Graduate Texts in Physics

Dirk Dubbers
Hans-Jürgen Stöckmann

Quantum Physics: The Bottom-Up Approach

From the Simple Two-Level System to Irreducible Representations

Springer

Graduate Texts in Physics

For further volumes:
www.springer.com/series/8431

Graduate Texts in Physics

Graduate Texts in Physics publishes core learning/teaching material for graduate- and advanced-level undergraduate courses on topics of current and emerging fields within physics, both pure and applied. These textbooks serve students at the MS- or PhD-level and their instructors as comprehensive sources of principles, definitions, derivations, experiments and applications (as relevant) for their mastery and teaching, respectively. International in scope and relevance, the textbooks correspond to course syllabi sufficiently to serve as required reading. Their didactic style, comprehensiveness and coverage of fundamental material also make them suitable as introductions or references for scientists entering, or requiring timely knowledge of, a research field.

Dirk Dubbers • Hans-Jürgen Stöckmann

Quantum Physics: The Bottom-Up Approach

From the Simple Two-Level System to Irreducible Representations

Springer

Dirk Dubbers
Fak. Physik und Astronomie
Physikalisches Institut
Universität Heidelberg
Heidelberg, Germany

Hans-Jürgen Stöckmann
Fachbereich Physik
Philipps-Universität Marburg
Marburg, Germany

ISSN 1868-4513
ISSN 1868-4521 (electronic)
Graduate Texts in Physics
ISBN 978-3-642-31059-1
ISBN 978-3-642-31060-7 (eBook)
DOI 10.1007/978-3-642-31060-7
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2012950772

Preface

Quantum mechanics pervades many branches of science, from physics, material science, informatics, to chemistry and molecular biology. Many products of everyday life derive from discoveries based on quantum physics, like silicon chips, magnetic storage devices, lasers, medical imaging devices, as well as chemicals and biochemicals. Therefore, scientists and engineers in many fields need a good understanding of quantum theory, but often they are overwhelmed by the sheer volume of most standard textbooks on quantum physics.

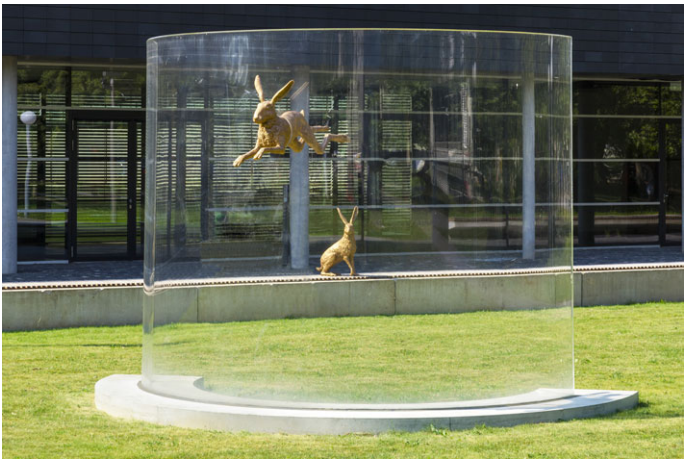
Our approach is to first limit discussion to the smallest systems in nature that still display the basic features of quantum theory. Hence this tutorial at first deals with systems in which only two quantum states are involved, subject to external perturbations. Such effective spin one-half systems are a valuable training ground to elucidate the subtleties of quantum theory and, indeed, the essence of quantum mechanics lies in the two-level system. We present basic quantum calculations step by step in a simple notation and in sufficient detail because the practitioner usually has no time to lose when proceeding from one equation to the next.

As a starting point we assume the reader to have taken introductory courses in quantum mechanics and linear algebra, and to be familiar with the Schrödinger equation and the essentials of angular momentum, which we recapitulate in Part I of the book. Part II covers essential topics of quantum physics based on the two-state approach, including subjects of high contemporary interest as quantum entanglement, quantum chaos, or geometric phases. In Part III, the results then are applied to various topics from atomic, condensed matter, and nuclear physics, and from quantum informatics.

We then proceed to the more general concepts of quantum theory. To this end, Part IV of this treatise restarts from first principles to develop the theory of angular momentum, spherical tensors, and irreducible representations. We derive a generalized spin precession equation that covers the higher multipole interactions, and apply the results to various topics in atomic and condensed matter physics. Chapters on multiple quantum transitions, dressed atom effects, spin relaxation and decoherence conclude the tutorial.

The text is based on various lectures given by the authors on the two-state system, irreducible tensors, and quantum chaos, and is complemented with an illustrative set of basic experiments, many of them done in the authors' respective laboratories. In short, the aim of this tutorial is to provide the bachelor student as well as the practitioner with a compact text that lets them understand a wealth of quantum physics.

Heidelberg, Germany — Dirk Dubbers
Marburg, Germany — Hans-Jürgen Stöckmann



“Dürer meets Einstein, travelling.” Installation by Sabrina Hohmann, for an interpretation see http://www.physi.uni-heidelberg.de/KunstamBau.pdf. With friendly permission of the artist. Foto: Wolf-Dieter Gericke

This page intentionally left blank

Contents

This page intentionally left blank

Part I
Prologue

The authors assume that the reader of this treatise has taken an introductory course on quantum mechanics. This Prologue summarizes the essentials of such a course.

This page intentionally left blank

Chapter 1
Recollections from Elementary Quantum Physics

Abstract We recall the prerequisites that we assume the reader to be familiar with, namely the Schrödinger equation in its time dependent and time independent form, the uncertainty relations, and the basic properties of angular momentum.

Introductory courses on quantum physics discuss the one-dimensional Schrödinger equation for the *wave function* $\Psi(x,t)$ of a particle of mass M moving in a potential V

$$\mathrm{i}\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2M}\frac{\partial^2\Psi}{\partial x^2} + V\Psi. \tag{1.1}$$

Therein $\hbar = h/2\pi$ is the reduced Planck constant. The function Ψ is understood as a probability amplitude whose absolute square $|\Psi(x,t)|^2 = \Psi^*(x,t)\Psi(x,t)$ gives the *probability density* for finding the particle at time t at position x. This probability density is insensitive to a *phase factor* $\mathrm{e}^{\mathrm{i}\varphi}$. With the *Hamilton operator*

$$\mathrm{H} = -\frac{\hbar^2}{2M}\frac{\partial^2}{\partial x^2} + V, \tag{1.2}$$

the Schrödinger equation reads

$$\dot{\Psi} = -\frac{\mathrm{i}}{\hbar}\mathrm{H}\Psi. \tag{1.3}$$

The dot denotes the time derivative.

In this text, we print operators and matrices in nonitalic type, like H, $\mathbf{p}$, or $\boldsymbol{\sigma}$, just to remind the reader that a simple letter may represent a mathematical object more complicated than a number or a function. Ordinary vectors in three-dimensional space are written in bold italic type, like $\boldsymbol{x}$ or $\boldsymbol{B}$.

