymbol{BC}] = \boldsymbol{B}[\boldsymbol{A}, \boldsymbol{C}] + [\boldsymbol{A}, \boldsymbol{B}]\boldsymbol{C}.$$

Together with $[\boldsymbol{A}, \boldsymbol{A}^\dagger] = \boldsymbol{I}$ from (6.10), we get

$$[\boldsymbol{A}, \boldsymbol{A}^\dagger\boldsymbol{A}] = \boldsymbol{A}^\dagger \underbrace{[\boldsymbol{A}, \boldsymbol{A}]}_{\boldsymbol{0}} + \underbrace{[\boldsymbol{A}, \boldsymbol{A}^\dagger]}_{\boldsymbol{I}}\boldsymbol{A} = \boldsymbol{A}.$$

2. $\underline{n - 1 \to n}$: Does $[\boldsymbol{A}^n, \boldsymbol{N}] = n\boldsymbol{A}^n$ follow from

$$[\boldsymbol{A}^{n-1}, \boldsymbol{N}] = (n-1)\boldsymbol{A}^{n-1}?$$

Again according to Problem 3.7,

$$[\boldsymbol{AB}, \boldsymbol{C}] = \boldsymbol{A}[\boldsymbol{B}, \boldsymbol{C}] + [\boldsymbol{A}, \boldsymbol{C}]\boldsymbol{B}.$$

We can apply the above induction assumption to find

$$[\boldsymbol{A}^n, \boldsymbol{N}] = \boldsymbol{A}[\boldsymbol{A}^{n-1}, \boldsymbol{N}] + [\boldsymbol{A}, \boldsymbol{N}]\boldsymbol{A}^{n-1} = \boldsymbol{A}(n-1)\boldsymbol{A}^{n-1} + (1 \cdot \boldsymbol{A}^1)\boldsymbol{A}^{n-1} =$$

$$= (n-1)\boldsymbol{A}^n + \boldsymbol{A}^n = n\boldsymbol{A}^n.$$

5.6. With (5.32) and (5.33), we get

$$\boldsymbol{N} = \boldsymbol{A}^\dagger \boldsymbol{A} = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & \\ 0 & \sqrt{2} & 0 & 0 & \\ 0 & 0 & \sqrt{3} & 0 & \ddots \\ \vdots & & \ddots & \ddots & \ddots \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & \sqrt{2} & 0 & \\ 0 & 0 & 0 & \sqrt{3} & \ddots \\ \vdots & & & \ddots & \ddots \end{pmatrix} =$$

$$= \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \\ 0 & 0 & 2 & 0 & \\ 0 & 0 & 0 & 3 & \ddots \\ \vdots & & \ddots & \ddots & \ddots \end{pmatrix}.$$

The eigenvalues are the numbers at the main diagonal, namely 0, 1, 2, 3, ...

5.7. The eigenvalue equations (5.29) and (5.30) are

$$\boldsymbol{A}\boldsymbol{e}_n = \sqrt{n}\boldsymbol{e}_{n-1}$$

and

$$\boldsymbol{A}^\dagger \boldsymbol{e}_n = \sqrt{n+1}\boldsymbol{e}_{n+1}.$$

With these equations, (5.34) supplies

$$\boldsymbol{X}\boldsymbol{e}_n = \sqrt{\frac{\hbar}{2m\omega_0}}(\boldsymbol{A} + \boldsymbol{A}^\dagger)\boldsymbol{e}_n = \sqrt{\frac{\hbar}{2m\omega_0}}\left(\sqrt{n}\boldsymbol{e}_{n-1} + \sqrt{n+1}\boldsymbol{e}_{n+1}\right). \quad \text{(A.5)}$$

Multiplying by the transposed eigenvector $\boldsymbol{e}_m$ from the left gives the matrix element in the mth row and nth column of $\boldsymbol{X}$ (with $\boldsymbol{e}_m \boldsymbol{e}_n = \delta_{mn}$):

$$\boldsymbol{X}_{mn} = \boldsymbol{e}_m^\intercal \boldsymbol{X}\boldsymbol{e}_n = \sqrt{\frac{\hbar}{2m\omega_0}}\left(\sqrt{n}\,\delta_{m,n-1} + \sqrt{n+1}\,\delta_{m,n+1}\right).$$

Multiplying (A.5) with $\boldsymbol{X}$ yields

$$\boldsymbol{X}^{\mathbf{2}}\boldsymbol{e}_n = \sqrt{\frac{\hbar}{2m\omega_0}}\left(\sqrt{n}\boldsymbol{X}\boldsymbol{e}_{n-1} + \sqrt{n+1}\boldsymbol{X}\boldsymbol{e}_{n+1}\right), \tag{A.6}$$

and with (5.29) and (5.30) we get

$$\boldsymbol{X}^{\mathbf{2}}\boldsymbol{e}_n = \frac{\hbar}{2m\omega_0}\left(\sqrt{n}(\boldsymbol{A}+\boldsymbol{A}^{\dagger})\boldsymbol{e}_{n-1} + \sqrt{n+1}(\boldsymbol{A}+\boldsymbol{A}^{\dagger})\boldsymbol{e}_{n+1}\right) =$$

$$= \frac{\hbar}{2m\omega_0}\left(\sqrt{n(n-1)}\boldsymbol{e}_{n-2} + (2n+1)\boldsymbol{e}_n + \sqrt{(n+1)(n+2)}\boldsymbol{e}_{n+2}\right).$$

Multiplying this equation also by the transposed eigenvector $\boldsymbol{e}_m$ from the left gives the matrix element in the mth row and nth column of $\boldsymbol{X}^2$:

$$\boldsymbol{X}^2_{mn} = \frac{\hbar}{2m\omega_0}\left(\sqrt{n(n-1)}\delta_{m,n-2} + (2n+1)\delta_{m,n} + \sqrt{(n+1)(n+2)}\delta_{m,n+2}\right).$$

In the same way, we obtain

$$\boldsymbol{X}^3_{mn} = \left(\frac{\hbar}{2m\omega_0}\right)^{3/2}\left(\sqrt{n(n-1)(n-2)}\delta_{m,n-3} + 3\sqrt{(n+1)^3}\delta_{m,n+1} + \right.$$

$$\left. +3\sqrt{n^3}\delta_{m,n-1} + \sqrt{(n+1)(n+2)(n+3)}\delta_{m,n+3}\right).$$

The matrix $\boldsymbol{X}^3$ therefore has the form

$$\boldsymbol{X}^3 = \left(\frac{\hbar}{2m\omega_0}\right)^{3/2}\begin{pmatrix} 0 & 3 & 0 & \sqrt{2\cdot 3} & 0 & \cdots \\ 3 & 0 & 6\sqrt{2} & 0 & \sqrt{2\cdot 3\cdot 4} & \cdots \\ 0 & 6\sqrt{2} & 0 & 9\sqrt{3} & 0 & \cdots \\ \sqrt{2\cdot 3} & 0 & 9\sqrt{3} & 0 & & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots \end{pmatrix}.$$

6.1.

$$\boldsymbol{\mathfrak{X}}\cdot\boldsymbol{\mathfrak{P}} - \boldsymbol{\mathfrak{P}}\cdot\boldsymbol{\mathfrak{X}} = \boldsymbol{X_1P_1 + X_2P_2 + X_3P_3 - P_1X_1 - P_2X_2 - P_3X_3} =$$

$$= i\hbar\boldsymbol{I} + i\hbar\boldsymbol{I} + i\hbar\boldsymbol{I} = \underline{\underline{3\,i\,\hbar\,\boldsymbol{I}}}.$$

6.2. We have

$$(\boldsymbol{\mathfrak{A}}\times\boldsymbol{\mathfrak{B}})\cdot\boldsymbol{\mathfrak{C}} =$$

$$= \left(\begin{pmatrix} \boldsymbol{A}_1 \\ \boldsymbol{A}_2 \\ \boldsymbol{A}_3 \end{pmatrix} \times \begin{pmatrix} \boldsymbol{B}_1 \\ \boldsymbol{B}_2 \\ \boldsymbol{B}_3 \end{pmatrix} \right) \cdot \begin{pmatrix} \boldsymbol{C}_1 \\ \boldsymbol{C}_2 \\ \boldsymbol{C}_3 \end{pmatrix} =$$

$$= \begin{pmatrix} \boldsymbol{A}_2\boldsymbol{B}_3 - \boldsymbol{A}_3\boldsymbol{B}_2 \\ \boldsymbol{A}_3\boldsymbol{B}_1 - \boldsymbol{A}_1\boldsymbol{B}_3 \\ \boldsymbol{A}_1\boldsymbol{B}_2 - \boldsymbol{A}_2\boldsymbol{B}_1 \end{pmatrix} \cdot \begin{pmatrix} \boldsymbol{C}_1 \\ \boldsymbol{C}_2 \\ \boldsymbol{C}_3 \end{pmatrix} =$$

$$= \boldsymbol{A}_2\boldsymbol{B}_3\boldsymbol{C}_1 - \boldsymbol{A}_3\boldsymbol{B}_2\boldsymbol{C}_1 + \boldsymbol{A}_3\boldsymbol{B}_1\boldsymbol{C}_2 - \boldsymbol{A}_1\boldsymbol{B}_3\boldsymbol{C}_2 + \boldsymbol{A}_1\boldsymbol{B}_2\boldsymbol{C}_3 - \boldsymbol{A}_2\boldsymbol{B}_1\boldsymbol{C}_3 =$$

$$= \begin{pmatrix} \boldsymbol{A}_1 \\ \boldsymbol{A}_2 \\ \boldsymbol{A}_3 \end{pmatrix} \cdot \begin{pmatrix} \boldsymbol{B}_2\boldsymbol{C}_3 - \boldsymbol{B}_3\boldsymbol{C}_2 \\ \boldsymbol{B}_3\boldsymbol{C}_1 - \boldsymbol{B}_1\boldsymbol{C}_3 \\ \boldsymbol{B}_1\boldsymbol{C}_2 - \boldsymbol{B}_2\boldsymbol{C}_1 \end{pmatrix} = \underline{\underline{\mathfrak{A} \cdot (\mathfrak{B} \times \mathfrak{C})}}.$$

6.3. $\boldsymbol{L}_\pm^\dagger = (\boldsymbol{L}_1 \pm i\boldsymbol{L}_2)^\dagger = \boldsymbol{L}_1^\dagger \pm (i\boldsymbol{L}_2)^\dagger = \boldsymbol{L}_1 \mp i\boldsymbol{L}_2 = \boldsymbol{L}_\mp \neq \boldsymbol{L}_\pm$, so neither $\boldsymbol{L}_+$ nor $\boldsymbol{L}_-$ is Hermitian.

6.4. We have

$$|\boldsymbol{L}_+\boldsymbol{e}(j,m)|^2 = |\alpha\boldsymbol{e}(j,m+1)|^2 = |\alpha|^2 \underbrace{\boldsymbol{e}(j,m+1)^\dagger\boldsymbol{e}(j,m+1)}_{1} = |\alpha|^2.$$

Also, it applies

$$(\boldsymbol{L}_+\boldsymbol{e}(j,m))^\dagger(\boldsymbol{L}_+\boldsymbol{e}(j,m)) = \boldsymbol{e}(j,m)^\dagger(\boldsymbol{L}_-\boldsymbol{L}_+)\boldsymbol{e}(j,m).$$

With $\boldsymbol{L}_-\boldsymbol{L}_+ = \boldsymbol{L}^2 - \boldsymbol{L}_3^2 - \hbar\boldsymbol{L}_3$ from (6.21), we obtain

$$\begin{aligned}(\boldsymbol{L}_+\boldsymbol{e}(j,m))^\dagger(\boldsymbol{L}_+\boldsymbol{e}(j,m)) &= \boldsymbol{e}(j,m)^\dagger(\boldsymbol{L}^2 - \boldsymbol{L}_3^2 - \hbar\boldsymbol{L}_3)\boldsymbol{e}(j,m) \\ &= \hbar^2(j(j+1) - m^2 - m)\underbrace{\boldsymbol{e}(j,m)^\dagger\boldsymbol{e}(j,m)}_{1}.\end{aligned}$$

This implies

$$|\alpha|^2 = \hbar^2(j(j+1) - m^2 - m)$$

and therefore

$$\alpha = \hbar\sqrt{j(j+1) - m^2 - m} = \hbar\sqrt{(j-m)(j+m+1)}.$$

The factor for $\boldsymbol{L}_-$ can be calculated in a similar way.

7.1. First, we know that

$$\dot{\boldsymbol{P}}_i = -\frac{\partial \boldsymbol{H}}{\partial \boldsymbol{X}_i}.$$

If the Hamiltonian is

$$\boldsymbol{H} = \frac{1}{2m}\mathfrak{P}^2 - Ze^2\boldsymbol{R}^{-1},$$

then we find

$$\frac{\partial \boldsymbol{H}}{\partial \boldsymbol{X}_i} = -Ze^2\frac{\partial \boldsymbol{R}^{-1}}{\partial \boldsymbol{X}_i} = -Ze^2\frac{\partial(\boldsymbol{X}_1^2 + \boldsymbol{X}_2^2 + \boldsymbol{X}_3^2)^{-\frac{1}{2}}}{\partial \boldsymbol{X}_i} =$$

$$= Ze^2\boldsymbol{X}_i(\boldsymbol{X}_1^2 + \boldsymbol{X}_2^2 + \boldsymbol{X}_3^2)^{-\frac{3}{2}} = Ze^2\boldsymbol{X}_i\boldsymbol{R}^{-3},$$

in other words,

$$\dot{\mathfrak{P}} = -Ze^2\mathfrak{R}\boldsymbol{R}^{-3}. \tag{A.7}$$

7.2. For the first component matrix $\boldsymbol{X}_1\boldsymbol{R}^{-1}$ of $\mathfrak{R}\boldsymbol{R}^{-1}$, we obtain with the help of the Heisenberg formula $\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A} = \frac{i}{\hbar}[\boldsymbol{E}, \boldsymbol{A}]$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\boldsymbol{X}_1\boldsymbol{R}^{-1}\right) = \frac{i}{\hbar}\left\{\boldsymbol{E}(\boldsymbol{X}_1\boldsymbol{R}^{-1}) - (\boldsymbol{X}_1\boldsymbol{R}^{-1})\boldsymbol{E}\right\} = \tag{A.8}$$

$$= \frac{i}{2m\hbar}\left\{\mathfrak{P}^2(\boldsymbol{X}_1\boldsymbol{R}^{-1}) - (\boldsymbol{X}_1\boldsymbol{R}^{-1})\mathfrak{P}^2\right\}. \tag{A.9}$$

Note that we used the fact that

$$\boldsymbol{R}\boldsymbol{X}_i = (\boldsymbol{X}_1^2\boldsymbol{X}_i^2 + \boldsymbol{X}_2^2\boldsymbol{X}_i^2 + \boldsymbol{X}_3^2\boldsymbol{X}_i{}^2)^{1/2} = \boldsymbol{X}_i\boldsymbol{R},$$

which implies $\boldsymbol{X}_i\boldsymbol{R}^{-1} = \boldsymbol{R}^{-1}\boldsymbol{X}_i$. Therefore, the component $-Ze^2\boldsymbol{R}^{-1}$ of the Hamiltonian

$$\boldsymbol{E} = \frac{1}{2m}\mathfrak{P}^2 - Ze^2\boldsymbol{R}^{-1}$$

vanishes in (A.8). Now with

$$\mathfrak{P}^2 = \boldsymbol{P}_1^2 + \boldsymbol{P}_2^2 + \boldsymbol{P}_3^2$$

Equation (A.9) can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\boldsymbol{X}_1\boldsymbol{R}^{-1}\right) = \frac{i}{2m\hbar}\left\{\sum_{j=1}^{3}\left(\boldsymbol{P}_j^2\left(\boldsymbol{X}_1\boldsymbol{R}^{-1}\right) - \left(\boldsymbol{X}_1\boldsymbol{R}^{-1}\right)\boldsymbol{P}_j^2\right)\right\}. \tag{A.10}$$

We can add the entity $\mathbf{0} = -\boldsymbol{P}_j\boldsymbol{X}_i\boldsymbol{R}^{-1}\boldsymbol{P}_j + \boldsymbol{P}_j\boldsymbol{X}_i\boldsymbol{R}^{-1}\boldsymbol{P}_j$ into (A.10), which leads to

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\boldsymbol{X}_1\boldsymbol{R}^{-1}\right) =$$

$$= \frac{i}{2m\hbar} \Biggl\{ \sum_{j=1}^{3} \Bigl\{ \Bigl(\boldsymbol{P}_j (\boldsymbol{P}_j \boldsymbol{X}_1 \boldsymbol{R}^{-1} - \boldsymbol{X}_1 \boldsymbol{R}^{-1} \boldsymbol{P}_j \Bigr) + (\boldsymbol{P}_j \boldsymbol{X}_1 \boldsymbol{R}^{-1} - \boldsymbol{X}_1 \boldsymbol{R}^{-1} \boldsymbol{P}_j \Bigr) \boldsymbol{P}_j \Biggr\}. \tag{A.11}$$