The time dependent Schrödinger equation reminds us of the law of energy conservation $E = p^2/2M + V$ if we associate the operator $\mathrm{i}\hbar\partial/\partial t$ with energy E and the operator $(\hbar/\mathrm{i})\partial/\partial x$ with momentum p. Take as an example the plane wave $\Psi(x,t) = \Psi_0\mathrm{e}^{\mathrm{i}(kx-\omega t)}$ of a photon propagating in the x-direction. The energy operation $\mathrm{i}\hbar\partial\Psi/\partial t = \hbar\omega\Psi$ then relates energy to frequency, $E = \hbar\omega$, and the momentum operation $(\hbar/\mathrm{i})\partial\Psi/\partial x = \hbar k\Psi$ relates momentum to wave number, $p = \hbar k$.

D. Dubbers, H.-J. Stöckmann, *Quantum Physics: The Bottom-Up Approach*, Graduate Texts in Physics,
DOI 10.1007/978-3-642-31060-7_1,

The probability amplitude $\Psi(x,t)$ for finding the particle at time t at position x and the amplitude $\Phi(p,E)$ for finding it with energy E and momentum p turn out to be *Fourier transforms* of each other. Pairs of Fourier transforms have widths that are reciprocal to each other. If the width Δx in position is large, the width Δp in momentum is small, and vice versa, and the same for the widths Δt and ΔE. The *conjugate observables* p and x or E and t obey the *uncertainty relations*

$$\Delta p \Delta x \geq \frac{1}{2}\hbar, \qquad \Delta E \Delta t \geq \frac{1}{2}\hbar. \tag{1.4}$$

The exact meaning of Δx, etc. will be defined in Sect. 3.4.

For *stationary*, that is, time independent potentials $V(x)$, the Hamilton operator (1.2) acts only on the position variable x. The solution of the Schrödinger equation then is *separable* in x and t,

$$\Psi(x,t) = \psi(x)\mathrm{e}^{-\mathrm{i}Et/\hbar}. \tag{1.5}$$

The probability density (in units of m^{-1}) for finding a particle at position x then is independent of time $|\Psi(x,t)|^2 = |\psi(x)|^2$. The amplitude $\psi(x)$ is a solution of the *time independent* Schrödinger equation

$$-\frac{\hbar^2}{2M}\frac{\partial^2\psi(x)}{\partial x^2} + V(x)\psi(x) = E\psi(x). \tag{1.6}$$

In operator notation, this reads

$$\mathrm{H}\psi = E\psi. \tag{1.7}$$

For particles trapped in a potential well $V(x)$, the requirement that total probability is *normalizable* to

$$\int_{-\infty}^{+\infty} \left|\psi(x)\right|^2 \mathrm{d}x = 1 \tag{1.8}$$

leads to the *quantization of energy* E such that only a discrete set of values E_n, with the corresponding wave functions $\psi_n(x)$, $n = 1, 2, 3, \ldots$, is allowed. n is called the *main quantum number*.

The mean value derived from repeated measurements of a physical quantity or *observable* A is given by its *expectation value*

$$\left\langle A(t)\right\rangle = \int_{-\infty}^{+\infty} \Psi^*(x,t) A \Psi(x,t)\,\mathrm{d}x. \tag{1.9}$$

For a stationary state Eq. (1.5), the expectation value $\langle A\rangle$ does not depend on time and is insensitive to any phase factor $\mathrm{e}^{\mathrm{i}\varphi}$. For example, the mean position $\langle x\rangle$ of a matter wave is given by the weighted average

$$\langle x\rangle = \int_{-\infty}^{+\infty} x\left|\psi(x)\right|^2 \mathrm{d}x. \tag{1.10}$$

Sometimes the Hamiltonian H can be divided into two parts $\mathrm{H} = \mathrm{H}_1(\boldsymbol{x}) + \mathrm{H}_2(\boldsymbol{y})$, one depending on one set of (not necessarily spatial) variables $\boldsymbol{x}$, the other on a disjoint set of variables $\boldsymbol{y}$. Then, as shown in standard textbooks, the solution of the Schrödinger equation is separable in $\boldsymbol{x}$ and $\boldsymbol{y}$ as $\psi(\boldsymbol{x}, \boldsymbol{y}) = \psi_1(\boldsymbol{x})\psi_2(\boldsymbol{y})$. As in Eq. (1.5), we must attach a phase factor to obtain the time dependent solution

$$\Psi(\boldsymbol{x}, \boldsymbol{y}, t) = \psi_1(x)\psi_2(\boldsymbol{y})\mathrm{e}^{-\mathrm{i}Et/\hbar}. \tag{1.11}$$

The probability densities for the joint occurrence of variable $\boldsymbol{x}$ and variable $\boldsymbol{y}$ then multiply to

$$\left|\Psi(\boldsymbol{x}, \boldsymbol{y}, t)\right|^2 = \left|\psi_1(\boldsymbol{x})\right|^2 \left|\psi_2(\boldsymbol{y})\right|^2. \tag{1.12}$$

This is consistent with classical probability theory, where probabilities multiply for independent events. The probability that two dice both show 5 is $\frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$.

If, on the other hand, the quantum system can evolve along two mutually exclusive paths ψ_1 or ψ_2, as is the case in the famous double-slit experiments, then the two probability amplitudes add *coherently*, and the total probability is

$$|\Psi|^2 = |\psi_1 + \psi_2|^2 = |\psi_1|^2 + |\psi_2|^2 + 2|\psi_1||\psi_2| \sin\delta, \tag{1.13}$$

with phase angle δ. Only if phase coherence between the partial waves ψ_1 and ψ_2 is lost, does the total probability equal the classical result

$$|\Psi|^2 = |\psi_1|^2 + |\psi_2|^2, \tag{1.14}$$

with probabilities adding for mutually exclusive events. The probability that a die shows either 5 or 3 is $\frac{1}{6} + \frac{1}{6} = \frac{1}{3}$.

Figure 1.1 shows the setup and result of a double-slit experiment with a beam of slow neutrons. What we see is the self-interference of the probability amplitudes ψ_1 and ψ_2 of single neutrons. The conditions for the appearance of quantum interference will be discussed in Sect. 6.6.

For most students the Schrödinger equation is no more difficult than the many other differential equations encountered in mechanics, electrodynamics, or transport theory. Therefore, students usually have fewer problems with the ordinary Schrödinger equation than with a presentation of quantum mechanics in the form of matrix mechanics. In everyday scientific life, however, spectroscopic problems that require matrix diagonalization are more frequent than problems like particle waves encountering step potentials and other scattering problems. Furthermore, as we shall see in Sect. 19.1.2, the Schrödinger equation can always equally well be expressed as a matrix equation. Therefore, matrix mechanics is the main topic of the present tutorial. Instead of starting with the Schrödinger equation, we could have started with the Heisenberg equation of motion (discovered a few weeks before the Schrödinger equation) because both are equivalent. We postpone the introduction of Heisenberg's equation to Chap. 10 because students are less familiar with this approach.