Multiplication with $\boldsymbol{R}^{\mathbf{3}} \boldsymbol{R}^{\mathbf{-3}} = \boldsymbol{I}$ from the right supplies

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\boldsymbol{X}_1 \boldsymbol{R}^{-1} \right) =$$

$$= \frac{i}{2m\hbar} \sum_{j=1}^{3} \bigl\{ \boldsymbol{P}_j \left(\boldsymbol{P}_j \boldsymbol{X}_1 \boldsymbol{R}^2 - \boldsymbol{X}_1 \boldsymbol{R}^{-1} \boldsymbol{P}_j \boldsymbol{R}^3 \right) +$$

$$+ \left(\boldsymbol{P}_j \boldsymbol{X}_1 \boldsymbol{R}^{-1} - \boldsymbol{X}_1 \boldsymbol{R}^{-1} \boldsymbol{P}_j \right) \boldsymbol{P}_j \boldsymbol{R}^3 \bigr\} \boldsymbol{R}^{-3}. \tag{A.12}$$

If we multiply (7.5), namely

$$\boldsymbol{P}_j \boldsymbol{R} - \boldsymbol{R} \boldsymbol{P}_j = \frac{h}{2\pi \boldsymbol{i}} \boldsymbol{X}_j \boldsymbol{R}^{-1},$$

with $\boldsymbol{R}^{-1}$ from both the left and the right, and if we also take into account that $\boldsymbol{R}^{-1} \boldsymbol{X}_j = \boldsymbol{X}_j \boldsymbol{R}^{-1}$, we get

$$\boldsymbol{R}^{-1} \boldsymbol{P}_j - \boldsymbol{P}_j \boldsymbol{R}^{-1} = \frac{\hbar}{\boldsymbol{i}} \boldsymbol{X}_j \boldsymbol{R}^{-3}.$$

Rearranging yields

$$\boldsymbol{R}^{-1} \boldsymbol{P}_j = \boldsymbol{P}_j \boldsymbol{R}^{-1} - i\hbar \boldsymbol{X}_j \boldsymbol{R}^{-3}. \tag{A.13}$$

We can now put this result into the round parentheses in (A.12) in order to get

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\boldsymbol{X}_1 \boldsymbol{R}^{-1} \right) =$$

$$= \frac{i}{2m\hbar} \sum_{j=1}^{3} \bigl\{ \boldsymbol{P}_j \left(\boldsymbol{P}_j \boldsymbol{X}_1 \boldsymbol{R}^2 - \boldsymbol{X}_1 \boldsymbol{P}_j \boldsymbol{R}^2 + i\hbar \boldsymbol{X}_1 \boldsymbol{X}_j \right) +$$

$$+ \left(\boldsymbol{P}_j \boldsymbol{X}_1 \boldsymbol{R}^{-1} - \boldsymbol{X}_1 \boldsymbol{P}_j \boldsymbol{R}^{-1} + i\hbar \boldsymbol{X}_1 \boldsymbol{X}_j \boldsymbol{R}^{-3} \right) \boldsymbol{P}_j \bigr\}. \tag{A.14}$$

Next, we know that $\boldsymbol{P}_j \boldsymbol{X}_1 - \boldsymbol{X}_1 \boldsymbol{P}_j = \frac{\hbar}{i} \boldsymbol{I}$ for $j = 1$ and $= \mathbf{0}$ for $j \neq 1$. Using this result as well as inserting $\boldsymbol{I} = \boldsymbol{R}^3 \boldsymbol{R}^{-3}$ in the second round parentheses, we get

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\boldsymbol{X}_1 \boldsymbol{R}^{-1} \right) =$$

$$= \frac{1}{2m}\left\{\left(\boldsymbol{P}_1(\boldsymbol{X}_2^2+\boldsymbol{X}_3^2)\boldsymbol{R}^{-3}-\boldsymbol{P}_2\boldsymbol{X}_{\mathbf{1}}\boldsymbol{X}_{\mathbf{2}}\boldsymbol{R}^{-3}-\boldsymbol{P}_3\boldsymbol{X}_{\mathbf{1}}\boldsymbol{X}_{\mathbf{3}}\boldsymbol{R}^{-3}\right)+\right.$$

$$\left.+\left((\boldsymbol{X}_2^2+\boldsymbol{X}_3^2)\boldsymbol{R}^{-3}\boldsymbol{P}_1-\boldsymbol{X}_{\mathbf{1}}\boldsymbol{X}_{\mathbf{2}}\boldsymbol{R}^{-3}\boldsymbol{P}_2-\boldsymbol{X}_{\mathbf{1}}\boldsymbol{X}_{\mathbf{3}}\boldsymbol{R}^{-3}\boldsymbol{P}_3\right)\right\}. \qquad \text{(A.15)}$$

It turns out that

$$(\boldsymbol{P}_1(\boldsymbol{X}_2^2+\boldsymbol{X}_3^2)\boldsymbol{R}^{-3}-\boldsymbol{P}_2\boldsymbol{X}_{\mathbf{1}}\boldsymbol{X}_{\mathbf{2}}\boldsymbol{R}^{-3}-\boldsymbol{P}_3\boldsymbol{X}_{\mathbf{1}}\boldsymbol{X}_{\mathbf{3}}\boldsymbol{R}^{-3})=$$

$$=\boldsymbol{X}_3\boldsymbol{P}_1\boldsymbol{X}_3\boldsymbol{R}^{-3}-\boldsymbol{X}_1\boldsymbol{P}_3\boldsymbol{X}_3\boldsymbol{R}^{-3}-\boldsymbol{X}_1\boldsymbol{P}_2\boldsymbol{X}_2\boldsymbol{R}^{-3}+\boldsymbol{X}_2\boldsymbol{P}_1\boldsymbol{X}_2\boldsymbol{R}^{-3}=$$

$$=(\boldsymbol{X}_3\boldsymbol{P}_1-\boldsymbol{X}_1\boldsymbol{P}_3)\boldsymbol{X}_3\boldsymbol{R}^{-3}-(\boldsymbol{X}_1\boldsymbol{P}_2-\boldsymbol{X}_2\boldsymbol{P}_1)\boldsymbol{X}_2\boldsymbol{R}^{-3}=$$

$$\boldsymbol{L}_2\boldsymbol{X}_3\boldsymbol{R}^{-3}-\boldsymbol{L}_3\boldsymbol{X}_2\boldsymbol{R}^{-3},$$

that is, the first component of $\mathfrak{L}\times\mathfrak{R}\boldsymbol{R}^{-3}$. So we finally get the desired result

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathfrak{R}\boldsymbol{R}^{-\mathbf{1}})=\frac{\mathbf{1}}{\mathbf{2}\boldsymbol{m}}\left\{\mathfrak{L}\times(\mathfrak{R}\boldsymbol{R}^{-\mathbf{3}})-(\mathfrak{R}\boldsymbol{R}^{-\mathbf{3}})\times\mathfrak{L}\right\}. \qquad \text{(A.16)}$$

7.3. According to (7.11), the Lenz matrix-vector is defined as

$$\mathfrak{A}\overset{\text{def}}{=}\frac{1}{Ze^2m}\frac{1}{2}(\mathfrak{L}\times\mathfrak{P}-\mathfrak{P}\times\mathfrak{L})+\mathfrak{R}\boldsymbol{R}^{-1},$$

where the angular momentum $\mathfrak{L}$ (in the round parentheses) is constant. For the time derivative of $\mathfrak{A}$, we therefore obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathfrak{A}=\frac{1}{Ze^2m}\frac{1}{2}\left(\mathfrak{L}\times\dot{\mathfrak{P}}-\dot{\mathfrak{P}}\times\mathfrak{L}\right)+\frac{\mathrm{d}}{\mathrm{d}t}\mathfrak{R}\boldsymbol{R}^{-1}. \qquad \text{(A.17)}$$

With (A.7) and (A.16), we get

$$\underline{\underline{\frac{\mathrm{d}}{\mathrm{d}t}\mathfrak{A}}}=-\frac{1}{2m}\left\{\mathfrak{L}\times(\mathfrak{R}\boldsymbol{R}^{-3})-(\mathfrak{R}\boldsymbol{R}^{-3})\times\mathfrak{L}\right\}+$$

$$+\frac{1}{2m}\left\{\mathfrak{L}\times(\mathfrak{R}\boldsymbol{R}^{-3})-(\mathfrak{R}\boldsymbol{R}^{-3})\times\mathfrak{L}\right\}=\underline{\underline{\mathbf{0}}}. \qquad \text{(A.18)}$$

8.1. The Pauli spin matrix $\boldsymbol{\sigma}_1$ has the form

$$\boldsymbol{\sigma}_1=\begin{pmatrix}0&1\\1&0\end{pmatrix}.$$

With α a real number, the power series expansion of the exponential function is

$$\exp(i\alpha\boldsymbol{\sigma}_1) = \sum_{\nu=0}^{\infty} \frac{i\alpha^{\nu}}{\nu!}\,\boldsymbol{\sigma}_1^{\nu}.$$

We find that

$$\boldsymbol{\sigma}_1^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \boldsymbol{I}_2,$$

and therefore

$$\boldsymbol{\sigma}_1^{2\nu} = \boldsymbol{I}_2 \quad \text{and} \quad \boldsymbol{\sigma}_1^{2\nu+1} = \boldsymbol{\sigma}_1. \tag{A.19}$$

We can now separate the sum into one part with even integers and one part with odd integers:

$$\exp(i\alpha\boldsymbol{\sigma}_1) = \sum_{\nu=0}^{\infty} \frac{(i\alpha)^{2\nu}}{(2\nu)!}\,\boldsymbol{\sigma}_1^{2\nu} + \sum_{\nu=0}^{\infty} \frac{(i\alpha)^{2\nu+1}}{(2\nu+1)!}\,\boldsymbol{\sigma}_1^{2\nu+1}.$$

With (A.19), we get

$$\begin{aligned}\exp(i\alpha\boldsymbol{\sigma}_1) &= \boldsymbol{I}_2 \sum_{\nu=0}^{\infty}(-1)^{\nu}\frac{(i\alpha)^{2\nu}}{(2\nu)!} + i\boldsymbol{\sigma}_1 \sum_{\nu=0}^{\infty}(-1)^{\nu}\frac{(i\alpha)^{2\nu+1}}{(2\nu+1)!} = \\ &= \boldsymbol{I}_2 \cos\alpha + i\boldsymbol{\sigma}_1 \sin\alpha \\ &= \begin{pmatrix} \cos\alpha & 0 \\ 0 & \cos\alpha \end{pmatrix} + \begin{pmatrix} 0 & i\sin\alpha \\ i\sin\alpha & 0 \end{pmatrix} \\ &= \begin{pmatrix} \cos\alpha & i\sin\alpha \\ i\sin\alpha & \cos\alpha \end{pmatrix}.\end{aligned}$$

8.2. Each of the three Pauli matrices has the two eigenvalues +1 and −1. The corresponding normalized eigenvectors are

$$\boldsymbol{e}_{1+} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \boldsymbol{e}_{1-} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

$$\boldsymbol{e}_{2+} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix}, \quad \boldsymbol{e}_{2-} = \frac{1}{\sqrt{2}}\begin{pmatrix} i \\ 1 \end{pmatrix},$$

$$\boldsymbol{e}_{3+} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \boldsymbol{e}_{3-} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

8.3. If we write the eigenvectors as columns of the transformation matrix $\boldsymbol{T}$, we obtain, for example, for the Pauli matrix $\boldsymbol{\sigma}_1$

$$\boldsymbol{T} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

This yields

$$\boldsymbol{T}^\dagger \boldsymbol{\sigma}_1 \boldsymbol{T} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

that is, a diagonal matrix with the eigenvalues $+1$ and -1 on the diagonal.

8.4. We start with the largest possible values $j = \frac{1}{2} + \frac{1}{2} = 1$ and $m = \frac{1}{2} + \frac{1}{2} = 1$. m can then have the values -1, 0, and $+1$. For the largest total angular momentum quantum number $j = m = \frac{1}{2} + \frac{1}{2} = 1$, there is exactly one state in the coupled and uncoupled basis:

$$\boldsymbol{e}_{j=1,m=1} = \boldsymbol{e}_{\frac{1}{2},\frac{1}{2}} \otimes \boldsymbol{e}_{\frac{1}{2},\frac{1}{2}}. \tag{A.20}$$

Now let us remember that for spin-$\frac{1}{2}$ systems the Pauli matrix $\boldsymbol{\sigma}_3$ has the two eigenvectors

$$\boldsymbol{e}_{\frac{1}{2},\frac{1}{2}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \boldsymbol{e}_{\frac{1}{2},-\frac{1}{2}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The four possible vector combinations are

$$\boldsymbol{e}_{\frac{1}{2},\frac{1}{2}} \otimes \boldsymbol{e}_{\frac{1}{2},\frac{1}{2}} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{e}_{\frac{1}{2},\frac{1}{2}} \otimes \boldsymbol{e}_{\frac{1}{2},-\frac{1}{2}} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix},$$

$$\boldsymbol{e}_{\frac{1}{2},-\frac{1}{2}} \otimes \boldsymbol{e}_{\frac{1}{2},\frac{1}{2}} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \boldsymbol{e}_{\frac{1}{2},-\frac{1}{2}} \otimes \boldsymbol{e}_{\frac{1}{2},-\frac{1}{2}} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

We can apply the lowering operator

$$\boldsymbol{J}_- = \boldsymbol{J}_1 + i\boldsymbol{J}_2 = \boldsymbol{S}_- \otimes \boldsymbol{I}_2 + \boldsymbol{I}_2 \otimes \boldsymbol{S}_-$$

to (A.20) and get

$$\boldsymbol{J}_- \boldsymbol{e}_{j=1,m=1} = (\boldsymbol{S}_- \boldsymbol{e}_{\frac{1}{2},\frac{1}{2}}) \otimes \boldsymbol{e}_{\frac{1}{2},\frac{1}{2}} + \boldsymbol{e}_{\frac{1}{2},\frac{1}{2}} \otimes (\boldsymbol{S}_- \boldsymbol{e}_{\frac{1}{2},\frac{1}{2}}). \tag{A.21}$$

With (6.49), namely

$$\boldsymbol{J}_- \boldsymbol{e}_{\boldsymbol{jm}} = [j(j+1) - m(m-1)]^{1/2}\, \hbar \boldsymbol{e}_{\boldsymbol{j,m-1}}$$

and

$$\boldsymbol{S_{-}e_{sm}} = [s(s+1) - m(m-1)]^{1/2}\, \hbar \boldsymbol{e_{s,m-1}},$$

we obtain from (A.21)

$$(2)^{\frac{1}{2}} \hbar \boldsymbol{e}_{j=1,m=0} = \hbar \boldsymbol{e}_{\frac{1}{2},-\frac{1}{2}} \otimes \boldsymbol{e}_{\frac{1}{2},\frac{1}{2}} + \hbar \boldsymbol{e}_{\frac{1}{2},\frac{1}{2}} \otimes \boldsymbol{e}_{\frac{1}{2},-\frac{1}{2}},$$

which leads to

$$\boldsymbol{e}_{j=1,m=0} = \frac{1}{\sqrt{2}} \boldsymbol{e}_{\frac{1}{2},-\frac{1}{2}} \otimes \boldsymbol{e}_{\frac{1}{2},\frac{1}{2}} + \frac{1}{\sqrt{2}} \boldsymbol{e}_{\frac{1}{2},\frac{1}{2}} \otimes \boldsymbol{e}_{\frac{1}{2},-\frac{1}{2}} =$$

$$= \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}.$$

According to (8.37), the Clebsch–Gordan coefficients are defined as

$$C(a, b; j, m) = (\boldsymbol{e}'_a \otimes \boldsymbol{e}''_b)^{\dagger} \boldsymbol{e}(j, m).$$

For our problem, this reads

$$C(a, b; j, m) = (\boldsymbol{e}_a \otimes \boldsymbol{e}_b)^{\dagger} \boldsymbol{e}(j, m).$$

In particular, we find the Clebsch–Gordan coefficients

$$\underline{\underline{C\left(\left(\frac{1}{2}, -\frac{1}{2}\right), \left(\frac{1}{2}, \frac{1}{2}\right); j = 1, m = 0\right)}} = (\boldsymbol{e}_{\frac{1}{2},-\frac{1}{2}} \otimes \boldsymbol{e}_{\frac{1}{2},\frac{1}{2}})^{\dagger} \boldsymbol{e}(1, 0) = \underline{\underline{\frac{1}{\sqrt{2}}}},$$

$$\underline{\underline{C\left(\left(\frac{1}{2}, \frac{1}{2}\right), \left(\frac{1}{2}, -\frac{1}{2}\right); j = 1, m = 0\right)}} = (\boldsymbol{e}_{\frac{1}{2},\frac{1}{2}} \otimes \boldsymbol{e}_{\frac{1}{2},-\frac{1}{2}})^{\dagger} \boldsymbol{e}(1, 0) = \underline{\underline{\frac{1}{\sqrt{2}}}}.$$

Another Clebsch–Gordan coefficient can be obtained directly from (A.20):

$$\underline{\underline{C\left(\left(\frac{1}{2}, \frac{1}{2}\right), \left(\frac{1}{2}, \frac{1}{2}\right); j = 1, m = 1\right) = 1.}}$$

Applying the lowering operator $\boldsymbol{J}_{-}$ once again yields

$$\underline{\underline{C\left(\left(\frac{1}{2}, \frac{1}{2}\right), \left(\frac{1}{2}, -\frac{1}{2}\right); j = 1, m = -1\right) = 1.}}$$

Until now we have had three basis vectors, namely

$$\frac{1}{\sqrt{2}}\begin{pmatrix}0\\1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \quad \text{and} \quad \begin{pmatrix}0\\0\\0\\1\end{pmatrix}.$$

By choosing an orthonormal vector corresponding to $\boldsymbol{e}_{j=1,m=0}$, we get the missing fourth basis vector

$$\frac{1}{\sqrt{2}}\begin{pmatrix}0\\1\\-1\\0\end{pmatrix}.$$

This means that the additional Clebsch–Gordan coefficients are $\underline{\underline{\pm\frac{1}{\sqrt{2}}}}$.