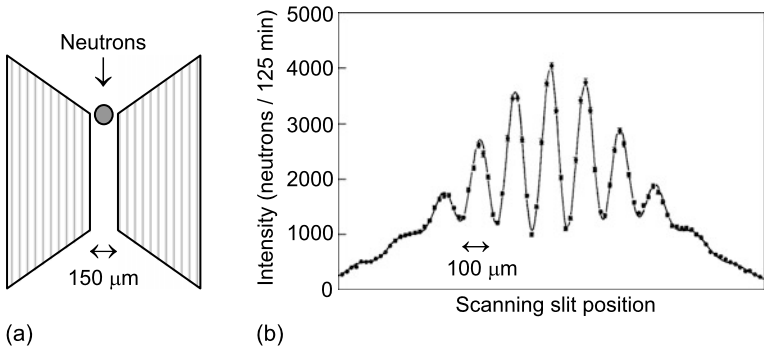


Fig. 1.1 (**a**) Double-slit experiment with neutrons: The monochromatic beam (*along the arrow*), of de Broglie wave length $\lambda = 2$ nm, $\Delta\lambda/\lambda = 8$ %, meets two slits, formed by a beryllium wire (*shaded circle*) and two neutron absorbing glass edges (*vertical hatching*), installed on an optical bench of 10 m length with 20 μm wide entrance and exit slits. (**b**) The measured neutron self-interference pattern follows theoretical expectation. From Zeilinger et al. (1988)

To begin we further assume acquaintance with some basic statements of quantum physics that are both revolutionary and demanding, and our hope is that the reader of this text will learn how to live and work with them. The first statement is: To any physical observable A corresponds an operator A. The measurement of the observable A leaves the system under study in one of the *eigenstates* or *eigenfunctions* ψ_n of this operator such that

$$\mathrm{A}\psi_n = a_n\psi_n. \tag{1.15}$$

The possible outcomes of the measurement are limited to the *eigenvalues* a_n. Which of these eigenfunctions and eigenvalues is singled out by the measurement is uncertain until the measurement is actually done.

If we regard the ordered list ψ_n as elements of a vector space, called the Hilbert space, the operation A just stretches an eigenvector ψ_n of this space by a factor a_n. The best-known example of an eigenvalue equation of this type is the time independent Schrödinger equation

$$\mathrm{H}\psi_n = E_n\psi_n. \tag{1.16}$$

H is the operator for the observable energy E, ψ_n are the eigenfunctions of energy, and E_n are the corresponding eigenvalues. In this text, we shall only treat the simplest case where the spectrum of the E_n is *discrete* and enumerable, as are atomic energy spectra, and in most cases limit the number of energy levels to two.

Another basic statement is derived in Sect. 3.4: Two physical quantities, described by operators A and B, can assume well defined values a_n and b_n, simultaneously measurable with arbitrarily high precision and unhampered by any uncertainty relation, if and only if they share a common set of eigenfunctions ψ_n, with

$$\mathrm{A}\psi_n = a_n\psi_n, \qquad \mathrm{B}\psi_n = b_n\psi_n. \tag{1.17}$$

An important operator is that of the *angular momentum* $\mathbf{J}$, which we shall derive from first principles in Chap. 16. Let us beforehand recapitulate the following properties of $\mathbf{J}$, as known from introductory quantum physics. The components J_x, J_y, and J_z of the angular momentum operator $\mathbf{J}$ cannot be measured simultaneously to arbitrary precision. Only the square magnitude $\mathbf{J}^2$ of the angular momentum and its component J_z along an arbitrary axis z are well defined and can be measured simultaneously without uncertainty. This means that the phase of $\mathbf{J}$ about this axis z remains uncertain. The operators of the two observables $\mathbf{J}^2$ and J_z then must have simultaneous eigenfunctions that we call ψ_{jm}, which obey

$$\mathbf{J}^2\psi_{jm} = j(j+1)\hbar^2\psi_{jm}, \tag{1.18a}$$

$$\mathrm{J}_z\psi_{jm} = m\hbar\psi_{jm}, \tag{1.18b}$$

with eigenvalues $j(j+1)\hbar^2$ and $m\hbar$, respectively.

Angular momentum is quantized, with possible *angular momentum quantum numbers* $j = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$. For a given value of j, the *magnetic quantum number* m can take on only the $2j+1$ different values $m = -j, -j+1, \ldots, j$.

For $j = \frac{1}{2}$ we have $m = \pm\frac{1}{2}$ and

$$\mathbf{J}^2\psi_{\frac{1}{2},\pm\frac{1}{2}} = \frac{3}{4}\hbar^2\psi_{\frac{1}{2},\pm\frac{1}{2}}, \tag{1.19a}$$

$$\mathrm{J}_z\psi_{\frac{1}{2},\pm\frac{1}{2}} = \pm\frac{1}{2}\hbar\psi_{\frac{1}{2},\pm\frac{1}{2}}. \tag{1.19b}$$

If we arrange the $2j+1$ eigenfunctions of the angular momentum ψ_{jm} into one column

$$\psi_j = \begin{pmatrix} \psi_{j,j} \\ \psi_{j,j-1} \\ \vdots \\ \psi_{j,-j} \end{pmatrix}, \tag{1.20}$$

we shall print this column vector ψ_j in nonitalic type, as we did for the operators. The corresponding row vector is $\psi_j^\dagger = (\psi_{j,j}^*, \psi_{j,j-1}^*, \ldots, \psi_{j,-j}^*)$, where the dagger signifies the *conjugate transpose* $\psi^\dagger = \psi^{*\mathrm{T}}$ of a complex vector (or matrix).

The components J_x and J_y of the angular momentum operator $\mathbf{J}$ do not have simultaneous eigenfunctions with $\mathbf{J}^2$ and J_z. If we form the linear combinations

$$\mathrm{J}_+ = \mathrm{J}_x + \mathrm{i}\mathrm{J}_y, \qquad \mathrm{J}_- = \mathrm{J}_x - \mathrm{i}\mathrm{J}_y, \tag{1.21}$$

these operators are found to act on the angular momentum state ψ_{jm} as

$$\mathrm{J}_+\psi_{jm} = \hbar\sqrt{j(j+1) - m(m+1)}\,\psi_{j,m+1}, \tag{1.22a}$$

$$\mathrm{J}_-\psi_{jm} = \hbar\sqrt{j(j+1) - m(m-1)}\,\psi_{j,m-1}, \tag{1.22b}$$

with $J_+\psi_{j,+j} = 0$ and $J_-\psi_{j,-j} = 0$. The J_+ and J_- are the *raising* and *lowering operators*, respectively. For $j = \frac{1}{2}$ we have $J_+\psi_{\frac{1}{2},-\frac{1}{2}} = \hbar\psi_{\frac{1}{2},+\frac{1}{2}}$ and $J_-\psi_{\frac{1}{2},+\frac{1}{2}} = \hbar\psi_{\frac{1}{2},-\frac{1}{2}}$.