8.5. Because $\boldsymbol{J}^2$ and $\boldsymbol{J}_3$ commute, they have identical eigenvectors $\boldsymbol{e}_{jm}$. The eigenvalue equations of $\boldsymbol{J}$ are then

$$\boldsymbol{J}^2\boldsymbol{e}_{jm} = j(j+1)\hbar^2\boldsymbol{e}_{jm}, \tag{A.22}$$

$$\boldsymbol{J}_3\boldsymbol{e}_{jm} = m\hbar\boldsymbol{e}_{jm}, \quad -j \le m \le j. \tag{A.23}$$

The largest total angular momentum quantum number is $j = m = \ell + \frac{1}{2}$. For these values, there is exactly one state in the coupled and uncoupled basis:

$$\boldsymbol{e}_{j=\ell+\frac{1}{2},m=\ell+\frac{1}{2}} = \boldsymbol{e}_{\ell,\ell} \otimes \boldsymbol{e}_{\frac{1}{2},\frac{1}{2}}.$$

As in the previous problem, we apply the lowering operator

$$\boldsymbol{J}_- = \boldsymbol{J}_1 + i\boldsymbol{J}_2 = \boldsymbol{L}_- \otimes \boldsymbol{I}_S + \boldsymbol{I}_L \otimes \boldsymbol{S}_-$$

and receive

$$\boldsymbol{J}_-\boldsymbol{e}_{j=\ell+\frac{1}{2},m=\ell+\frac{1}{2}} = \boldsymbol{L}_-\boldsymbol{e}_{\ell,\ell} \otimes \boldsymbol{e}_{\frac{1}{2},\frac{1}{2}} + \boldsymbol{e}_{\ell,\ell} \otimes \boldsymbol{S}_-\boldsymbol{e}_{\frac{1}{2},\frac{1}{2}}. \tag{A.24}$$

With (6.49), namely

$$\boldsymbol{L}_-\boldsymbol{e}_{\boldsymbol{\ell m}} = [\ell(\ell+1) - m(m-1)]^{1/2}\,\hbar\boldsymbol{e}_{\boldsymbol{\ell,m-1}},$$

$$\boldsymbol{J}_-\boldsymbol{e}_{\boldsymbol{jm}} = [j(j+1) - m(m-1)]^{1/2}\,\hbar\boldsymbol{e}_{\boldsymbol{j,m-1}},$$

and

$$\boldsymbol{S}_-\boldsymbol{e}_{\boldsymbol{sm}} = [s(s+1) - m(m-1)]^{1/2}\,\hbar\boldsymbol{e}_{\boldsymbol{s,m-1}},$$

we obtain from (A.24)

$$(2\ell+1)^{\frac{1}{2}}\hbar\boldsymbol{e}_{j=\ell+\frac{1}{2},m=\ell-\frac{1}{2}} = (2\ell)^{\frac{1}{2}}\hbar\boldsymbol{e}_{\ell,\ell-1}\otimes\boldsymbol{e}_{\frac{1}{2},\frac{1}{2}} + (1)^{\frac{1}{2}}\hbar\boldsymbol{e}_{\ell,\ell}\otimes\boldsymbol{e}_{\frac{1}{2},-\frac{1}{2}}.$$

This can be written as

$$\boldsymbol{e}_{j=\ell+\frac{1}{2},m=\ell-\frac{1}{2}} = \left(\frac{2\ell}{2\ell+1}\right)^{\frac{1}{2}}\boldsymbol{e}_{\ell,\ell-1}\otimes\boldsymbol{e}_{\frac{1}{2},\frac{1}{2}} + \left(\frac{1}{2\ell+1}\right)^{\frac{1}{2}}\boldsymbol{e}_{\ell,\ell}\otimes\boldsymbol{e}_{\frac{1}{2},-\frac{1}{2}}.$$

The vectors $\boldsymbol{e}_{\ell,\ell-1}\otimes\boldsymbol{e}_{\frac{1}{2},\frac{1}{2}}$ and $\boldsymbol{e}_{\ell,\ell}\otimes\boldsymbol{e}_{\frac{1}{2},-\frac{1}{2}}$ are orthogonal to each other. The orthogonal linear combination is therefore the state of the total angular momentum $j=\ell-1/2$ with the same $m=\ell-1/2$:

$$\boldsymbol{e}_{j=\ell-\frac{1}{2},m=\ell-\frac{1}{2}} = \left(\frac{\ell}{2\ell+1}\right)^{\frac{1}{2}}\boldsymbol{e}_{\ell,\ell-1}\otimes\boldsymbol{e}_{\frac{1}{2},\frac{1}{2}} - \left(\frac{2\ell}{2\ell+1}\right)^{\frac{1}{2}}\boldsymbol{e}_{\ell,\ell}\otimes\boldsymbol{e}_{\frac{1}{2},-\frac{1}{2}}.$$

Applying the lowering operator $\boldsymbol{J}_-$ as above and writing down the orthogonal linear combination, we find

$$\boldsymbol{e}_{j=\ell+\frac{1}{2},m=\ell-\frac{3}{2}} = \left(\frac{2\ell-1}{2\ell+1}\right)^{\frac{1}{2}}\boldsymbol{e}_{\ell,\ell-2}\otimes\boldsymbol{e}_{\frac{1}{2},\frac{1}{2}} + \left(\frac{2}{2\ell+1}\right)^{\frac{1}{2}}\boldsymbol{e}_{\ell,\ell}\otimes\boldsymbol{e}_{\frac{1}{2},-\frac{1}{2}},$$

$$\boldsymbol{e}_{j=\ell-\frac{1}{2},m=\ell-\frac{3}{2}} = \left(\frac{2}{2\ell+1}\right)^{\frac{1}{2}}\boldsymbol{e}_{\ell,\ell-2}\otimes\boldsymbol{e}_{\frac{1}{2},\frac{1}{2}} - \left(\frac{2\ell-1}{2\ell+1}\right)^{\frac{1}{2}}\boldsymbol{e}_{\ell,\ell-1}\otimes\boldsymbol{e}_{\frac{1}{2},-\frac{1}{2}}.$$

With this method, we get the general result

$$\boldsymbol{e}_{j=\ell+\frac{1}{2},m} = \left(\frac{\ell+m+\frac{1}{2}}{2\ell+1}\right)^{\frac{1}{2}}\boldsymbol{e}_{\ell,m-\frac{1}{2}}\otimes\boldsymbol{e}_{\frac{1}{2},\frac{1}{2}} + \left(\frac{\ell-m+\frac{1}{2}}{2\ell+1}\right)^{\frac{1}{2}}\boldsymbol{e}_{\ell,m+\frac{1}{2}}\otimes\boldsymbol{e}_{\frac{1}{2},-\frac{1}{2}},$$

$$\boldsymbol{e}_{j=\ell-\frac{1}{2},m} = \left(\frac{\ell-m+\frac{1}{2}}{2\ell+1}\right)^{\frac{1}{2}}\boldsymbol{e}_{\ell,m-\frac{1}{2}}\otimes\boldsymbol{e}_{\frac{1}{2},\frac{1}{2}} - \left(\frac{\ell+m+\frac{1}{2}}{2\ell+1}\right)^{\frac{1}{2}}\boldsymbol{e}_{\ell,m+\frac{1}{2}}\otimes\boldsymbol{e}_{\frac{1}{2},-\frac{1}{2}}.$$

There are $2(2\ell+1)$ states in total. The Clebsch–Gordan coefficients are

$$\underline{\underline{C\left(\left(\ell,m\mp\frac{1}{2}\right),\left(\frac{1}{2},\pm\frac{1}{2}\right);j=\ell+\frac{1}{2},m\right)=}}$$

$$(\boldsymbol{e}_{\ell,m\mp\frac{1}{2}}\otimes\boldsymbol{e}_{\frac{1}{2},\pm\frac{1}{2}})^{\dagger}\boldsymbol{e}_{j=\ell+\frac{1}{2},m} = \underline{\underline{\left(\frac{\ell\pm m+\frac{1}{2}}{2\ell+1}\right)^{\frac{1}{2}}}},$$

and

$$\underline{\underline{C\left(\left(\ell, m \mp \frac{1}{2}\right), \left(\frac{1}{2}, \pm\frac{1}{2}\right); j = \ell - \frac{1}{2}, m\right) =}}$$

$$= (\boldsymbol{e}_{\ell, m\mp\frac{1}{2}} \otimes \boldsymbol{e}_{\frac{1}{2},\pm\frac{1}{2}})^\dagger \boldsymbol{e}_{j=\ell-\frac{1}{2},m} = \underline{\underline{\pm\left(\frac{\ell \mp m + \frac{1}{2}}{2\ell+1}\right)^{\frac{1}{2}}}}.$$

9.1. In the normal Zeeman effect, the energy level E_n splits into $2\ell+1$ levels under the influence of a magnetic field. After all, the highest ℓ-value for a fixed n is $\ell = n-1$, and these include $2\ell+1 = 2n-1$ values of m. For $\ell = 2$, we therefore find $2\ell+1 = 5$ values of m, namely $m = +2, +1, 0, -1$ and -2. In other words, the level $\ell = 2$ is split into $\underline{\underline{5}}$ levels.

10.1. From (10.7), we find in matrix form

$$\begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} = \boldsymbol{C} = \begin{pmatrix} a_1 b_1 & a_1 b_2 \\ a_2 b_1 & a_2 b_2 \end{pmatrix} = \begin{pmatrix} a_1 \boldsymbol{b}^\intercal \\ a_2 \boldsymbol{b}^\intercal \end{pmatrix} = \boldsymbol{a}\boldsymbol{b}^\intercal.$$

In the second matrix, the second line is a multiple of the first line, therefore the determinant of the matrix is zero. This implies that the determinant of $\boldsymbol{C}$ must also be equal to zero:

$$c_{11}\, c_{22} - c_{12}\, c_{21} = 0.$$

By the way, an algebraic computer program for the calculation of the a_i and b_j can be found in [23].

10.2.

$$\boldsymbol{A}_{1,2}\boldsymbol{B}_{1,2} = (\boldsymbol{A} \otimes \boldsymbol{I}_B)(\boldsymbol{I}_A \otimes \boldsymbol{B}) = \boldsymbol{A} \otimes \boldsymbol{B},$$

$$\boldsymbol{B}_{1,2}\boldsymbol{A}_{1,2} = (\boldsymbol{I}_A \otimes \boldsymbol{B})(\boldsymbol{A} \otimes \boldsymbol{I}_B) = \boldsymbol{A} \otimes \boldsymbol{B}.$$

It follows that

$$[\boldsymbol{A}_{1,2}, \boldsymbol{B}_{1,2}] = \boldsymbol{A} \otimes \boldsymbol{B} - \boldsymbol{A} \otimes \boldsymbol{B} = \boldsymbol{0}.$$

10.3.

$$\langle \boldsymbol{A}_{1,2} \rangle = (\boldsymbol{\xi}_1 \otimes \boldsymbol{\xi}_2)^\dagger (\boldsymbol{A} \otimes \boldsymbol{I}_B)(\boldsymbol{\xi}_1 \otimes \boldsymbol{\xi}_2) = (\boldsymbol{\xi}_1^\dagger \boldsymbol{A} \boldsymbol{\xi}_1)\underbrace{(\boldsymbol{\xi}_2^\dagger \boldsymbol{\xi}_2)}_{1} = \langle \boldsymbol{A} \rangle,$$

$$\langle \boldsymbol{B}_{1,2} \rangle = (\boldsymbol{\xi}_1 \otimes \boldsymbol{\xi}_2)^\dagger (\boldsymbol{I}_A \otimes \boldsymbol{B})(\boldsymbol{\xi}_1 \otimes \boldsymbol{\xi}_2) = \underbrace{(\boldsymbol{\xi}_1^\dagger \boldsymbol{\xi}_1)}_{1}(\boldsymbol{\xi}_2^\dagger \boldsymbol{B} \boldsymbol{\xi}_2) = \langle \boldsymbol{B} \rangle.$$

10.4. According to (11.8), we have

$$\boldsymbol{P}(\boldsymbol{\xi}_1 \otimes \boldsymbol{\xi}_2) = \boldsymbol{\xi}_2 \otimes \boldsymbol{\xi}_1$$

with the permutation matrix

$$\boldsymbol{P} = \boldsymbol{U}_{2\times 2} = \left(\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right).$$

Because $\boldsymbol{P}$ is symmetric and real, we find $\boldsymbol{P}^{\dagger} = \boldsymbol{P}$. Due to $\boldsymbol{P}^2 = \boldsymbol{I}_4$, we get $\boldsymbol{P}^{-1} = \boldsymbol{P}$. The matrix $\boldsymbol{P}$ therefore only has eigenvalues equal to ±1.

11.1. Schrödinger would assign a wave function to each state of the cat, that is, "alive after an hour" and "dead after an hour." The probabilities for both states would be equal. The total state of the system would then be a superposition of the two wave functions. In other words, the ψ-function of the system would express that the living and the dead cat are mixed in equal parts.

According to the *Copenhagen interpretation* of quantum mechanics, the wave function of the system collapses at the time of measurement. Once you open the chamber and observe the system (i.e., a measurement), the atomic nucleus jumps from the earlier superposition state into one of the eigenstates according to the measurement device. It is therefore only at the time of the measurement (by an outside observer) that it is decided whether the cat is dead or alive. Prior to the measurement, we can only make a probability statement about the status of the cat.

From the perspective of *ensemble theory*, the experiment would be described by an ensemble of identical systems, say 10,000 boxes with one cat in each box. After a certain time interval, approximately 5000 cats would be dead and approximately 5000 cats would be alive. This result is due to the empirical law of large numbers. According to this law, the more often you repeat the experiment, the better the results approach the theoretical probability of 50%. To me, this is the interpretation of a physicist, whereas other interpretations (such as Everett's parallel universe, with one universe for the living cat and one universe for the dead cat) rather belong to the field of philosophy. In any case, we had better speak of Schrödinger's cats, that is, in *plural*.

11.2.

$$(a)\,[p,x]\psi(x) = (px - xp)\psi(x) = \frac{\hbar}{i}\frac{\partial}{\partial x}x\psi(x) - x\frac{\hbar}{i}\frac{\partial}{\partial x}\psi(x) =$$

$$= \frac{\hbar}{i}\psi(x) + x\frac{\hbar}{i}\frac{\partial}{\partial x}\psi(x) - x\frac{\hbar}{i}\frac{\partial}{\partial x}\psi(x) = \frac{\hbar}{i}\psi(x).$$

In other words, $\underline{\underline{[p,x] = \frac{\hbar}{i}}}$.

$$(b)\,[p,x^n]\psi(x) = \frac{\hbar}{i}\frac{\partial}{\partial x}x^n\psi(x) - x^n\frac{\hbar}{i}\frac{\partial}{\partial x}\psi(x) =$$

$$= nx^{n-1}\frac{\hbar}{i}\psi(x) + x^n\frac{\hbar}{i}\frac{\partial}{\partial x}\psi(x) - x^n\frac{\hbar}{i}\frac{\partial}{\partial x}\psi(x) = nx^{n-1}\frac{\hbar}{i}\psi(x);$$

that is, $\underline{\underline{[p, x^n] = n\frac{\hbar}{i}x^{n-1}}}$.