Angular momentum may be composed of the *orbital* angular momentum $\mathbf{L}$ and of *spin* angular momentum $\mathbf{S}$, which can be added to the *total* angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{S}$. The orbital angular momentum quantum number l can only have integer values, spin quantum number s can have either integer or half-integer values. The *triangle rule* of vector addition tells us that the (integer or half-integer) angular momentum quantum number j has the allowed range

$$|l - s| \leq j \leq l + s. \tag{1.23}$$

The total angular momentum quantum number is

$$m = m_l + m_s, \tag{1.24}$$

going in unit steps from $m = -j$ to $m = j$. For example, two angular momenta with $l = 1$ and $s = \frac{1}{2}$ can be added to $j = \frac{1}{2}$ or $j = \frac{3}{2}$. For $l = 0$, $j = s$ we have $m = m_s$, and for $s = 0$, $j = l$ we have $m = m_l$. For the sake of simplicity, in these two special cases we shall always write m for the respective magnetic quantum numbers m_s or m_l.

References

Zeilinger, A., Gähler, R., Shull, C.G., Treimer, W., Mampe, W.: Single- and double-slit diffraction of neutrons. Rev. Mod. Phys. **60**, 1067–1073 (1988)

Part II
Two-State Quantum Systems

Now we get started and derive the peculiar features of quantum physics for the special case of a system having two distinct quantum states. Such systems are found to behave always as if they had an effective spin of $\frac{1}{2}$. We first study the general evolution of spin-$\frac{1}{2}$ systems, and learn about quantum operators in all varieties. Then we discuss several quantum features that play a dominant role in contemporary quantum physics. To conclude this part, we retrace the simple arguments on which Heisenberg had based the concepts of matrix mechanics.

This page intentionally left blank

Chapter 2
A Most Simple Two-Level System

Abstract A first example of a two-state quantum system is an ensemble of independent spin-$\frac{1}{2}$ particles, electrons, for instance, interacting with an external magnetic field via their magnetic dipole moment. We discuss the potential energy of the magnetic dipole interaction and recall the basic experiments that discovered the quantization of energy and angular momentum.

2.1 Magnetic Moment and Spin

The spin and the magnetic moment of the electron and their quantization are natural consequences of the relativistic Dirac equation. This equation is beyond the scope of this text, but in the nonrelativistic limit, the magnetic dipole interaction enters the Schrödinger equation via the classical *magnetic potential*

$$V_{\mathrm{M}} = -\boldsymbol{\mu} \cdot \boldsymbol{B}, \tag{2.1}$$

or $V_{\mathrm{M}}(\theta) = -\mu B \cos\theta$.

In *classical* physics, the angle θ between the magnetic moment $\boldsymbol{\mu}$ and the magnetic field $\boldsymbol{B}$ can assume any value $0 \le \theta \le \pi$, see Fig. 2.1. The magnetic moment of a current I that forms a loop of area $\boldsymbol{a}$ is $\boldsymbol{\mu} = I\boldsymbol{a}$. The magnetic moment $\boldsymbol{\mu}$ of a classical (spinless) electron, orbiting a nucleus with velocity $\boldsymbol{v}$ at relative position $\boldsymbol{r}$, is related to its orbital angular momentum $\boldsymbol{L} = m_e \boldsymbol{r} \times \boldsymbol{v}$ as

$$\boldsymbol{\mu} = -\frac{e}{2m_e}\boldsymbol{L}, \tag{2.2}$$

with the positive elementary charge e, and the mass m_e of the electron.

The classical magnetic force $\boldsymbol{F} = -\nabla V_{\mathrm{M}}$ acting on $\boldsymbol{\mu}$ is

$$\boldsymbol{F} = \nabla(\boldsymbol{\mu} \cdot \boldsymbol{B}), \tag{2.3}$$

and the torque $\boldsymbol{T}$ on $\boldsymbol{\mu}$ is

$$\boldsymbol{T} = \boldsymbol{\mu} \times \boldsymbol{B}, \tag{2.4}$$

with $T = \mu B \sin\theta = \partial V_{\mathrm{M}}/\partial\theta$. Torque $\boldsymbol{T}$ disappears at the positions of stable and unstable equilibrium $\theta = 0$ and $\theta = \pm\pi$.

D. Dubbers, H.-J. Stöckmann, *Quantum Physics: The Bottom-Up Approach*, Graduate Texts in Physics,
DOI 10.1007/978-3-642-31060-7_2, © Springer-Verlag Berlin Heidelberg 2013

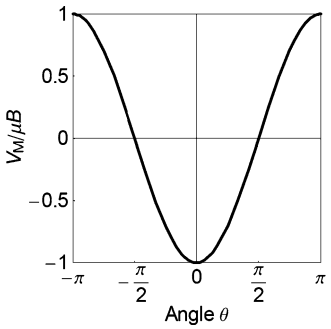


Fig. 2.1 Magnetic potential $V_{\mathrm{M}}(\theta)$ seen by the magnetic moment of a compass in a magnetic field. Classical magnetic energy can take on any value between $-\mu B$ and $+\mu B$, with stable and unstable equilibrium positions at $\theta = 0$ and $\theta = \pm\pi$, while for a quantum spin-$\frac{1}{2}$ system only two discrete energy values are allowed

In *quantum* physics, in analogy to (2.2), we write the *magnetic moment operator* of an electron with angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{S}$ as $\boldsymbol{\mu} = g(e/2m_e)\mathbf{J}$, or

$$\boldsymbol{\mu} = \frac{g\mu_{\mathrm{B}}}{\hbar}\mathbf{J}. \tag{2.5}$$

The magnetic moment is given in units of the *Bohr magneton*

$$\mu_{\mathrm{B}} = \frac{e\hbar}{2m_e} \tag{2.6}$$

of magnitude $\mu_{\mathrm{B}} = 5.79 \times 10^{-5}$ eV/T.

As angular momenta are quantized, so are, via (2.5), magnetic moments. The g-factor may be looked upon as a "fudge factor" which takes care of possible deviations from this crude model. Indeed, the Dirac equation predicts $g_l = -1$ for the orbital angular momentum $\mathbf{L}$ of an electron, but $g_s = -2$ for its spin angular momentum $\mathbf{S}$. For total angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{S}$ the g-factor, called the *Landé factor* g_j, is found to take on values between -1 and -2. We follow the official CODATA sign convention for the g-factor (Mohr et al. 2008).

The tabulated values of magnetic moments with angular momentum quantum number j give the quantities

$$\mu = g\mu_{\mathrm{B}} j, \tag{2.7}$$

which is the eigenvalue of the operator $\mu_z = g\mu_{\mathrm{B}}\mathrm{J}_z/\hbar$ for the maximum magnetic quantum number $m = j$. Equation (2.5) then reads

$$\boldsymbol{\mu} = \mu\frac{\mathbf{J}}{\hbar j}. \tag{2.8}$$

The *magnetic energy operator* $\mathrm{H_M}$ is obtained from the classical magnetic potential $V_\mathrm{M} = -\boldsymbol{\mu} \cdot \boldsymbol{B}$ by replacing the magnetic moment $\boldsymbol{\mu}$ with the magnetic moment operator $\boldsymbol{\mu}$:

$$\mathrm{H_M} = -\boldsymbol{\mu} \cdot \boldsymbol{B} = -\frac{\mu}{\hbar j} \mathbf{J} \cdot \boldsymbol{B}. \tag{2.9}$$

For the free electron with $s = \frac{1}{2}$, the g-factor is $g \approx -2$ within one percent (the tiny deviation from $g = -2$ will be discussed in Sect. 4.1). Its magnetic moment then is

$$\mu = \frac{1}{2} g \mu_\mathrm{B} \approx -\mu_\mathrm{B} \quad \text{(free electron)}. \tag{2.10}$$

Another frequently used quantity is the *gyromagnetic ratio*

$$\gamma = -\frac{g\mu_\mathrm{B}}{\hbar}. \tag{2.11}$$

The gyromagnetic ratio $\gamma = -g\mu/\hbar$ links, via (2.8), the magnetic moment of a gyrating particle, endowed with some internal charge distribution, to its angular momentum, as $\boldsymbol{\mu} = -\gamma \mathbf{J}$. For the free electron, $\gamma/2\pi = 28.0$ GHz/T > 0. With three quantities g, μ, and γ at hand, notation is somewhat redundant. (CODATA define all $\gamma = |g|\mu_\mathrm{B}/\hbar$ as positive, which, however, complicates established NMR notation.)

For daily work, a physicist should know by heart a set of natural constants. The following six constants, precise to about 1 percent or better, are easy to memorize and sufficient for essentially all calculations in microscopic physics:
(i) Velocity of light $c = 3 \times 10^8$ m/s; (ii) Electron mass $m_e = 0.511$ MeV/c^2; (iii) Nucleon mass $m_\mathrm{N} = 1836 m_e$; (iv) Fine structure constant $\alpha = e^2/(4\pi\varepsilon_0\hbar c) = 1/137$; (v) Boltzmann's constant $k_\mathrm{B} = 26$ meV/300 K; (vi) $\hbar c = 200$ eV nm (= 200 MeV fm for nuclear physicists).
The eV is officially accepted for use with the SI system. As an example, we calculate the Bohr magneton, (2.6), to $\mu_\mathrm{B} = c(\hbar c/2m_e c^2)e \approx 5.9 \times 10^{-5}$ eV/T, where we used 1 T = 1 Vs/m^2 and absorbed e in the eV.
Notice that in small print we bring, for the interested reader, information that can be skipped without loosing the main line of argumentation.

The results of this section were mainly established by two historic experiments. The *Zeeman effect* showed quantization of magnetic energy, while the *Stern–Gerlach effect* proved the directional quantization of angular momentum.

2.2 Zeeman Effect

The magnetic-field induced splitting of atomic spectral lines was first observed by Zeeman towards the end of the nineteenth century, when atomic structure was still unknown. In classical electrodynamics, for a charge that oscillates in the presence

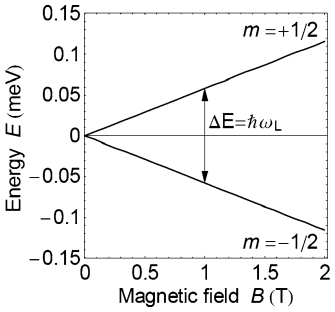


Fig. 2.2 Zeeman effect of the free electron. The energy degeneracy at zero field is lifted by application of a magnetic field $B \neq 0$. The observed Zeeman splitting is linear in B. The electron's gyromagnetic ratio is $\gamma = \omega_\mathrm{L}/B = 2\pi \times 28$ GHz/T

of a magnetic field, one expects a three-fold splitting of the frequency of the emitted dipole radiation, called the *normal* Zeeman effect. In most experiments, however, atoms show an *anomalous* Zeeman effect, which was explained only years later after electron spin had been discovered and its role in spin-orbit coupling was understood.

Quantum mechanically, with $\mathbf{J}$ and $\boldsymbol{\mu}$ quantized, the magnetic energy $\mathrm{H_M} = -\boldsymbol{\mu} \cdot \boldsymbol{B}$ from (2.9) is also quantized, which leads to a splitting of atomic energy levels when the atoms are exposed to a field $\boldsymbol{B}$, and with it to a splitting of the spectrum of the emitted photons, see also Fig. 4.2. If we choose the z-axis to lie along the magnetic field of amplitude $B = B_z$, the magnetic energy operator is $\mathrm{H_M} = -(g\mu_\mathrm{B}/\hbar)\mathrm{J}_z B_z$. With the eigenvalues $m\hbar$ of J_z from (1.18b), the $2j+1$ energy eigenvalues of the atom become

$$E_m = E_0 - mg\mu_\mathrm{B} B, \quad m = -j, \ldots, j. \tag{2.12}$$

E_0 is the energy of the unperturbed level. For zero external field, the atoms have no preferred axis, surrounding space is rotationally symmetric, and the $2j+1$ magnetic substates of the unperturbed atoms, despite their different spatial orientations, all have the same energy E_0. One says that the energy levels are *degenerate*. With the magnetic field turned on, the $(2j+1)$-fold energy splitting is linear in B, Fig. 2.2, and is equidistant, i.e., the separation of neighboring energy levels is independent of m,

$$\Delta E = E_m - E_{m-1} = -g\mu_\mathrm{B} B. \tag{2.13}$$

One shortly writes

$$\Delta E = \hbar\omega_\mathrm{L}, \tag{2.14}$$

with the *Larmor angular frequency* ω_L, written vectorially as

$$\boldsymbol{\omega}_\mathrm{L} = \gamma \boldsymbol{B}. \tag{2.15}$$

The magnetic Hamiltonian then becomes $\mathrm{H_M} = \mathbf{J} \cdot \boldsymbol{\omega}_\mathrm{L}$. For the electron, $\omega_\mathrm{L} > 0$, and in a magnetic field energy levels are split in two as

$$E_\pm = E_0 \pm \frac{1}{2}\hbar\omega_\mathrm{L}. \tag{2.16}$$

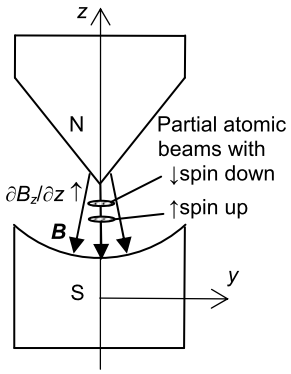


Fig. 2.3 In the Stern–Gerlach apparatus, silver atoms from a heated piece of silver in a steel box evaporate through a small hole into a vacuum tube, where 1/20 mm wide orifices shape a very narrow beam. The beam passes a magnet of 3 cm length, here shown in *cross-section*. The magnetic field gradient of 2 mT/cm pulls the spin up and spin down atoms into opposite directions. The two partial beams formed in this way then are condensed onto a small glass plate

Zeeman observed the magnetic-field induced splitting of the prominent yellow D_1, D_2 doublet in the spectrum of Na. The origin of this doublet will be discussed in Sect. 18.3, see also Fig. 18.1. The optical transmission spectrum of a sodium vapor in an optical cell was analyzed in an optical grating of high resolution. In a field of magnitude $B = 2$ T (the saturation field of the magnet's iron core), the Zeeman splitting of the Na ground state is $\hbar\omega_L \approx 2\mu_B B = 2.4 \times 10^{-4}$ eV, much less than the energy of 2.1 eV of the optical transition.