(c)With the commutator rule from Problem 3.7,

$$[\boldsymbol{AB}, \boldsymbol{C}] = \boldsymbol{A}[\boldsymbol{B}, \boldsymbol{C}] + [\boldsymbol{A}, \boldsymbol{C}]\boldsymbol{B},$$

we prove the statement inductively.

$\underline{n = 2 :}$

$$[p^2, x] = [p \cdot p, x] = p\underbrace{[p, x]}_{\frac{\hbar}{i}\boldsymbol{I}} + [p, x]p = 2\frac{\hbar}{i}p.$$

$\underline{n - 1 \rightarrow n :}$

$$\underline{\underline{[p^n, x]}} = [p \cdot p^{n-1}, x] = p\underbrace{[p^{n-1}, x]}_{(n-1)\frac{\hbar}{i}p^{n-2}} + \underbrace{[p, x]}_{\frac{\hbar}{i}}p^{n-1} = \underline{\underline{n\frac{\hbar}{i}p^{n-1}}}.$$

11.3. (a) The differentiation of an operator A with respect to time t is defined as

$$\frac{\mathrm{d}A}{\mathrm{d}t} \overset{\text{def}}{=} \lim_{\epsilon \to 0} \frac{A(t + \epsilon) - A(t)}{\epsilon}.$$

For two time-dependent operators $A(t)$ and $B(t)$, we find

$$\underline{\underline{\frac{\mathrm{d}}{\mathrm{d}t}(A(t)B(t))}} = \lim_{\epsilon \to 0} \frac{A(t + \epsilon)B(t + \epsilon) - A(t)B(t)}{\epsilon} =$$

$$= \lim_{\epsilon \to 0} \left[\frac{[A(t + \epsilon) - A(t)]B(t)}{\epsilon} + \frac{A(t + \epsilon)[B(t + \epsilon) - B(t)]}{\epsilon} \right] =$$

$$= \underline{\underline{\frac{\mathrm{d}A(t)}{\mathrm{d}t}B(t) + A(t)\frac{\mathrm{d}B(t)}{\mathrm{d}t}}}. \tag{A.25}$$

Next, we show inductively that $\frac{\mathrm{d}A(t)^n}{\mathrm{d}t} = \sum_{\nu=1}^{n} A(t)^{\nu-1}\frac{\mathrm{d}A}{\mathrm{d}t}A(t)^{n-\nu}$

$\underline{\mathrm{n} = 2}$:

$$\frac{\mathrm{d}A(t)^2}{\mathrm{d}t} = \frac{\mathrm{d}A(t)}{\mathrm{d}t}A(t) + A(t)\frac{\mathrm{d}A(t)}{\mathrm{d}t} = \sum_{\nu=1}^{2} A(t)^{\nu-1}\frac{\mathrm{d}A}{\mathrm{d}t}A(t)^{2-\nu}.$$

$\underline{n \rightarrow n + 1}$:

$$\underline{\underline{\frac{\mathrm{d}A(t)^{n+1}}{\mathrm{d}t}}} \overset{(A.25)}{=} \frac{\mathrm{d}A(t)^n}{\mathrm{d}t}A(t) + A(t)^n\frac{\mathrm{d}A(t)}{\mathrm{d}t} =$$

$$\overset{n}{=} \sum_{\nu=1}^{n} A(t)^{\nu-1} \frac{\mathrm{d}A}{\mathrm{d}t} A(t)^{n-\nu} + A(t)^n \frac{\mathrm{d}A(t)}{\mathrm{d}t} =$$

$$= \underline{\underline{\sum_{\nu=1}^{n+1} A(t)^{\nu-1} \frac{\mathrm{d}A}{\mathrm{d}t} A(t)^{n+1-\nu}}}.$$

The differentiation of a function $f(A)$ of an operator A with respect to the operator A is defined as

$$\frac{\mathrm{d}f(A)}{\mathrm{d}A} \overset{\text{def}}{=} \lim_{\epsilon \to 0} \frac{f(A + \epsilon \mathbf{1}) - f(A)}{\epsilon},$$

with the neutral element $\mathbf{1}$. We obtain a similar product rule

$$\underline{\underline{\frac{\mathrm{d}}{\mathrm{d}A}(f(A)g(A))}} = \lim_{\epsilon \to 0} \frac{f(A + \epsilon \mathbf{1})g(A + \epsilon \mathbf{1}) - f(A)g(A)}{\epsilon} =$$

$$= \lim_{\epsilon \to 0} \left[\frac{[f(A + \epsilon \mathbf{1}) - f(A)]g(A)}{\epsilon} + \frac{f(A + \epsilon \mathbf{1})[g(t + \epsilon \mathbf{1}) - g(A)]}{\epsilon} \right] =$$

$$= \underline{\underline{\frac{\mathrm{d}f(A)}{\mathrm{d}A} g(A) + f(A) \frac{\mathrm{d}g(A)}{\mathrm{d}A}}}.$$

12.1. The gamma matrices are defined as

$$\boldsymbol{\gamma}_i \overset{\text{def}}{=} \boldsymbol{\beta} \boldsymbol{\alpha}_i.$$

With (12.38) to (12.41), where $\boldsymbol{\sigma}_j$ are the Pauli matrices, we get

$$\boldsymbol{\alpha}_1 = \boldsymbol{\sigma}_1 \otimes \boldsymbol{\sigma}_1 = \begin{pmatrix} \mathbf{0} & \boldsymbol{\sigma}_1 \\ \boldsymbol{\sigma}_1 & \mathbf{0} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix},$$

$$\boldsymbol{\alpha}_2 = \boldsymbol{\sigma}_1 \otimes \boldsymbol{\sigma}_2 = \begin{pmatrix} \mathbf{0} & \boldsymbol{\sigma}_2 \\ \boldsymbol{\sigma}_2 & \mathbf{0} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix},$$

$$\boldsymbol{\alpha}_3 = \boldsymbol{\sigma}_1 \otimes \boldsymbol{\sigma}_3 = \begin{pmatrix} \mathbf{0} & \boldsymbol{\sigma}_3 \\ \boldsymbol{\sigma}_3 & \mathbf{0} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix},$$

$$\boldsymbol{\beta} = \boldsymbol{\sigma}_3 \otimes \boldsymbol{I}_2 = \begin{pmatrix} \boldsymbol{I}_2 & \mathbf{0} \\ \mathbf{0} & -\boldsymbol{I}_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

The gamma matrices are therefore

$$\gamma_0 = \boldsymbol{\beta} = \begin{pmatrix} \boldsymbol{I}_2 & \mathbf{0} \\ \mathbf{0} & -\boldsymbol{I}_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \tag{A.26}$$

$$\gamma_1 = \boldsymbol{\beta}\boldsymbol{\alpha}_1 = \begin{pmatrix} \boldsymbol{I}_2 & \mathbf{0} \\ \mathbf{0} & -\boldsymbol{I}_2 \end{pmatrix} \begin{pmatrix} \mathbf{0} & \boldsymbol{\sigma}_1 \\ \boldsymbol{\sigma}_1 & \mathbf{0} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \tag{A.27}$$

$$\gamma_2 = \boldsymbol{\beta}\boldsymbol{\alpha}_2 = \begin{pmatrix} \boldsymbol{I}_2 & \mathbf{0} \\ \mathbf{0} & -\boldsymbol{I}_2 \end{pmatrix} \begin{pmatrix} \mathbf{0} & \boldsymbol{\sigma}_2 \\ \boldsymbol{\sigma}_2 & \mathbf{0} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}, \tag{A.28}$$

$$\gamma_3 = \boldsymbol{\beta}\boldsymbol{\alpha}_3 = \begin{pmatrix} \boldsymbol{I}_2 & \mathbf{0} \\ \mathbf{0} & -\boldsymbol{I}_2 \end{pmatrix} \begin{pmatrix} \mathbf{0} & \boldsymbol{\sigma}_3 \\ \boldsymbol{\sigma}_3 & \mathbf{0} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \tag{A.29}$$

Note that for $j = 1, 2$, and 3, we can also write the gamma matrices as

$$\gamma_j = \begin{pmatrix} \mathbf{0} & \boldsymbol{\sigma}_j \\ -\boldsymbol{\sigma}_j & \mathbf{0} \end{pmatrix}. \tag{A.30}$$

In this form, we immediately find that

$$\gamma_j^2 = \begin{pmatrix} \mathbf{0} & \boldsymbol{\sigma}_j \\ -\boldsymbol{\sigma}_j & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{0} & \boldsymbol{\sigma}_j \\ -\boldsymbol{\sigma}_j & \mathbf{0} \end{pmatrix} = \begin{pmatrix} -\boldsymbol{I}_2 & \mathbf{0} \\ \mathbf{0} & -\boldsymbol{I}_2 \end{pmatrix} = -\boldsymbol{I}_4, \tag{A.31}$$

because $\boldsymbol{\sigma}_j^2 = \boldsymbol{I}_2$ for all j. With the *Levi-Civita symbol*

$$\varepsilon_{ijk} = \begin{cases} +1 & \text{if } (i, j, k) \text{ is } (1, 2, 3), (2, 3, 1) \text{ or } (3, 1, 2), \\ -1 & \text{if } (i, j, k) \text{ is } (3, 2, 1), (1, 3, 2) \text{ or } (2, 1, 3), \\ 0 & \text{if } i = j \text{ or } j = k \text{ or } k = i, \end{cases}$$

we find for the product of two Pauli matrices ($\mu \neq \nu$)

$$\boldsymbol{\sigma}_\mu \boldsymbol{\sigma}_\nu = i \sum_\kappa \varepsilon_{\mu\nu\kappa} \boldsymbol{\sigma}_\kappa. \tag{A.32}$$

With (A.31), we finally get ($\mu \neq \nu$)

$$\gamma_\mu \gamma_\nu = \begin{pmatrix} \mathbf{0} & \boldsymbol{\sigma}_\mu \\ -\boldsymbol{\sigma}_\mu & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{0} & \boldsymbol{\sigma}_\nu \\ -\boldsymbol{\sigma}_\nu & \mathbf{0} \end{pmatrix} =$$

$$= \begin{pmatrix} -\boldsymbol{\sigma}_\mu \boldsymbol{\sigma}_\nu & \mathbf{0} \\ \mathbf{0} & -\boldsymbol{\sigma}_\mu \boldsymbol{\sigma}_\nu \end{pmatrix} = \begin{pmatrix} -i \sum_\kappa \varepsilon_{\mu\nu\kappa} \boldsymbol{\sigma}_\kappa & \mathbf{0} \\ \mathbf{0} & -i \sum_\kappa \varepsilon_{\mu\nu\kappa} \boldsymbol{\sigma}_\kappa \end{pmatrix}. \tag{A.33}$$

For example, we get

$$\gamma_1 \gamma_2 = \begin{pmatrix} -i\boldsymbol{\sigma}_3 & \mathbf{0} \\ \mathbf{0} & -i\boldsymbol{\sigma}_3 \end{pmatrix} = \begin{pmatrix} -i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & 0 & 0 & i \end{pmatrix}.$$

From (A.33), we immediately see that

$$\gamma_\mu \gamma_\nu = -\gamma_\nu \gamma_\mu, \tag{A.34}$$

which implies that the matrices *anti-commute*:

$$\{\gamma_\mu, \gamma_\nu\} \overset{\text{def}}{=} \gamma_\mu \gamma_\nu + \gamma_\nu \gamma_\mu = \mathbf{0}. \tag{A.35}$$

The commutation law also follows from (A.33) as

$$[\gamma_\mu, \gamma_\nu] \overset{\text{def}}{=} \gamma_\mu \gamma_\nu - \gamma_\nu \gamma_\mu = 2 \begin{pmatrix} -i \sum_\kappa \varepsilon_{\mu\nu\kappa} \boldsymbol{\sigma}_\kappa & \mathbf{0} \\ \mathbf{0} & -i \sum_\kappa \varepsilon_{\mu\nu\kappa} \boldsymbol{\sigma}_\kappa \end{pmatrix}. \tag{A.36}$$

12.2. Using (2.5) to (2.4), we first obtain

$$i\,\gamma_0 \frac{\partial}{\partial x_0} = \begin{pmatrix} i\frac{\partial}{\partial x_0} & 0 & 0 & 0 \\ 0 & i\frac{\partial}{\partial x_0} & 0 & 0 \\ 0 & 0 & -i\frac{\partial}{\partial x_0} & 0 \\ 0 & 0 & 0 & -i\frac{\partial}{\partial x_0} \end{pmatrix}, \tag{A.37}$$

$$i\,\gamma_1\frac{\partial}{\partial x_1}=\begin{pmatrix} 0 & 0 & 0 & i\frac{\partial}{\partial x_1} \\ 0 & 0 & i\frac{\partial}{\partial x_1} & 0 \\ 0 & -i\frac{\partial}{\partial x_1} & 0 & 0 \\ -i\frac{\partial}{\partial x_1} & 0 & 0 & 0 \end{pmatrix}, \tag{A.38}$$

$$i\,\gamma_1\frac{\partial}{\partial x_2}=\begin{pmatrix} 0 & 0 & 0 & \frac{\partial}{\partial x_2} \\ 0 & 0 & -\frac{\partial}{\partial x_2} & 0 \\ 0 & -\frac{\partial}{\partial x_2} & 0 & 0 \\ \frac{\partial}{\partial x_2} & 0 & 0 & 0 \end{pmatrix}, \tag{A.39}$$

$$i\,\gamma_1\frac{\partial}{\partial x_3}=\begin{pmatrix} 0 & 0 & i\frac{\partial}{\partial x_3} & 0 \\ 0 & 0 & 0 & -i\frac{\partial}{\partial x_3} \\ -i\frac{\partial}{\partial x_3} & 0 & 0 & 0 \\ 0 & i\frac{\partial}{\partial x_3} & 0 & 0 \end{pmatrix}. \tag{A.40}$$

Adding these four matrices yields

$$i\not{\partial}=i\sum_{j=0}^{3}\gamma_j\frac{\partial}{\partial x_j}=$$

$$=\begin{pmatrix} i\frac{\partial}{\partial x_0} & 0 & i\frac{\partial}{\partial x_3} & i\frac{\partial}{\partial x_1}+\frac{\partial}{\partial x_2} \\ 0 & i\frac{\partial}{\partial x_0} & i\frac{\partial}{\partial x_1}-\frac{\partial}{\partial x_2} & -i\frac{\partial}{\partial x_3} \\ -i\frac{\partial}{\partial x_3} & -i\frac{\partial}{\partial x_1}-\frac{\partial}{\partial x_2} & -i\frac{\partial}{\partial x_0} & 0 \\ \frac{\partial}{\partial x_2}-i\frac{\partial}{\partial x_1} & i\frac{\partial}{\partial x_3} & 0 & -i\frac{\partial}{\partial x_0} \end{pmatrix}. \tag{A.41}$$

This takes us to four linear differential equations:

$$(i\not{\partial}-\frac{m_0}{\hbar}\boldsymbol{I}_4)\psi=$$

$$\begin{pmatrix} i\dfrac{\partial}{\partial x_0}-\dfrac{m_0}{\hbar} & 0 & i\dfrac{\partial}{\partial x_3} & i\dfrac{\partial}{\partial x_1}+\dfrac{\partial}{\partial x_2} \\ 0 & i\dfrac{\partial}{\partial x_0}-\dfrac{m_0}{\hbar} & i\dfrac{\partial}{\partial x_1}-\dfrac{\partial}{\partial x_2} & -i\dfrac{\partial}{\partial x_3} \\ -i\dfrac{\partial}{\partial x_3} & -i\dfrac{\partial}{\partial x_1}-\dfrac{\partial}{\partial x_2} & -i\dfrac{\partial}{\partial x_0}-\dfrac{m_0}{\hbar} & 0 \\ \dfrac{\partial}{\partial x_2}-i\dfrac{\partial}{\partial x_1} & i\dfrac{\partial}{\partial x_3} & 0 & -i\dfrac{\partial}{\partial x_0}-\dfrac{m_0}{\hbar} \end{pmatrix} \begin{pmatrix} \psi_0 \\ \psi_1 \\ \psi_2 \\ \psi_3 \end{pmatrix} =$$

$$\begin{pmatrix} i\dfrac{\partial\psi_0}{\partial x_0}+i\dfrac{\partial\psi_3}{\partial x_1}+\dfrac{\partial\psi_3}{\partial x_2}+i\dfrac{\partial\psi_2}{\partial x_3}-\dfrac{m_0}{\hbar}\psi_0 \\ i\dfrac{\partial\psi_1}{\partial x_0}+i\dfrac{\partial\psi_2}{\partial x_1}-\dfrac{\partial\psi_2}{\partial x_2}-i\dfrac{\partial\psi_3}{\partial x_3}-\dfrac{m_0}{\hbar}\psi_1 \\ -i\dfrac{\partial\psi_2}{\partial x_0}-i\dfrac{\partial\psi_1}{\partial x_1}-\dfrac{\partial\psi_1}{\partial x_2}-i\dfrac{\partial\psi_0}{\partial x_3}-\dfrac{m_0}{\hbar}\psi_2 \\ -i\dfrac{\partial\psi_3}{\partial x_0}-i\dfrac{\partial\psi_0}{\partial x_1}+\dfrac{\partial\psi_0}{\partial x_2}+i\dfrac{\partial\psi_1}{\partial x_3}-\dfrac{m_0}{\hbar}\psi_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{A.42}$$

With the ansatz of a plane wave

$$\boldsymbol{\psi}(\boldsymbol{x},t) = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} \exp(i(\boldsymbol{k}^{\mathsf{T}}\boldsymbol{x}-\omega t)), \tag{A.43}$$

we obtain for the first derivative

$$\frac{\partial}{\partial x_0}\,\boldsymbol{\psi} = \boldsymbol{c}\,(-i\,\omega/c)\exp(i(\boldsymbol{k}^{\mathsf{T}}\boldsymbol{x}-\omega t)). \tag{A.44}$$

With $\boldsymbol{k}^{\mathsf{T}}\boldsymbol{x} = k_1x_1+k_2x_2+k_3x_3$, we further get

$$\frac{\partial}{\partial x_j}\,\boldsymbol{\psi} = \boldsymbol{c}\,(i\,k_j)\exp(i(\boldsymbol{k}^{\mathsf{T}}\boldsymbol{x}-\omega t)). \tag{A.45}$$

If we insert these derivatives into (A.42), we finally get the linear algebraic equation

$$\begin{pmatrix} \frac{\omega}{c}-\frac{m_0}{\hbar} & 0 & -k_3 & ik_2-k_1 \\ 0 & \frac{\omega}{c}-\frac{m_0}{\hbar} & -k_1-ik_2 & k_3 \\ k_3 & k_1-ik_2 & -\frac{\omega}{c}-\frac{m_0}{\hbar} & 0 \\ k_1+ik_2 & -k_3 & 0 & -\frac{\omega}{c}-\frac{m_0}{\hbar} \end{pmatrix} \boldsymbol{c} = \boldsymbol{0}. \tag{A.46}$$

Note that there is an algebraic computer program (written in MAXIMA) in [23].

Appendix B
The Kronecker Product

B.1 Definitions

The Kronecker product of two matrices $\boldsymbol{A} \in \mathbb{C}^{n \times m}$ and $\boldsymbol{B} \in \mathbb{C}^{p \times q}$ leads to a matrix $\boldsymbol{C} \in \mathbb{C}^{np \times mq}$ and is denoted by

$$\boldsymbol{A} \otimes \boldsymbol{B} = \boldsymbol{C}.$$

The submatrices $\boldsymbol{C}_{ij} \in \mathbb{C}^{p \times q}$ for $i = 1$ to n and for $j = 1$ to m are defined as

$$\boldsymbol{C}_{ij} \overset{\text{def}}{=} a_{ij} \boldsymbol{B},$$

the matrix $\boldsymbol{C}$ therefore has the form

$$\boldsymbol{C} = \begin{pmatrix} a_{11}\boldsymbol{B} & a_{12}\boldsymbol{B} & \ldots & a_{1m}\boldsymbol{B} \\ a_{21}\boldsymbol{B} & a_{22}\boldsymbol{B} & \ldots & a_{2m}\boldsymbol{B} \\ \ldots & & & \\ a_{n1}\boldsymbol{B} & a_{n2}\boldsymbol{B} & \ldots & a_{nm}\boldsymbol{B} \end{pmatrix}.$$

B.2 Properties of the Kronecker Product

Without proof, the following important properties also hold (see, e.g., [7]):

$$(\boldsymbol{A} \otimes \boldsymbol{B}) \otimes \boldsymbol{C} = \boldsymbol{A} \otimes (\boldsymbol{B} \otimes \boldsymbol{C}), \tag{B.1}$$

$$(\boldsymbol{A} \otimes \boldsymbol{B})^{\intercal} = \boldsymbol{A}^{\intercal} \otimes \boldsymbol{B}^{\intercal}, \tag{B.2}$$

$$(\boldsymbol{A} \otimes \boldsymbol{B})(\boldsymbol{C} \otimes \boldsymbol{D}) = \boldsymbol{A}\boldsymbol{C} \otimes \boldsymbol{B}\boldsymbol{D}. \tag{B.3}$$

G. Ludyk, *Quantum Mechanics in Matrix Form*,
https://doi.org/10.1007/978-3-319-26366-3

Appendix C
Fourier Decomposition of Periodic Functions

Let $f(t)$ be an arbitrary periodic function. Its Fourier decomposition is given by the sum of simple sine waves

$$f(t)=\frac{y_0}{2}+\sum_{n=1}^{\infty} y_n\cdot\sin(n\omega t+\varphi_n),$$

where the y_n are the *amplitudes* and the φ_n are the *phases*. With the help of the trigonometric addition theorem, the function can be decomposed into

$$f(t)=\frac{y_0}{2}+\sum_{n=1}^{\infty} a_n\cdot\cos(n\omega t)+\sum_{n=1}^{\infty} b_n\cdot\sin(n\omega t). \tag{C.1}$$

The goal of *Fourier analysis* is to compute the coefficients a_n and b_n (with the help of integrals over a full period).

We can also write the Fourier series with the help of complex numbers as

$$f(t)=\sum_{n=-\infty}^{\infty} c_n e^{in\omega t}.$$

The connection with the above notation with sine and cosine expressions is obtained as

$$\sum_{n=-\infty}^{\infty} c_n e^{in\omega t}=c_0+\sum_{n=1}^{\infty} c_n e^{in\omega t}+\sum_{n=-1}^{-\infty} c_n e^{in\omega t}$$

$$=c_0+\sum_{n=1}^{\infty} c_n(\cos n\omega t+i\,\sin n\omega t)+\sum_{n=-1}^{-\infty} c_n(\cos n\omega t+i\,\sin n\omega t)$$

G. Ludyk, *Quantum Mechanics in Matrix Form*,
https://doi.org/10.1007/978-3-319-26366-3

$$= c_0 + \sum_{n=1}^{\infty} c_n (\cos n\omega t + i \sin n\omega t) + \sum_{n=1}^{\infty} c_n (\cos n\omega t - i \sin n\omega t)$$

$$= c_0 + \sum_{n=1}^{\infty} \left[(c_n + c_{-n}) \cos n\omega t + i (c_n - c_{-n}) \sin n\omega t \right]. \tag{C.2}$$

A comparison of (C.1) with (C.2) yields

$$c_o = \frac{y_0}{2},$$

$$a_n = c_n + c_{-n}$$

and

$$b_n = i(c_n - c_{-n}),$$

or

$$c_n = \frac{a_n - ib_n}{2} \quad \text{and} \quad c_{-n} = \frac{a_n + ib_n}{2} = c_n^*.$$

The main idea of Fourier analysis is to compare the signal with sinusoids of various frequencies ω_n. Each such sinusoid may be thought of as a prototype oscillation. As a result, we obtain for each considered frequency parameter ω_n a magnitude coefficient $y_n \geq 0$, along with a phase coefficient φ_n. In the case that the coefficient y_n is large, there is a high similarity between the signal and the sinusoid of frequency ω_n, and the signal contains a periodic oscillation at that frequency. In the case that y_n is small, the signal does not contain a remarkable periodic component at that frequency. The original signal and the Fourier transform contain the same amount of information. This information, however, is represented in different ways. Whereas the signal displays the information across time, the Fourier transform displays the information across frequency.

Appendix D
Laplace–Runge–Lenz Vector

In a potential of the form $V(r) = -\alpha/r$ (e.g., the Coulomb potential or the gravitational potential), the Laplace–Runge–Lenz vector is a conserved entity. In classical mechanics, the vector is primarily used to describe the shape and orientation of orbits of astronomical bodies, such as the path of a planet around its star. For two bodies that move according to Newtonian physics, the Laplace–Runge–Lenz vector is a constant of motion; that is, it is the same on every point of the path. With

$$\boldsymbol{\ell} = \boldsymbol{r} \times \boldsymbol{p} \tag{D.1}$$

the (time-constant) angular momentum of the electron around the nucleus, and

$$\boldsymbol{p} = m\,\boldsymbol{v}$$

its linear momentum, the respective (constant) Laplace–Runge–Lenz vector is defined as

$$\boldsymbol{a} \overset{\text{def}}{=} \frac{1}{Ze^2m}(\boldsymbol{\ell} \times \boldsymbol{p}) + \boldsymbol{r}/r. \tag{D.2}$$

Let us now prove that the Laplace–Runge–Lenz vector is indeed constant over time. With $k \overset{\text{def}}{=} Z^2e$, we get

$$\boldsymbol{a} = \frac{1}{k\,m}(\boldsymbol{\ell} \times \boldsymbol{p}) + \boldsymbol{r}/r. \tag{D.3}$$

In addition, the time derivative of the momentum for the potential $V = -k/r$ is

$$\dot{\boldsymbol{p}} = -k\,\boldsymbol{r}/r^3, \tag{D.4}$$

that is, Newton's law of motion. Now

G. Ludyk, *Quantum Mechanics in Matrix Form*,
https://doi.org/10.1007/978-3-319-26366-3

$$\dot{\boldsymbol{a}} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{k\,m}(\boldsymbol{\ell}\times\boldsymbol{p}) + \boldsymbol{r}/r\right) =$$

$$= \frac{1}{k\,m}(\underbrace{\dot{\boldsymbol{\ell}}}_{0}\times\boldsymbol{p} + \boldsymbol{\ell}\times\dot{\boldsymbol{p}}) + \frac{\dot{\boldsymbol{r}}}{r} + \boldsymbol{r}\left(\left(\frac{\partial}{\partial r}\frac{1}{r}\right)\cdot\dot{\boldsymbol{r}}\right) =$$

$$= \frac{1}{k\,m}(\boldsymbol{r}\times\boldsymbol{p})\times\dot{\boldsymbol{p}} + \frac{1}{m}\frac{\boldsymbol{p}}{r} + \boldsymbol{r}\left(\boldsymbol{r}\frac{-1}{r^3}\cdot\dot{\boldsymbol{r}}\right).$$

With (D.4) and the vector product rule $\boldsymbol{\alpha}\times\boldsymbol{\beta}\times\boldsymbol{\gamma} = (\boldsymbol{\alpha}\cdot\boldsymbol{\gamma})\boldsymbol{\beta} - (\boldsymbol{\alpha}\cdot\boldsymbol{\beta})\boldsymbol{\gamma}$, we find indeed

$$\dot{\boldsymbol{a}} = -\frac{1}{mr^3}\underbrace{(\boldsymbol{r}\times\boldsymbol{p})\times\boldsymbol{r}}_{(\boldsymbol{r}\cdot\boldsymbol{r})\boldsymbol{p}-(\boldsymbol{r}\cdot\boldsymbol{p})\,\boldsymbol{r}} + \frac{1}{m}\frac{\boldsymbol{p}}{r} + \boldsymbol{r}\left(\boldsymbol{r}\frac{-1}{r^3}\cdot\underbrace{\dot{\boldsymbol{r}}}_{\frac{1}{m}\boldsymbol{p}}\right) =$$

$$= -\frac{1}{m\,r}\boldsymbol{p} + \frac{1}{m\,r^3}(\boldsymbol{r}\cdot\boldsymbol{p})\,\boldsymbol{r} + \frac{1}{m\,r}\boldsymbol{p} - \boldsymbol{r}(\boldsymbol{r}\cdot\boldsymbol{p})\frac{1}{m\,r^3} = 0.$$

D.1 Further Properties of the Vector *a*

Lemma *The vector $\boldsymbol{a}$ is perpendicular to the angular momentum vector $\boldsymbol{\ell}$.*

Proof The vector $\boldsymbol{\ell}\times\boldsymbol{p}$ is perpendicular to $\boldsymbol{\ell}$. Also, the vector $\boldsymbol{\ell} = \boldsymbol{r}\times\boldsymbol{p}$ is perpendicular to $\boldsymbol{r}$. This implies

$$\boldsymbol{\ell}\cdot\boldsymbol{a} = \frac{1}{Ze^2m}\underbrace{\boldsymbol{\ell}\cdot(\boldsymbol{\ell}\times\boldsymbol{p})}_{0} + \underbrace{\boldsymbol{\ell}\cdot\boldsymbol{r}}_{0}/r = 0. \quad \textbf{q.e.d.}$$

Lemma *The scalar product of $\boldsymbol{a}$ and $\boldsymbol{r}$ is*

$$\boldsymbol{a}\cdot\boldsymbol{r} = -\frac{1}{Ze^2m}\boldsymbol{\ell}^2 + r.$$

Proof With $(\boldsymbol{\alpha}\times\boldsymbol{\beta})\cdot\boldsymbol{\gamma} = \boldsymbol{\alpha}\cdot(\boldsymbol{\beta}\times\boldsymbol{\gamma})$, we find

$$\boldsymbol{a}\cdot\boldsymbol{r}=\frac{1}{Ze^2m}\underbrace{(\boldsymbol{\ell}\times\boldsymbol{p})\cdot\boldsymbol{r}}_{\boldsymbol{\ell}\,\underbrace{(\boldsymbol{p}\times\boldsymbol{r})}_{-\boldsymbol{\ell}}}+\frac{(\boldsymbol{r}\cdot\boldsymbol{r})}{r}=-\frac{1}{Ze^2m}\boldsymbol{\ell}^2+r.\quad \mathbf{q.e.d.}$$

Lemma

$$|\boldsymbol{a}|=\sqrt{\frac{2\,H}{Z^2e^4m}\boldsymbol{\ell}^2+1}. \tag{D.5}$$

Proof With the Lagrange identity

$$(\boldsymbol{\alpha}\times\boldsymbol{\beta})\cdot(\boldsymbol{\gamma}\times\boldsymbol{\delta})=(\boldsymbol{\alpha}\cdot\boldsymbol{\gamma})(\boldsymbol{\beta}\cdot\boldsymbol{\delta})-(\boldsymbol{\alpha}\cdot\boldsymbol{\delta})(\boldsymbol{\beta}\cdot\boldsymbol{\gamma})$$

and the Hamilton function $H=\frac{1}{2m}\boldsymbol{p}^2-\frac{Ze^2}{r}$, we find

$$|\boldsymbol{a}|^2=\boldsymbol{a}\cdot\boldsymbol{a}=\left|\frac{1}{Ze^2m}(\boldsymbol{\ell}\times\boldsymbol{p})+\boldsymbol{r}/r\right|^2=$$

$$=\frac{1}{Z^2e^4m^2}\underbrace{(\boldsymbol{\ell}\times\boldsymbol{p})\cdot(\boldsymbol{\ell}\times\boldsymbol{p})}_{(\boldsymbol{\ell}^2)(\boldsymbol{p}^2)-\underbrace{(\boldsymbol{\ell}\boldsymbol{p})}_{0}(\boldsymbol{p}\boldsymbol{\ell})}+\frac{2}{Ze^2m}\underbrace{(\boldsymbol{\ell}\times\boldsymbol{p})\cdot\boldsymbol{r}}_{-\boldsymbol{\ell}^2}/r+1=$$

$$=\frac{2}{Z^2e^4m}\boldsymbol{\ell}^2\underbrace{\left(\frac{1}{2m}\boldsymbol{p}^2-\frac{Ze^2}{r}\right)}_{H}+1.\quad \mathbf{q.e.d.}$$

Appendix E
Permutation

A permutation without repetition is an arrangement of n objects. For the first object, we have n placement possibilities, for the second object only $n - 1$ possibilities, for the third object only $n-2$, and so on. For the last item, there is only one empty seat left. The number of possible permutations of n objects is therefore $n! = n \cdot (n-1) \cdot \cdots \cdot 1$. For example, there are $4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$ possible ways to arrange four differently colored balls.