2.3 Stern–Gerlach Effect

A quarter century after Zeeman's discovery, though still before the advent of a veritable quantum theory, Stern and Gerlach observed what is called the *directional* or *space quantization* of angular momentum. In a Stern–Gerlach apparatus, a beam of atoms is exposed to a transverse magnetic field with a strong gradient across the beam axis. While the atoms fly through the magnetic field region, classically a constant transverse magnetic force $\boldsymbol{F} = \boldsymbol{\nabla}(\boldsymbol{\mu} \cdot \boldsymbol{B})$, (2.3), acts on their magnetic moment. Hence, instead of flying on straight lines, the atoms follow parabolic trajectories within the magnetic field region.

If the field $\boldsymbol{B}$ is locally approximated by a radial field, confined to the y–z plane as shown in Fig. 2.3, and if the atomic beam is limited to the field region where $B \approx B_z$, then the classical magnetic dipole force reduces to

$$F_z = \mu \frac{\partial B_z}{\partial z} \cos\theta. \tag{2.17}$$

The atoms in the beam are not noticeably magnetized because their thermal energy is much higher than their magnetic energy. Therefore, the classical expectation is that their magnetic moments point into all directions, their orientation is *isotropic*.

The forces on the atoms, (2.17), then are isotropic as well, which should lead to a smooth distribution of deflection angles θ. Hence the beam profile should simply be broadened when the field is turned on.

Instead, rather unexpected at the time, Stern and Gerlach observed that the atomic beam leaving the field region is split into two well-defined partial beams, schematically shown in Fig. 2.3. Evidently, instead of having the directions of the magnetic moment equally distributed over all angles, the quantity $\mu_z = \mu\cos\theta$ in (2.17) can assume only two quantized values. Today we know that with the magnetic moment $\boldsymbol{\mu}$ also the force $\boldsymbol{F}$ on the atoms in the beam is quantized. The silver atoms used have quantum numbers $l = 0$ and $j = s = 1/2$, and have the g-factor $g = g_s \approx -2$ of the free electron, hence $\mu_z = g\mu_\mathrm{B} J_z/\hbar$ from (2.5) has the two quantized eigenvalues $mg\mu_\mathrm{B} \approx \mp\mu_\mathrm{B}$, and

$$F_{z\pm} \approx \mp\mu_\mathrm{B}\frac{\partial B_z}{\partial z}. \tag{2.18}$$

This leads to the observed splitting into two partial beams. One partial beam is polarized *spin up*, the other *spin down*. That this two-fold splitting is due to an intrinsic spin of the electron was realized only several years after Stern and Gerlach's discovery.

The Stern–Gerlach effect can be seen with neutral atoms or with thermal neutrons, but not with free electrons, because the Lorentz force

$$\mathbf{F} = q\boldsymbol{v} \times \boldsymbol{B} \tag{2.19}$$

on a particle of charge q at thermal velocity $\boldsymbol{v}$ is much stronger than the force (2.18) due to the magnetic dipole interaction.

References

Mohr, P.J., Taylor, B.N., Newell, D.B.: CODATA recommended values of the fundamental physical constants: 2006. Rev. Mod. Phys. **80**, 633 (2008)

Chapter 3
Quantum Theory in a Nutshell

Abstract In principle, the two-state quantum system can be discussed in an elegant and short manner by employing various subtle operator techniques. We do not follow this route. Instead, we go one by one through all essential steps. We first establish the spin matrices for spin one half. It turns out that the most general 2×2 Hamiltonian matrix can be expressed in terms of the Pauli spin matrices. Therefore, every two-state quantum system, whatever the underlying interaction, can be treated as an $s = \frac{1}{2}$ effective-spin system. We derive the expectation values of energy- and spin-operators and their uncertainties, and we solve the two-state Schrödinger equation for two simple cases.

3.1 Spin Matrices

As the first example, we take the valence electron of a neutral alkaline atom in its ground state. Let H_0 be the (unknown) Hamiltonian operator of the free atom in zero magnetic field. When the magnetic field is turned on, then $\mathrm{H} = \mathrm{H}_0 + \mathrm{H}_\mathrm{M}$, with the magnetic operator $\mathrm{H}_\mathrm{M} = -(\mu/\hbar j)\mathbf{J} \cdot \boldsymbol{B} = \gamma \mathbf{J} \cdot \boldsymbol{B}$ from (2.9), and the Schrödinger equation (1.3) becomes

$$\dot{\Psi} = -\frac{\mathrm{i}}{\hbar}(\mathrm{H}_0 + \mathrm{H}_\mathrm{M})\Psi. \tag{3.1}$$

If H_0 is not affected by the magnetic field (we neglect the small diamagnetism of the atom), then H_0 and H_M are independent of each other, and separation of variables is possible, as pointed out in Chap. 1. H_0 then only contributes an overall phase factor $\exp(-\mathrm{i}E_0 t/\hbar)$ with ground state energy E_0. We here limit discussion to the case $\mathbf{L} = 0$, that is, $\mathbf{J} = \mathbf{S}$, with angular momentum quantum numbers $j = s$ and $m = m_s$. The spin dependent part of the Schrödinger equation $\dot{\psi}_s = -(\mathrm{i}/\hbar)\mathrm{H}_\mathrm{M}\psi_s$ then reads

$$\dot{\psi}_s = \frac{\mathrm{i}}{\hbar}\frac{\mu}{\hbar s}\mathbf{S} \cdot \boldsymbol{B}\psi_s. \tag{3.2}$$

Our task is to find the properties of the mathematical object $\mathbf{S} \cdot \boldsymbol{B}$.