Let us now introduce a more precise representation of an n digit permutation π as Cauchy's two-line notation with two rows and n columns. In the top row, we place the numbers 1 to n (in any order). Under each number j, we put the function value $\pi(j)$ in the second line. For example,

$$\pi = \begin{pmatrix} 1 & 2 & \cdots & n \\ \pi(1) & \pi(2) & \cdots & \pi(n) \end{pmatrix}.$$

Obviously, the second row also contains each number 1 to n exactly once. As an example, let us consider the permutation

$$\pi\colon \{1, 2, 3, 4\} \rightarrow \{2, 4, 3, 1\}$$

with $\pi(1) = 2$, $\pi(2) = 4$, $\pi(3) = 3$, and $\pi(4) = 1$. In Cauchy's two-line notation, this permutation looks like

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix}.$$

One can represent a permutation of n objects as an $n \times n$ matrix. If a permutation $\pi : \{1, \ldots, n\} \rightarrow \{1, \ldots, n\}$ of n objects is given, then the permutation matrix $\boldsymbol{P}_\pi$ is defined as

$$\boldsymbol{P}_\pi \overset{\text{def}}{=} \begin{pmatrix} \vec{p}_1 & \ldots & \vec{p}_n \end{pmatrix},$$

G. Ludyk, *Quantum Mechanics in Matrix Form*,
https://doi.org/10.1007/978-3-319-26366-3

where $\vec{p}_i$ is a *canonical basis vector*,[1] thus a permutation matrix is a square binary matrix that has exactly one entry of 1 in each row and each column and 0s elsewhere. As an example, consider the permutation

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 1 & 5 & 3 \end{pmatrix}.$$

The corresponding permutation matrix now has the form

$$\boldsymbol{P}_\pi = \begin{pmatrix} \vec{p}_1 & \vec{p}_2 & \vec{p}_3 & \vec{p}_4 & \vec{p}_5 \end{pmatrix} =$$

$$= \begin{pmatrix} \vec{e}_4 & \vec{e}_2 & \vec{e}_1 & \vec{e}_5 & \vec{e}_3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

For a vector $\vec{v} = (v_1, v_2, v_3, v_4, v_5)^\intercal$, this yields

$$\boldsymbol{P}_\pi \vec{v} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{pmatrix} = \begin{pmatrix} v_3 \\ v_2 \\ v_5 \\ v_1 \\ v_4 \end{pmatrix}.$$

---

[1] The canonical basis vector $\vec{e}_k$ is the kth column of the $n \times n$-identity matrix $\boldsymbol{I}_n$.

Appendix F
Determinants

F.1 Axiomatic Definition of the Determinant of a Square Matrix

A mapping from the space of square matrices onto the underlying field maps each matrix $\boldsymbol{A} = (\boldsymbol{v}_1|\ldots|\boldsymbol{v}_n)$ on its determinant $det\boldsymbol{A}$, if it satisfies the following three properties.

- It is *multilinear*, that is,

$$\det(\boldsymbol{v}_1|\ldots|\boldsymbol{v}_{i-1}|b\boldsymbol{v}_i + c\boldsymbol{w}|\boldsymbol{v}_{i+1}|\ldots|\boldsymbol{v}_n) =$$

$$= b\det(\boldsymbol{v}_1|\ldots|\boldsymbol{v}_{i-1}|\boldsymbol{v}_i|\boldsymbol{v}_{i+1}|\ldots|\boldsymbol{v}_n) +$$

$$+ c\det(\boldsymbol{v}_1|\ldots|\boldsymbol{v}_{i-1}|\boldsymbol{w}|\boldsymbol{v}_{i+1}|\ldots|\boldsymbol{v}_n)$$

  for all $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n, \boldsymbol{w} \in V$, and for all $b, c \in K$.
- If we interchange two columns or two rows, the sign is changed

$$\det(\boldsymbol{v}_1|\ldots|\boldsymbol{v}_i|\ldots|\boldsymbol{v}_j|\ldots|\boldsymbol{v}_n) = -\det(\boldsymbol{v}_1|\ldots|\boldsymbol{v}_j|\ldots|\boldsymbol{v}_i|\ldots|\boldsymbol{v}_n).$$

- If two columns or two rows are equal, the determinant is zero.
- It is normalized, that is,

$$\det \boldsymbol{I}_n = 1$$

  for the identity matrix $\boldsymbol{I}_n$.

G. Ludyk, *Quantum Mechanics in Matrix Form*,
https://doi.org/10.1007/978-3-319-26366-3

F.2 The Leibniz Formula

For an $n \times n$ matrix $\boldsymbol{A}$ with elements $a_{i,j}$, the determinant can be defined by the Leibniz formula

$$\det \boldsymbol{A} = \sum_{\sigma \in S_n} \left(\operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)} \right).$$

The sum stretches over all permutations[2] σ of the symmetric group S_n of degree n. $\operatorname{sgn}(\sigma)$ denotes the sign of the permutation σ, that is, $+1$ if σ is an even permutation, and -1 if it is odd. We can recognize whether a permutation is even or odd by counting the number of transpositions that are required to generate the permutation. An even number of permutations means that the permutation is even; an odd number of permutations means that the permutation is odd.

---

[2] See Appendix E.

Appendix G
Dirac's Bra-Ket Notation

In [9], Dirac introduced a special notation for quantum mechanical states. Today, this notation is used in many books on quantum mechanics and should therefore be briefly described here as well.

The state space in quantum mechanics is a complex finite or infinite vector space. Dirac denotes an element $\boldsymbol{f}$ of the vector space by $|f\,\rangle$, which he then calls a *ket* vector. An example for a one-dimensional ket is Schrödinger's wave function $|\psi\,\rangle$, whose representation in position space is the well-known complex-valued wave function $\psi(\boldsymbol{x})$. An example for a four-dimensional ket is the vector

$$|\psi\,\rangle \overset{\text{def}}{=} \boldsymbol{\psi} = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix}$$

in the Dirac equation. Note that the ket $|\psi\,\rangle$ stands for the *entire* wave function $\boldsymbol{\psi}$!

Dirac then defines a *dual* to each ket called the *bra*. We get a bra from the respective ket by taking its conjugate complex (if the ket is a vector, we also need to transpose):

$$\langle f\,| \overset{\text{def}}{=} (\boldsymbol{f}^*)^\intercal = \boldsymbol{f}^\dagger.$$

The *scalar product* of two vectors can then be written with the bra and the ket as

$$\langle f\,|\,g\,\rangle \overset{\text{def}}{=} \boldsymbol{f}^\dagger \boldsymbol{g}.$$

We then have

$$\langle f\,|g\,\rangle = \langle g\,|f\,\rangle^*.$$

For an operator C, the definition can be extended to

$$\langle f\,|C|g\,\rangle \overset{\text{def}}{=} \boldsymbol{f}^\dagger \boldsymbol{C} \boldsymbol{g}.$$

G. Ludyk, *Quantum Mechanics in Matrix Form*,
https://doi.org/10.1007/978-3-319-26366-3

This definition allows us to write the operation of C in two equivalent versions, depending on to which state we prefer to apply the operator:

$$\langle f \,|C|g\,\rangle = \langle C^\dagger f \,|g\,\rangle = \langle f \,|Cg\,\rangle.$$

The length of a vector $\boldsymbol{f}$ is given by

$$|\boldsymbol{f}| = \sqrt{\langle f \,|f\,\rangle}.$$

For the case of wave functions in position space, the scalar product is defined as

$$\langle f \,|g\,\rangle \overset{\text{def}}{=} \int_{-\infty}^{+\infty} f^*(x)g(x)\mathrm{d}x.$$

Appendix H
Proofs of Pauli's Formulas

H.1 'Commutator Gymnastics'

First we notate formulas that we need to perform the proofs. We have ($k = 1, 2$, and 3) for $\boldsymbol{F} = \boldsymbol{F}(\boldsymbol{X}_k, \boldsymbol{P}_k)$

$$(3.27) \quad [\boldsymbol{X}_k, \boldsymbol{F}] = i\hbar \frac{\partial \boldsymbol{F}}{\partial \boldsymbol{P}_k},$$

$$(3.28) \quad [\boldsymbol{P}_k, \boldsymbol{F}] = -i\hbar \frac{\partial \boldsymbol{F}}{\partial \boldsymbol{X}_k}.$$

For $\boldsymbol{F} = \boldsymbol{R} = \sqrt{\boldsymbol{X}_1^2 + \boldsymbol{X}_2^2 + \boldsymbol{X}_3^2}$ we get

$$\underline{\underline{[\boldsymbol{X}_k, \boldsymbol{R}]}} = i\hbar \frac{\partial \boldsymbol{R}}{\partial \boldsymbol{P}_k} = \underline{\underline{\boldsymbol{0}}}, \tag{H.1}$$

and

$$\underline{\underline{[\boldsymbol{P}_k, \boldsymbol{R}]}} = -i\hbar \frac{\partial \boldsymbol{R}}{\partial \boldsymbol{X}_k} = \underline{\underline{-i\hbar \boldsymbol{X}_k \boldsymbol{R}^{-1}}}. \tag{H.2}$$

Also

$$\underline{\underline{[\boldsymbol{P}_k, \boldsymbol{R}^{-1}]}} = -i\hbar \frac{\partial (\boldsymbol{R}^{-1})}{\partial \boldsymbol{X}_k} = \underline{\underline{-i\hbar \boldsymbol{X}_k (\boldsymbol{R}^3)^{-1}}}. \tag{H.3}$$

Furthermore

$$[\boldsymbol{L}_1, \boldsymbol{R}] = (\boldsymbol{X}_2\boldsymbol{P}_3 - \boldsymbol{X}_3\boldsymbol{P}_2)\boldsymbol{R} - \boldsymbol{R}(\boldsymbol{X}_2\boldsymbol{P}_3 - \boldsymbol{X}_3\boldsymbol{P}_2) = \boldsymbol{X}_2 \underbrace{[\boldsymbol{P}_3, \boldsymbol{R}]}_{-i\hbar \boldsymbol{X}_3 \boldsymbol{R}^{-1}} + \boldsymbol{X}_3 \underbrace{[\boldsymbol{R}, \boldsymbol{P}_2]}_{i\hbar \boldsymbol{X}_2 \boldsymbol{R}^{-1}} = \boldsymbol{0},$$

generally

$$\underline{\underline{[\boldsymbol{L}_k, \boldsymbol{R}] = \boldsymbol{0}}}. \tag{H.4}$$

G. Ludyk, *Quantum Mechanics in Matrix Form*,
https://doi.org/10.1007/978-3-319-26366-3

This implies that $\boldsymbol{R}^{-1}$ commutes with $\boldsymbol{L}_k$. For example, from $\boldsymbol{L}_k\boldsymbol{R} = \boldsymbol{R}\boldsymbol{L}_k$ it follows that

$$\boldsymbol{R}^{-1}\boldsymbol{L}_k\boldsymbol{R}\boldsymbol{R}^{-1} = \boldsymbol{R}^{-1}\boldsymbol{R}\boldsymbol{L}_k\boldsymbol{R}^{-1}$$

so

$$\boldsymbol{R}^{-1}\boldsymbol{L}_k = \boldsymbol{L}_k\boldsymbol{R}^{-1}$$

or

$$\underline{\underline{[\boldsymbol{L}_k, \boldsymbol{R}^{-1}] = \boldsymbol{0}.}} \tag{H.5}$$

It is

$$[\boldsymbol{X}_1, \boldsymbol{L}_1] = i\hbar\frac{\partial \boldsymbol{L_1}}{\partial \boldsymbol{P}_k} = i\hbar\frac{\partial(\boldsymbol{X}_2\boldsymbol{P}_3 - \boldsymbol{X}_3\boldsymbol{P}_2)}{\partial \boldsymbol{P}_1} = \boldsymbol{0}$$

and generally

$$\underline{\underline{[\boldsymbol{X}_k, \boldsymbol{L}_k] = \boldsymbol{0}.}} \tag{H.6}$$

Next we have

$$[\boldsymbol{X}_1, \boldsymbol{L}_2] = i\hbar\frac{\partial \boldsymbol{L_2}}{\partial \boldsymbol{P}_1} = i\hbar\frac{\partial(\boldsymbol{X}_3\boldsymbol{P}_1 - \boldsymbol{X}_1\boldsymbol{P}_3)}{\partial \boldsymbol{P}_1} = i\hbar\boldsymbol{X}_3,$$

and general[3]

$$\underline{\underline{[\boldsymbol{X}_i, \boldsymbol{L}_j] = i\hbar\varepsilon_{ijk}\boldsymbol{X}_k.}} \tag{H.7}$$

Similarly for the commutator with $\boldsymbol{P}$ we get

$$\underline{\underline{[\boldsymbol{P}_k, \boldsymbol{L}_k] = \boldsymbol{0}}} \tag{H.8}$$

and

$$\underline{\underline{[\boldsymbol{P}_i, \boldsymbol{L}_j] = i\hbar\varepsilon_{ijk}\boldsymbol{P}_k.}} \tag{H.9}$$

Notice these commutator properties [29]

$$\underline{\underline{[\boldsymbol{A}, \boldsymbol{B}\boldsymbol{C}] = \boldsymbol{B}[\boldsymbol{A}, \boldsymbol{C}] + [\boldsymbol{A}, \boldsymbol{B}]\boldsymbol{C}.}}$$

---

[3] The Levi-Civita symbol ε_{ijk} is defined as follows.

$$\varepsilon_{ijk} = \begin{cases} +1 & \text{if } (i, j, k) \text{ is } (1, 2, 3), (2, 3, 1), \text{ or } (3, 1, 2), \\ -1 & \text{if } (i, j, k) \text{ is } (3, 2, 1), (1, 3, 2), \text{ or } (2, 1, 3), \\ 0 & \text{if } i = j \text{ or } j = k \text{ or } k = i \end{cases};$$

that is, ε_{ijk} is 1 if (i, j, k) is an even permutation of $(1, 2, 3)$, -1 if it is an odd permutation, and 0 if any index is repeated.

Proof: $\boldsymbol{B}[\boldsymbol{A}, \boldsymbol{C}] + [\boldsymbol{A}, \boldsymbol{B}]\boldsymbol{C} = \boldsymbol{B}(\boldsymbol{AC} - \boldsymbol{CA}) + (\boldsymbol{AB} - \boldsymbol{BA})\boldsymbol{C} =$
$= \boldsymbol{ABC} - \boldsymbol{BCA} = [\boldsymbol{A}, \boldsymbol{BC}]$. q.e.d.
Similarly,

$$\underline{\underline{[\boldsymbol{AB}, \boldsymbol{C}] = \boldsymbol{A}[\boldsymbol{B}, \boldsymbol{C}] + [\boldsymbol{A}, \boldsymbol{C}]\boldsymbol{B}.}} \tag{H.10}$$

Suppose now that $[\boldsymbol{A}, \boldsymbol{B}] = \boldsymbol{0}$. If n is a positive integer, then so is

$$\underline{\underline{[\boldsymbol{A}^n, \boldsymbol{B}] = \boldsymbol{0}}}.$$

Proof By induction: The base case, $n = 1$, is given. Then using the inductive hypothesis, we get

$$[\boldsymbol{A}^n, \boldsymbol{B}] = [\boldsymbol{A}^{n-1}, \boldsymbol{B}]\boldsymbol{A} + \boldsymbol{A}^{n-1}[\boldsymbol{A}, \boldsymbol{B}] = \boldsymbol{0}. \quad \text{q.e.d.}$$

H.2 Proof of Pauli's Formula (7.14)

$$[\boldsymbol{A}_i, \boldsymbol{L}_i] = \boldsymbol{0} \quad (7.14)$$

First we prove

$$-\mathfrak{P} \times \mathfrak{L} = \mathfrak{L} \times \mathfrak{P} - 2i\hbar\mathfrak{P}. \tag{H.11}$$

We have with (H.9) for the first component, and similarly for the other two components,

$$(\mathfrak{L} \times \mathfrak{P})_1 - 2i\hbar\boldsymbol{P}_1 = \underbrace{\boldsymbol{L}_2\boldsymbol{P}_3}_{-\boldsymbol{P}_3\boldsymbol{L}_2 + i\hbar\boldsymbol{P}_1} - \underbrace{\boldsymbol{L}_3\boldsymbol{P}_2}_{-\boldsymbol{P}_2\boldsymbol{L}_3 - i\hbar\boldsymbol{P}_1} - 2i\hbar\boldsymbol{P}_1 = -(\mathfrak{P} \times \mathfrak{L})_1.$$

With (H.11) we can write (7.11) now

$$\mathfrak{A} = \frac{1}{Ze^2 m}(\mathfrak{L} \times \mathfrak{P} - i\hbar\mathfrak{P}) + \mathfrak{R}\boldsymbol{R}^{-1}. \tag{H.12}$$