D. Dubbers, H.-J. Stöckmann, *Quantum Physics: The Bottom-Up Approach*, Graduate Texts in Physics,
DOI 10.1007/978-3-642-31060-7_3, © Springer-Verlag Berlin Heidelberg 2013

We know from (1.20) that the solutions ψ_s of (3.2) are $(2s+1)$-tuples that are eigenfunctions of $\mathbf{S}^2$ and S_z. For $s=\frac{1}{2}$ we therefore choose the two-valued states

$$\alpha = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \beta = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{3.3}$$

as eigenvectors of H_M for the magnetic quantum numbers $m=+\frac{1}{2}$ and $m=-\frac{1}{2}$. These states form an *orthonormal basis* with

$$|\alpha|^2 = \alpha^\dagger \alpha = (1 \quad 0)\begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1, \tag{3.4a}$$

$$\alpha^\dagger \beta = (1 \quad 0)\begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0, \tag{3.4b}$$

and equally $|\beta|^2 = 1$ and $\beta^\dagger \alpha = 0$. In later chapters, we shall revert to the widespread notation $|+\frac{1}{2}\rangle$ and $|-\frac{1}{2}\rangle$ for the column vectors α and β, and $\langle +\frac{1}{2}|$ and $\langle -\frac{1}{2}|$ for their conjugate transpose, the row vectors $\alpha^\dagger$ and $\beta^\dagger$.

For $s=\frac{1}{2}$ we next establish the matrices for the spin operator $\mathbf{S}$ in (3.2). With respect to the basis states α and β, (1.19b) reads $S_z\alpha = \frac{1}{2}\hbar\alpha$ and $S_z\beta = -\frac{1}{2}\hbar\beta$, hence S_z is represented by the diagonal 2×2 matrix

$$S_z = \frac{1}{2}\hbar \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.5}$$

To find the other components S_x and S_y of $\mathbf{S}$, we use the spin raising and lowering operators $S_\pm = S_x \pm \mathrm{i}S_y$ from (1.22a), (1.22b) with $S_+\beta = \hbar\alpha$, $S_-\alpha = \hbar\beta$, $S_+\alpha = 0$, and $S_-\beta = 0$. This means that the corresponding matrices S_+ and S_- each have only one off-diagonal nonzero matrix element,

$$S_+ = \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad S_- = \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \tag{3.6}$$

hence

$$S_x = \frac{1}{2}(S_+ + S_-) = \frac{1}{2}\hbar \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{3.7a}$$

$$S_y = -\frac{1}{2}\mathrm{i}(S_+ - S_-) = \frac{1}{2}\hbar \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}. \tag{3.7b}$$

Shortening out the common factor $\frac{1}{2}\hbar$ in these equations, we arrive at

$$\mathbf{S} = \frac{1}{2}\hbar\boldsymbol{\sigma}, \tag{3.8a}$$

with the three *Pauli matrices*

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.8b}$$

For spin $s = \frac{1}{2}$ it is customary to express the spin operator $\mathbf{S}$ in terms of the dimensionless *Pauli spin operator* $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$.

When we apply the elements of the Pauli operator to α and β, we find

$$\sigma_x\alpha = \beta, \qquad \sigma_x\beta = \alpha, \tag{3.9a}$$

$$\sigma_y\alpha = \mathrm{i}\beta, \qquad \sigma_y\beta = -\mathrm{i}\alpha, \tag{3.9b}$$

$$\sigma_z\alpha = \alpha, \qquad \sigma_z\beta = -\beta. \tag{3.9c}$$

The square of each Pauli matrix is the 2×2 unit matrix I; therefore,

$$\boldsymbol{\sigma}^2 = \sigma_x^2 + \sigma_y^2 + \sigma_z^2 = 3\mathrm{I}. \tag{3.10}$$

For later use, we list the products of Pauli spin matrices

$$\sigma_x\sigma_y = \mathrm{i}\sigma_z, \qquad \sigma_y\sigma_z = \mathrm{i}\sigma_x, \qquad \sigma_z\sigma_x = \mathrm{i}\sigma_y, \tag{3.11}$$

with $\sigma_x\sigma_y = -\sigma_y\sigma_x$, etc. In terms of the Pauli spin operator, the magnetic moment operator of a spin-$\frac{1}{2}$ particle reads, from (2.5) and (2.7),

$$\boldsymbol{\mu} = \mu\boldsymbol{\sigma}. \tag{3.12}$$

For the free electron, $\mu \approx -\mu_\mathrm{B}$. The Schrödinger equation (3.2) then becomes

$$\dot{\psi}_s = \frac{\mathrm{i}}{\hbar}\mu\boldsymbol{\sigma} \cdot \boldsymbol{B}\psi_s. \tag{3.13}$$

3.2 Energy Matrices

With respect to the basis states α and β, any two-state operator A can be expressed in terms of its matrix elements

$$A_{11} = \alpha^\dagger \mathrm{A}\alpha, \qquad A_{12} = \alpha^\dagger \mathrm{A}\beta, \qquad A_{21} = \beta^\dagger \mathrm{A}\alpha, \qquad A_{22} = \beta^\dagger \mathrm{A}\beta, \tag{3.14}$$

which then are arranged in form of a 2×2 matrix. Let the general spin-$\frac{1}{2}$ Hamiltonian be represented by the matrix

$$\mathrm{H} = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}. \tag{3.15}$$

This Hamiltonian is *hermitian*, which means that it equals its conjugate transpose matrix, $\mathrm{H}^\dagger = \mathrm{H}$. Its diagonal elements H_{11} and H_{22} therefore are real, and its off-diagonal elements $H_{21} = H_{12}^*$ are the complex conjugates of each other. Hermiticity

of quantum operators assures that their eigenvalues, which are the possible outcomes of measurements, are real numbers.

With the spin matrices at hand, it is straightforward to derive the matrix elements of $\mathrm{H} = \mathrm{H}_0 + \mathrm{H}_\mathrm{M}$ used in (3.1). With spin independent $\mathrm{H}_0 = E_0\mathrm{I}$, and with the magnetic dipole interaction from (2.9) and (3.12)

$$\mathrm{H}_\mathrm{M} = -\mu\boldsymbol{\sigma}\cdot\boldsymbol{B}, \tag{3.16a}$$

$$\mathrm{H}_\mathrm{M} = -\mu(\sigma_x B_x + \sigma_y B_y + \sigma_z B_z), \tag{3.16b}$$

$$\mathrm{H}_\mathrm{M} = -\mu\begin{pmatrix} B_z & B_x - \mathrm{i}B_y \\ B_x + \mathrm{i}B_y & -B_z \end{pmatrix} = -\mu\begin{pmatrix} B_z & B_- \\ B_+ & -B_z \end{pmatrix}, \tag{3.16c}$$

with the complex field components $B_\pm = B_x \pm \mathrm{i}B_y$. Here and in the following, whenever suitable, we give the operator equation and the corresponding matrix equation one on top of the other. The reader then can easily see what is going on.

An arbitrary Hamiltonian H, (3.15), like any hermitian 2×2 matrix, can be expanded in terms of the Pauli matrices and the unit matrix I as

$$\mathrm{H} = E_0\mathrm{I} + \boldsymbol{\alpha}\cdot\boldsymbol{\sigma}, \tag{3.17a}$$

$$\mathrm{H} = E_0\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \alpha_x\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \alpha_y\begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} + \alpha_z\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{3.17b}$$

with coefficients

$$\begin{aligned} E_0 &= \frac{1}{2}(H_{11} + H_{22}), \qquad & \alpha_x &= \mathrm{Re}\, H_{12}, \\ \alpha_y &= -\,\mathrm{Im}\, H_{12}, & \alpha_z &= \frac{1}{2}(H_{11} - H_{22}). \end{aligned} \tag{3.18}$$

For a magnetic dipole interaction, (3.16a), the vector of coefficients is $\boldsymbol{\alpha} = -\mu\boldsymbol{B}$. The magnetic dipole Hamiltonian H_M therefore represents the most general two-state problem of quantum mechanics (Feynman et al. 1957).