With

$$\boldsymbol{A}_1 = \frac{1}{mZe^2}(\boldsymbol{L}_2\boldsymbol{P}_3 - \boldsymbol{L}_3\boldsymbol{P}_2 - i\hbar\boldsymbol{P}_1) + \boldsymbol{X}_1\boldsymbol{R}^{-1}$$

we have

$$[\boldsymbol{A}_1, \boldsymbol{L}_1] = \left[\frac{1}{mZe^2}(\boldsymbol{L}_2\boldsymbol{P}_3 - \boldsymbol{L}_3\boldsymbol{P}_2 - i\hbar\boldsymbol{P}_1), \boldsymbol{L}_1\right] + \left[\boldsymbol{X}_1\boldsymbol{R}^{-1}, \boldsymbol{L}_1\right],$$

and with (H.8)

$$[\boldsymbol{A}_1, \boldsymbol{L}_1] = \frac{1}{mZe^2}\left(\underbrace{\boldsymbol{L}_2\boldsymbol{P}_3\boldsymbol{L}_1 - \boldsymbol{L}_1\boldsymbol{L}_2\boldsymbol{P}_3}_{[\boldsymbol{L}_2\boldsymbol{P}_3, \boldsymbol{L}_1]}\underbrace{-\boldsymbol{L}_3\boldsymbol{P}_2\boldsymbol{L}_1 + \boldsymbol{L}_1\boldsymbol{L}_3\boldsymbol{P}_2}_{-[\boldsymbol{L}_3\boldsymbol{P}_2, \boldsymbol{L}_1]}\right) + [\boldsymbol{X}_1\boldsymbol{R}^{-1}, \boldsymbol{L}_1],$$

and with (H.13)

$$[\boldsymbol{A}_1, \boldsymbol{L}_1] = \frac{1}{mZe^2}\left(\underbrace{[\boldsymbol{L}_2\boldsymbol{P}_3, \boldsymbol{L}_1]}_{\boldsymbol{L}_2\underbrace{[\boldsymbol{P}_3, \boldsymbol{L}_1]}_{i\hbar\boldsymbol{P}_2} + \underbrace{[\boldsymbol{L}_2, \boldsymbol{L}_1]}_{-i\hbar\boldsymbol{L}_3}\boldsymbol{P}_3} - \underbrace{[\boldsymbol{L}_3\boldsymbol{P}_2, \boldsymbol{L}_1]}_{\boldsymbol{L}_3\underbrace{[\boldsymbol{P}_2, \boldsymbol{L}_1]}_{-i\hbar\boldsymbol{P}_3} + \underbrace{[\boldsymbol{L}_3, \boldsymbol{L}_1]}_{i\hbar\boldsymbol{L}_2}\boldsymbol{P}_2}\right) + [\cdots] =$$

$$= \boldsymbol{0} + [\boldsymbol{X}_1\boldsymbol{R}^{-1}, \boldsymbol{L}_1],$$

and with (H.13), (H.5), and (H.6) we finally get

$$[\boldsymbol{A}_1, \boldsymbol{L}_1] = \boldsymbol{X}_1\underbrace{[\boldsymbol{R}^{-1}, \boldsymbol{L}_1]}_{\boldsymbol{0}} + \underbrace{[\boldsymbol{X}_1, \boldsymbol{L}_1]}_{\boldsymbol{0}}\boldsymbol{R}^{-1} = \boldsymbol{0}.$$

Similarly we get for $k = 2$, and 3

$$[\boldsymbol{A}_k, \boldsymbol{L}_k] = \boldsymbol{0}.$$

H.3 Proof of Pauli's Formula (7.15)

(7.15) $\quad [\boldsymbol{A}_i, \boldsymbol{L}_j] = i\hbar\varepsilon_{ijk}\boldsymbol{A}_k.$

with

$$\boldsymbol{A}_1 = \frac{1}{mZe^2}(\boldsymbol{L}_2\boldsymbol{P}_3 - \boldsymbol{L}_3\boldsymbol{P}_2 - i\hbar\boldsymbol{P}_1) + \boldsymbol{X}_1\boldsymbol{R}^{-1}$$

we have

$$[\boldsymbol{A}_1, \boldsymbol{L}_2] = \left[\frac{1}{mZe^2}(\boldsymbol{L}_2\boldsymbol{P}_3 - \boldsymbol{L}_3\boldsymbol{P}_2 - i\hbar\boldsymbol{P}_1), \boldsymbol{L}_2\right] + \left[\boldsymbol{X}_1\boldsymbol{R}^{-1}, \boldsymbol{L}_2\right],$$

and with (H.8)

$$[\boldsymbol{A}_1, \boldsymbol{L}_2] = \frac{1}{mZe^2}\left(\underbrace{\boldsymbol{L}_2\boldsymbol{P}_3\boldsymbol{L}_2 - \boldsymbol{L}_2\boldsymbol{L}_2\boldsymbol{P}_3}_{[\boldsymbol{L}_2\boldsymbol{P}_3, \boldsymbol{L}_2]}\underbrace{-\boldsymbol{L}_3\boldsymbol{P}_2\boldsymbol{L}_2 + \boldsymbol{L}_2\boldsymbol{L}_3\boldsymbol{P}_2}_{-[\boldsymbol{L}_3\boldsymbol{P}_2, \boldsymbol{L}_2]}\underbrace{-i\hbar\boldsymbol{P}_1\boldsymbol{L}_2 + i\hbar\boldsymbol{L}_2\boldsymbol{P}_1}_{-i\hbar\underbrace{[\boldsymbol{P}_1, \boldsymbol{L}_2]}_{i\hbar\boldsymbol{P}_3}}\right) +$$

$$+[\boldsymbol{X}_1\boldsymbol{R}^{-1}, \boldsymbol{L}_2],$$

and with (H.13)

$$[\boldsymbol{A}_1, \boldsymbol{L}_2] = \frac{1}{mZe^2}\left(\underbrace{[\boldsymbol{L}_2\boldsymbol{P}_3, \boldsymbol{L}_2]}_{\boldsymbol{L}_2\underbrace{[\boldsymbol{P}_3, \boldsymbol{L}_2]}_{-i\hbar\boldsymbol{P}_1}+\underbrace{[\boldsymbol{L}_2, \boldsymbol{L}_2]}_{\boldsymbol{0}}\boldsymbol{P}_3} - \underbrace{[\boldsymbol{L}_3\boldsymbol{P}_2, \boldsymbol{L}_2]}_{\boldsymbol{L}_3\underbrace{[\boldsymbol{P}_2, \boldsymbol{L}_2]}_{\boldsymbol{0}}+\underbrace{[\boldsymbol{L}_3, \boldsymbol{L}_2]}_{-i\hbar\boldsymbol{L}_1}\boldsymbol{P}_2} -(i\hbar)^2\boldsymbol{P}_3\right) + [\cdots] =$$

$$= \frac{i\hbar}{mZe^2}\left(-\boldsymbol{L}_2\boldsymbol{P}_1 + \boldsymbol{L}_1\boldsymbol{P}_2 - i\hbar\boldsymbol{P}_3\right) + [\boldsymbol{X}_1\boldsymbol{R}^{-1}, \boldsymbol{L}_2].$$

With (H.13), (H.5), and (H.6) we finally get

$$\underline{\underline{[\boldsymbol{A}_1, \boldsymbol{L}_2]}} = \frac{i\hbar}{mZe^2}\left(\underbrace{-\boldsymbol{L}_2\boldsymbol{P}_1 + \boldsymbol{L}_1\boldsymbol{P}_2}_{(\mathfrak{L}\times\mathfrak{P})_3} - i\hbar\boldsymbol{P}_3\right) + \boldsymbol{X}_1\underbrace{[\boldsymbol{R}^{-1}, \boldsymbol{L}_2]}_{\boldsymbol{0}} + \underbrace{[\boldsymbol{X}_1, \boldsymbol{L}_2]}_{i\hbar\boldsymbol{X}_3}\boldsymbol{R}^{-1} = \underline{\underline{i\hbar\boldsymbol{A}_3}}.$$

Similarly we get

$$\underline{\underline{[\boldsymbol{A}_i, \boldsymbol{L}_j] = i\hbar\varepsilon_{ijk}\boldsymbol{A}_k.}}$$

H.4 Proof of Pauli's Formula (7.16)

(7.16) $\mathfrak{A}\cdot\mathfrak{L} = \mathfrak{L}\cdot\mathfrak{A} = \boldsymbol{0}.$

We have

$$\mathfrak{A}\cdot\mathfrak{L} = \left\{\frac{1}{Ze^2m}\left((\mathfrak{L}\times\mathfrak{P}) - i\hbar\,\mathfrak{P}\right) + (\mathfrak{R}\boldsymbol{R}^{-1})\right\}\cdot\mathfrak{L} \tag{H.13}$$

The angular momentum vector matrix $\mathfrak{L}$ is orthogonal to each of the three terms in (H.13). To see this, we first look for

$$(\mathfrak{L}\times\mathfrak{P})\cdot\mathfrak{L} = \begin{pmatrix} \boldsymbol{L}_2\boldsymbol{P}_3 - \boldsymbol{L}_3\boldsymbol{P}_2 \\ \boldsymbol{L}_3\boldsymbol{P}_1 - \boldsymbol{L}_1\boldsymbol{P}_3 \\ \boldsymbol{L}_1\boldsymbol{P}_2 - \boldsymbol{L}_2\boldsymbol{P}_1 \end{pmatrix}\cdot\begin{pmatrix} \boldsymbol{L}_1 \\ \boldsymbol{L}_2 \\ \boldsymbol{L}_3 \end{pmatrix} = \tag{H.14}$$

$$= \underbrace{\underbrace{\boldsymbol{L}_2\boldsymbol{P}_3}_{\boldsymbol{P}_3\boldsymbol{L}_2 - i\hbar\boldsymbol{P}_1}\boldsymbol{L}_1 - \underbrace{\boldsymbol{L}_1\boldsymbol{P}_3}_{\boldsymbol{P}_3\boldsymbol{L}_1 - i\hbar\boldsymbol{P}_2}\boldsymbol{L}_2}_{-i\hbar\boldsymbol{P}_1\boldsymbol{L}_1 + i\hbar\boldsymbol{P}_2\boldsymbol{L}_2 - i\hbar\boldsymbol{P}_3\boldsymbol{L}_3} + \underbrace{\boldsymbol{L}_3\boldsymbol{P}_1\boldsymbol{L}_2 - \boldsymbol{L}_2\boldsymbol{P}_1\boldsymbol{L}_3}_{-i\hbar\boldsymbol{P}_1\boldsymbol{L}_1 + i\hbar\boldsymbol{P}_3\boldsymbol{L}_3 - i\hbar\boldsymbol{P}_2\boldsymbol{L}_2} +$$

$$+ \underbrace{\boldsymbol{L}_1\boldsymbol{P}_2\boldsymbol{L}_3 - \boldsymbol{L}_3\boldsymbol{P}_2\boldsymbol{L}_1}_{-i\hbar\boldsymbol{P}_2\boldsymbol{L}_2 + i\hbar\boldsymbol{P}_1\boldsymbol{L}_1 + i\hbar\boldsymbol{P}_3\boldsymbol{L}_3} = -i\hbar\boldsymbol{P}_1\boldsymbol{L}_1 - i\hbar\boldsymbol{P}_2\boldsymbol{L}_2 - i\hbar\boldsymbol{P}_3\boldsymbol{L}_3 = -i\hbar\,\mathfrak{P}\cdot\mathfrak{L}.$$

This is the same as the second term in (H.13). But for this term we get

$$\mathfrak{P} \cdot \mathfrak{L} = \mathfrak{P} \cdot (\mathfrak{R} \times \mathfrak{P}) = \begin{pmatrix} \boldsymbol{P}_1 \\ \boldsymbol{P}_2 \\ \boldsymbol{P}_3 \end{pmatrix} \cdot \begin{pmatrix} \boldsymbol{X}_2\boldsymbol{P}_3 - \boldsymbol{X}_3\boldsymbol{P}_2 \\ \boldsymbol{X}_3\boldsymbol{P}_1 - \boldsymbol{X}_1\boldsymbol{P}_3 \\ \boldsymbol{X}_1\boldsymbol{P}_2 - \boldsymbol{X}_2\boldsymbol{P}_1 \end{pmatrix},$$

and because $\boldsymbol{X}_i$ commutes with $\boldsymbol{P}_k$ for $i \neq k$, we finally get

$$\mathfrak{P} \cdot \mathfrak{L} = \mathbf{0}. \tag{H.15}$$

The same is true for the third term in (H.13), inasmuch as $\boldsymbol{X}_i$ commutes with $\boldsymbol{X}_k$,

$$\mathfrak{R} \cdot \mathfrak{L} = \mathfrak{R} \cdot (\mathfrak{R} \times \mathfrak{P}) = \mathbf{0}. \tag{H.16}$$

H.5 Proof of Pauli's Formula (7.17)

(7.17) $\mathfrak{A} \times \mathfrak{A} = -i\hbar\frac{2}{mZ^2e^4}\mathfrak{L}\boldsymbol{E}.$

With

$$\mathfrak{A} = \frac{1}{Ze^2m}\frac{1}{2}\left(\mathfrak{L} \times \mathfrak{P} - i\hbar\mathfrak{P}\right) + \mathfrak{R}\boldsymbol{R}^{-1},$$

$$\mathfrak{B} \overset{\text{def}}{=} \frac{1}{2}(\mathfrak{L} \times \mathfrak{P} - i\hbar\mathfrak{P},$$

and

$$\mathfrak{A} = \frac{1}{Ze^2m}\mathfrak{B} + \mathfrak{R}\boldsymbol{R}^{-1},$$

we have

$$\mathfrak{A} \times \mathfrak{A} = \frac{1}{Z^2e^4m^2}(\mathfrak{B} \times \mathfrak{B}) + \frac{1}{Ze^2m}\left\{(\mathfrak{R}\boldsymbol{R}^{-1}) \times (\mathfrak{L} \times \mathfrak{P}) + (\mathfrak{L} \times \mathfrak{P}) \times (\mathfrak{R}\boldsymbol{R}^{-1})\right\}. \tag{H.17}$$

First, we obtain for $(\mathfrak{B} \times \mathfrak{B})$

$$\mathfrak{B} \times \mathfrak{B} = ((\mathfrak{L} \times \mathfrak{P}) \times (\mathfrak{L} \times \mathfrak{P})) - i\hbar((\mathfrak{L} \times \mathfrak{P}) \times \mathfrak{P}) - i\hbar(\mathfrak{P} \times (\mathfrak{L} \times \mathfrak{P})). \tag{H.18}$$

and for the first component

$$((\mathfrak{L} \times \mathfrak{P}) \times (\mathfrak{L} \times \mathfrak{P}))_1 = (\mathfrak{L} \times \mathfrak{P})_2(\mathfrak{L} \times \mathfrak{P})_3 - (\mathfrak{L} \times \mathfrak{P})_3(\mathfrak{L} \times \mathfrak{P})_2 =$$

$$= [(\mathfrak{L} \times \mathfrak{P})_2, (\mathfrak{L} \times \mathfrak{P})_3] = [\boldsymbol{L}_3\boldsymbol{P}_1 - \boldsymbol{L}_1\boldsymbol{P}_3, \boldsymbol{L}_1\boldsymbol{P}_2 - \boldsymbol{L}_2\boldsymbol{P}_1] =$$

$$= (\boldsymbol{L}_3\boldsymbol{P}_1 - \boldsymbol{L}_1\boldsymbol{P}_3)(\boldsymbol{L}_1\boldsymbol{P}_2 - \boldsymbol{L}_2\boldsymbol{P}_1) - (\boldsymbol{L}_1\boldsymbol{P}_2 - \boldsymbol{L}_2\boldsymbol{P}_1)(\boldsymbol{L}_3\boldsymbol{P}_1 - \boldsymbol{L}_1\boldsymbol{P}_3).$$

With (7.13), (H.8), and (H.9) we obtain

$$((\mathfrak{L}\times\mathfrak{P})\times(\mathfrak{L}\times\mathfrak{P}))_1 = -i\hbar\boldsymbol{L}_1\boldsymbol{P}_1^2 - i\hbar\boldsymbol{L}_1\boldsymbol{P}_2^2 - i\hbar\boldsymbol{L}_1\boldsymbol{P}_3^2 = -i\hbar\boldsymbol{L}_1\mathfrak{P}^2, \quad \text{(H.19)}$$

and generally

$$((\mathfrak{L}\times\mathfrak{P})\times(\mathfrak{L}\times\mathfrak{P})) = -i\hbar\,\mathfrak{L}\,\mathfrak{P}^2. \quad \text{(H.20)}$$

Next, we obtain for the first component of the second and third terms in (H.18)

$$((\mathfrak{L}\times\mathfrak{P})\times\mathfrak{P})_1 + (\mathfrak{P}\times(\mathfrak{L}\times\mathfrak{P}))_1 =$$

$$= \{(\mathfrak{L}\times\mathfrak{P})_2\boldsymbol{P}_3 - (\mathfrak{L}\times\mathfrak{P})_3\boldsymbol{P}_2\} + \{\boldsymbol{P}_2(\mathfrak{L}\times\mathfrak{P})_3 - \boldsymbol{P}_3(\mathfrak{L}\times\mathfrak{P})_2\} =$$

$$= [\underbrace{(\mathfrak{L}\times\mathfrak{P})_2}_{\boldsymbol{L}_3\boldsymbol{P}_1-\boldsymbol{L}_1\boldsymbol{P}_3}, \boldsymbol{P}_3] - [\underbrace{(\mathfrak{L}\times\mathfrak{P})_3}_{\boldsymbol{L}_1\boldsymbol{P}_2-\boldsymbol{L}_2\boldsymbol{P}_1}, \boldsymbol{P}_2] = \boldsymbol{0}.$$

Finally, we need to examine the last term in (H.17)

$$(\mathfrak{X}\boldsymbol{R}^{-1})\times(\mathfrak{L}\times\mathfrak{P}) + (\mathfrak{L}\times\mathfrak{P})\times(\mathfrak{X}\boldsymbol{R}^{-1}) \overset{\text{def}}{=} \mathfrak{C}.$$

With

$$\mathfrak{B} \overset{\text{def}}{=} \mathfrak{L}\times\mathfrak{P}$$

we have for the first component $\boldsymbol{C}_1$ of $\mathfrak{C}$

$$\boldsymbol{C}_1 = ((\mathfrak{X}\boldsymbol{R}^{-1})\times\mathfrak{B})_1 + (\mathfrak{B}\times(\mathfrak{X}\boldsymbol{R}^{-1}))_1 =$$

$$= (\boldsymbol{X}_2\boldsymbol{R}^{-1}\boldsymbol{B}_3 - \boldsymbol{X}_3\boldsymbol{R}^{-1}\boldsymbol{B}_2) + (\boldsymbol{B}_2\boldsymbol{X}_3\boldsymbol{R}^{-1} - \boldsymbol{B}_3\boldsymbol{X}_2\boldsymbol{R}^{-1}) =$$

$$= [\boldsymbol{X}_2\boldsymbol{R}^{-1}, \boldsymbol{B}_3] - [\boldsymbol{X}_3\boldsymbol{R}^{-1}, \boldsymbol{B}_2],$$

and with (H.13) and (3.27)

$$\boldsymbol{C}_1 = \boldsymbol{X}_2[\boldsymbol{R}^{-1}, \boldsymbol{B}_3] + \underbrace{[\boldsymbol{X}_2, \boldsymbol{B}_3]}_{i\hbar\frac{\partial\boldsymbol{B}_3}{\partial\boldsymbol{P}_2}}\boldsymbol{R}^{-1} - \boldsymbol{X}_3[\boldsymbol{R}^{-1}, \boldsymbol{B}_2] - \underbrace{[\boldsymbol{X}_3, \boldsymbol{B}_2]}_{i\hbar\frac{\partial\boldsymbol{B}_2}{\partial\boldsymbol{P}_3}}\boldsymbol{R}^{-1} =$$

$$= i\hbar\frac{\partial\boldsymbol{B}_3}{\partial\boldsymbol{P}_2}\boldsymbol{R}^{-1} - i\hbar\frac{\partial\boldsymbol{B}_2}{\partial\boldsymbol{P}_3}\boldsymbol{R}^{-1} + \boldsymbol{X}_2[\boldsymbol{R}^{-1}, \boldsymbol{L}_1\boldsymbol{P}_2 - \boldsymbol{L}_2\boldsymbol{P}_1] - \boldsymbol{X}_3[\boldsymbol{R}^{-1}, \boldsymbol{L}_1\boldsymbol{P}_3 - \boldsymbol{L}_3\boldsymbol{P}_1]$$

$$-i\hbar\frac{\partial\boldsymbol{R}^{-1}}{\partial\boldsymbol{X}_3}\boldsymbol{X}_2 + i\hbar\frac{\partial\boldsymbol{R}^{-1}}{\partial\boldsymbol{X}_2}\boldsymbol{X}_3,$$

or

$$\boldsymbol{C}_1 = i\hbar\frac{\partial\boldsymbol{B}_3}{\partial\boldsymbol{P}_2}\boldsymbol{R}^{-1} - i\hbar\frac{\partial\boldsymbol{B}_2}{\partial\boldsymbol{P}_3}\boldsymbol{R}^{-1}$$

$$+i\hbar \boldsymbol{L}_1\left(\frac{\partial \boldsymbol{R}^{-1}}{\partial \boldsymbol{X}_2}\boldsymbol{X}_2+\frac{\partial \boldsymbol{R}^{-1}}{\partial \boldsymbol{X}_3}\boldsymbol{X}_3\right)-i\hbar(\boldsymbol{L}_2\boldsymbol{X}_2+\boldsymbol{L}_3\boldsymbol{X}_3)\frac{\partial \boldsymbol{R}^{-1}}{\partial \boldsymbol{X}_1}.$$

With $\mathfrak{L}\mathfrak{R}=\boldsymbol{L}_1\boldsymbol{X}_1+\boldsymbol{L}_2\boldsymbol{X}_2+\boldsymbol{L}_3\boldsymbol{X}_3=\boldsymbol{0}$ and $\dfrac{\partial \boldsymbol{R}^{-1}}{\partial \boldsymbol{X}_k}=-\boldsymbol{X}_k\boldsymbol{R}^{-3}$ we obtain for the last line

$$i\hbar \boldsymbol{L}_1\left(\boldsymbol{X}_1^2+\boldsymbol{X}_2^2+\boldsymbol{X}_3^2\right)\boldsymbol{R}^{-3}=\imath\hbar \boldsymbol{L}_1\boldsymbol{R}^{-1}.$$

Furthermore, one has

$$\frac{\partial \boldsymbol{B}_3}{\partial \boldsymbol{P}_2}-\frac{\partial \boldsymbol{B}_2}{\partial \boldsymbol{P}_3}=\frac{\partial}{\partial \boldsymbol{P}_2}(\boldsymbol{L}_1\boldsymbol{P}_2)+\frac{\partial}{\partial \boldsymbol{P}_3}(\boldsymbol{L}_1\boldsymbol{P}_3)=$$

$$=2\boldsymbol{L}_1+(\boldsymbol{P}_2\frac{\partial}{\partial \boldsymbol{P}_2}+\boldsymbol{P}_3\frac{\partial}{\partial \boldsymbol{P}_3})\boldsymbol{L}_1=3\boldsymbol{L}_1,$$

thus in total

$$\boldsymbol{C}_1=2i\hbar \boldsymbol{L}_1.$$

Putting this in (H.17), we have

$$\mathfrak{A}\times\mathfrak{A}=-i\hbar\frac{2}{mZ^2e^4}\mathfrak{L}\left(\frac{1}{2m}\mathfrak{P}^2-e\boldsymbol{R}^{-1}\right).$$

H.6 Proof of Pauli's Formula (7.18)

(7.18) $\mathfrak{A}^2=\frac{2}{mZ^2e^4}\boldsymbol{E}\left(\mathfrak{L}^2+\frac{h^2}{4\pi^2}\boldsymbol{I}\right)+\boldsymbol{I}.$

With

$$\mathfrak{A}=\frac{1}{Ze^2m}\frac{1}{2}\left(\mathfrak{L}\times\mathfrak{P}-i\hbar\mathfrak{P}\right)+\mathfrak{R}\boldsymbol{R}^{-1},$$

$$\mathfrak{B}\overset{\text{def}}{=}\frac{1}{Ze^2m}\frac{1}{2}(\mathfrak{L}\times\mathfrak{P}-i\hbar\mathfrak{P}),$$

and

$$\mathfrak{A}=\mathfrak{B}+\mathfrak{R}\boldsymbol{R}^{-1},$$

we have

$$\mathfrak{A}\cdot\mathfrak{A}=\left\{\mathfrak{B}^2+\mathfrak{R}\boldsymbol{R}^{-1}\mathfrak{B}+\mathfrak{B}\mathfrak{R}\boldsymbol{R}^{-1}+\boldsymbol{I}\right\} \tag{H.21}$$

For the first term on the right-hand side of (H.21) we get

$$\mathfrak{B}^2 = \frac{1}{Z^2e^4m^2}\frac{1}{4}(\mathfrak{L}\times\mathfrak{P} - i\hbar\mathfrak{P})^2 =$$

$$= \frac{1}{Z^2e^4m^2}\frac{1}{4}\left((\mathfrak{L}\times\mathfrak{P})^2 - i\hbar\mathfrak{P}(\mathfrak{L}\times\mathfrak{P}) - i\hbar(\mathfrak{L}\times\mathfrak{P})\mathfrak{P} - (\hbar)^2\mathfrak{P}^2\right). \tag{H.22}$$

As with formula (H.14), these formulas can be proven.

$$(\mathfrak{L}\times\mathfrak{P})^2 = \mathfrak{P}^2\mathfrak{L}^2, \tag{H.23}$$

$$\mathfrak{P}(\mathfrak{L}\times\mathfrak{P}) = 2i\hbar\mathfrak{P}^2, \tag{H.24}$$

and

$$(\mathfrak{L}\times\mathfrak{P})\mathfrak{P} = \mathbf{0}. \tag{H.25}$$

With these three formulas, we have

$$\mathfrak{B}^2 = \frac{1}{Z^2e^4m^2}\frac{1}{4}\left(\mathfrak{P}^2\mathfrak{L}^2 + 2\hbar^2\mathfrak{P}^2 - (\hbar)^2\mathfrak{P}^2\right) = \frac{1}{Z^2e^4m^2}\frac{1}{4}\mathfrak{P}^2(\mathfrak{L}^2 + (\hbar)^2\boldsymbol{I}). \tag{H.26}$$

Next we get

$$(\mathfrak{R}\boldsymbol{R}^{-1})\cdot(\mathfrak{L}\times\mathfrak{P}) = \begin{pmatrix} \boldsymbol{X}_1\boldsymbol{R}^{-1} \\ \boldsymbol{X}_2\boldsymbol{R}^{-1} \\ \boldsymbol{X}_3\boldsymbol{R}^{-1} \end{pmatrix}\cdot\begin{pmatrix} \boldsymbol{L}_2\boldsymbol{P}_3 - \boldsymbol{L}_3\boldsymbol{P}_2 \\ \boldsymbol{L}_3\boldsymbol{P}_1 - \boldsymbol{L}_1\boldsymbol{P}_3 \\ \boldsymbol{L}_1\boldsymbol{P}_2 - \boldsymbol{L}_2\boldsymbol{P}_1 \end{pmatrix} = 2i\hbar(\mathfrak{R}\cdot\mathfrak{P})\boldsymbol{R}^{-1} + \boldsymbol{L}^2\boldsymbol{R}^{-1},$$

and

$$(\mathfrak{L}\times\mathfrak{P})\cdot(\mathfrak{R}\boldsymbol{R}^{-1}) = \boldsymbol{L}^2\boldsymbol{R}^{-1}.$$

Thus we have

$$\mathfrak{R}\boldsymbol{R}^{-1}\mathfrak{B} = \frac{1}{Ze^2m}\frac{1}{2}(\boldsymbol{L}^2\boldsymbol{R}^{-1} + 2i\hbar\mathfrak{R}\cdot\mathfrak{P})\boldsymbol{R}^{-1}),$$

and

$$\mathfrak{B}\mathfrak{R}\boldsymbol{R}^{-1} = \frac{1}{Ze^2m}\frac{1}{2}(\boldsymbol{L}^2\boldsymbol{R}^{-1} - 2i\hbar\mathfrak{R}\cdot\mathfrak{P})\boldsymbol{R}^{-1}).$$

Putting these in (H.21) we obtain the final result

$$\mathfrak{A}\cdot\mathfrak{A} = \frac{2}{mZ^2e^4}\left(\frac{1}{2m}\boldsymbol{P}^2 - Ze^2\boldsymbol{R}^{-1}\right)(\boldsymbol{L}^2 + \hbar^2\boldsymbol{I}) + \boldsymbol{I} = \frac{2}{mZ^2e^4}\boldsymbol{E}(\boldsymbol{L}^2 + \hbar^2\boldsymbol{I}) + \boldsymbol{I}.$$

Appendix I
Physical Quantities and Units

- Bohr Magneton: $\mu_B = \frac{e\hbar}{2m_e} = 9.27400968 \cdot 10^{-24}\mathrm{Am}^2$ (or JT^{-1});
- Fine Structure Constant: $\alpha = \frac{e^2}{4\pi\epsilon_0 c\hbar} = \frac{e^2 c\mu_0}{2\hbar} = 7.29735357 \cdot 10^{-3} \approx \frac{1}{137}$;
- Bohr Radius: $a_0 = \frac{4\pi\epsilon_0\hbar^2}{m_e e^2} = \frac{\hbar}{\alpha m_e c} = 5.2917721092 \cdot 10^{-11}\,\mathrm{m} \approx 52.9\,\mathrm{pm}$;

Force: $1\mathrm{N}(Newton) = 1\mathrm{m\,kg/s^2}$;
Work, Energy: $1\mathrm{J}(Joule) = 1\mathrm{N\,m}$;
Magnetic Flux Density: $1\mathrm{T}(Tesla) = 1\mathrm{N/(Am)}$;
Power: $1\mathrm{W}(Watt) = 1\mathrm{J/s}$;
Voltage: $1\mathrm{V}(Volt) = 1\mathrm{J/C}$;
Charge Quantity: $1\mathrm{C}(Coulomb) = 1\mathrm{A\,s}$;
Current: $1\mathrm{A}(Ampere) = 1\mathrm{C/s}$.

G. Ludyk, *Quantum Mechanics in Matrix Form*,
https://doi.org/10.1007/978-3-319-26366-3

References

1. G. Aruldhas, *Quantum Mechanics* (PHI Learning, New Delhi, 2012)
2. P. Atkins, R. Friedman, *Molecular Quantum Mechanics* (Oxford University Press, Oxford, 2005)
3. L.E. Ballentine, *Quantum Mechanics. A Modern Development* (World Scientific, Singapore, 1998)
4. M. Born, P. Jordan, Zur Quantenmechanik. ZS f. Physik (1925). The English translation of the paper is included as paper 13 in [27]
5. M. Born, P. Jordan, *Elementare Quantenmechanik* (Springer, German, 1930)
6. M. Born, W. Heisenberg, P. Jordan, Zur Quantenmechanik, (II). ZS f. Physik (1925). The English translation of the paper is included as paper 15 in [27]
7. J.W. Brewer, Kronecker products and matrix calculus in system theory. IEEE Trans. Circuits Syst. pp. 772–781 (1978)
8. P.A.M. Dirac, The fundamental equations of quantum mechanics. Proc. R. Soc. London, Ser. A (1925)
9. P.A.M. Dirac, *The Principles of Quantum Mechanics* (Oxford, Oxford, 1967)
10. D. Dubbers, H.-J. Stöckmann, *Quantum Physics, the Bottom-Up Approach* (Springer, Berlin, 2013)
11. S. Gasiorowicz, *Quantum Physics* (Wiley, New York, 2003)
12. H.S. Green, *Matrix Methods in Quantum Mechanics* (Barnes & Noble, New York, 1965)
13. W. Heisenberg, Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen. ZS f. Physik (1925). The English translation of the paper is included as paper 12 in [27]
14. P. Jordan, *Anschauliche Quantentheorie* (Springer, German, 1936)
15. G. Ludyk, *Einstein in Matrix Form* (Springer, Berlin, 2013)
16. E.L. Michelsen, *Quircy Quantum Concepts* (Springer, Berlin, 2014)
17. J. Pade, *Quantum Mechanics for Pedestrians (1 and 2)* (Springer, Berlin, 2013)
18. W. Pauli, Über das Wasserstoffspektrum vom Standpunkt der neuen Quantenmechanik. ZS f. Physik (1926). The English translation of the paper is included as paper 16 in [27]
19. M. Razavy, *Heisenberg's Quantum Mechanics* (World Scientific, Singapore, 2011)
20. F. Scheck, *Quantum Physics* (Springer, Berlin, 2013)
21. F. Schwabl, *Quantum Mechanics* (Springer, Berlin, 2007)
22. A. Sommerfeld, *Atomic Structure and Spectral Lines* (Methuen, London, 1923)
23. W.-H. Steeb, Y. Hardy, *Quantum Mechanics Using Computer Algebra* (World Scientific, Singapore, 2010)

G. Ludyk, *Quantum Mechanics in Matrix Form*,
https://doi.org/10.1007/978-3-319-26366-3

24. Th.F. Jordan, *Quantummechanics in Simple Matrix Form* (Dover, 1986)
25. S. Weinberg, *Lectures on Quantum Mechanics* (Cambridge, Cambridge, 2013)
26. S. Weinberg, *Dreams of a Final Theory* (Pantheon Books, New York, 1992)
27. B.L. van der Waerden, *Sources of Quantum Mechanics* (Dover, Amsterdam, 1967)
28. P.A. Schilpp, *Albert Einstein: Philosopher - Scientist* (Kohlhammer, New York, 1949)
29. G.H. Shortley, G.E. Kimball, Analysis of non-commuting vectors with application to quantum mechanics and vector calculus. Proc. N. A. S. (1934)

Index

G. Ludyk, *Quantum Mechanics in Matrix Form*,
https://doi.org/10.1007/978-3-319-26366-3