Any quantum system consisting of two distinct states can be regarded as an $s = \frac{1}{2}$ *effective-spin system* acted upon by an *effective magnetic field*.

The eigenvalues of a quadratic matrix are derived from the zeros of the *characteristic polynomial*, given by the *secular equation*

$$\det(\mathrm{H} - E\mathrm{I}) = 0, \tag{3.19a}$$

$$\begin{aligned} \begin{vmatrix} H_{11} - E & H_{12} \\ H_{12}^* & H_{22} - E \end{vmatrix} &= (H_{11} - E)(H_{22} - E) - |H_{12}|^2 \\ &= E^2 - (H_{11} + H_{22})E + H_{11}H_{22} - |H_{12}|^2 = 0, \end{aligned} \tag{3.19b}$$

whose solutions are the two eigenenergies

$$E_{\pm} = \frac{1}{2}(H_{11} + H_{22}) \pm \frac{1}{2}\sqrt{(H_{11} - H_{22})^2 + 4|H_{12}|^2}. \tag{3.20}$$

In terms of the matrix elements of (3.16c), with $H_{11} - H_{22} = -2\mu B_z$ and $|H_{12}|^2 = \mu^2 B_+ B_- = \mu^2(B_x^2 + B_y^2)$, this becomes

$$E_{\pm} = E_0 \mp \mu\sqrt{B_x^2 + B_y^2 + B_z^2} = E_0 \mp \mu B. \tag{3.21}$$

In the following, without loss of generality, we shall set $E_0 = 0$, and $\mathrm{H} = \mathrm{H_M}$. We recall that for $s = \frac{1}{2}$ the following expressions are in use:

$$E_{\pm} = \mp\mu B = \mp\frac{1}{2} g\mu_{\mathrm{B}} B = \pm\frac{1}{2}\hbar\gamma B = \pm\frac{1}{2}\hbar\omega_{\mathrm{L}}. \tag{3.22}$$

These measurable eigenvalues $E_{\pm}$ are independent of the choice of the basis states of the matrix H, that is, of the *choice of representation* of H. In the energy representation, the Hamiltonian matrix is diagonal

$$\mathrm{H_D} = \begin{pmatrix} E_+ & 0 \\ 0 & E_- \end{pmatrix} = -\mu B_z \sigma_z = \frac{1}{2}\hbar\omega_{\mathrm{L}}\sigma_z, \tag{3.23}$$

which is the case when the *quantization axis* is chosen along z, and our task will be to transform the matrix in (3.16c) into diagonal form. We use the subscript D to signal the diagonal representation of matrices and their eigenvectors.

3.3 Expectation Values

Before we solve the Schrödinger equation (3.13) for spin $s = \frac{1}{2}$, let us stop for a moment and ask what to do with the solution $\psi(t)$ once we have found it (from now on, we drop the subscript s on ψ). Any solution of the Schrödinger equation is a linear superposition of basis states α and β

$$\psi(t) = a(t)\alpha + b(t)\beta = \begin{pmatrix} a(t) \\ b(t) \end{pmatrix}, \tag{3.24}$$

with appropriate, in general complex coefficients $a(t)$ and $b(t)$. Every such state vector ψ is normalized to

$$|\psi|^2 = \psi^\dagger \psi = |a|^2 + |b|^2 = 1. \tag{3.25}$$

The only meaningful quantities in quantum physics are the results of measurements done on a system with a given apparatus. In a spin measurement, a spin-$\frac{1}{2}$ particle will always be found in one of its two basis states α or β. The occurrence

of such an outcome or "event" is defined only in a probabilistic way: The wave function ψ, and with it a and b, are interpreted as probability amplitudes, with the probabilities

$$p_{\text{up}} = |a|^2, \qquad p_{\text{down}} = |b|^2 \tag{3.26}$$

for finding the system after measurement either in state α (spin up) or state β (spin down), and similarly for more than two states. Equation (3.25) says that these probabilities add up to a total probability

$$p_{\text{up}} + p_{\text{down}} = 1. \tag{3.27}$$

We want to know the solution ψ of the Schrödinger equation in order to know these probabilities and their dependence on time and other parameters.

Assume that an observable A is measured repeatedly under identical conditions, or measured simultaneously for many independent identical systems or *ensembles* of particles. The mean result of measurements is the expectation value

$$\langle \mathrm{A} \rangle = \psi^\dagger \mathrm{A} \psi, \quad \text{with } \psi^\dagger \psi = 1. \tag{3.28}$$

The only quantities relevant in quantum physics are expectation values.

If A is a scalar operator like the Hamiltonian H, then $\langle \mathrm{A} \rangle$ is a simple number, if A is a vector operator like the spin operator $\mathbf{S}$ in Hilbert space, then $\langle \mathrm{A} \rangle$ is a conventional vector in ordinary space, etc. An example of an expectation value was given in (1.10) for the case of a space-dependent wave function $\psi(x, t)$.

3.3.1 Expectation Value of Energy

Let the observable in (3.28) be energy. If the Hamiltonian operator is given in the energy representation with diagonal matrix $\mathrm{H_D}$ from (3.23), then the expectation value of energy is

$$\langle \mathrm{H} \rangle = \psi_{\mathrm{D}}^\dagger \mathrm{H_D} \psi_{\mathrm{D}}, \tag{3.29a}$$

$$\langle \mathrm{H} \rangle = (a^* \quad b^*) \begin{pmatrix} E_+ & 0 \\ 0 & E_- \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = |a|^2 E_+ + |b|^2 E_-. \tag{3.29b}$$

Indeed, this is the weighted average of energy. For H given in a different representation, with a different choice of basis states, one finds the same expectation value, as will be checked at the end of Sect. 5.1.

The only energy measurement that we have met so far is that of the Zeeman effect; therefore, let us find the expectation value of energy for this case. With the $\boldsymbol{B}$-field turned off, it is $\langle \mathrm{H} \rangle = E_0$, but unfortunately it is also $\langle \mathrm{H} \rangle = E_0$ with the

$\boldsymbol{B}$-field turned on because the Zeeman splitting leaves the mean energy of the m-states unaffected. Does this mean that the different energies measured in the Zeeman effect are not expectation values? If we want to stick to our postulate that all measurement results in quantum physics are expectation values of the corresponding operators, we must generalize the meaning of (3.28). Indeed, the result (3.29a) is that of a bolometric, i.e., total-energy measurement. In atomic physics, bolometers are out of use since the early days of optical spectroscopy.

In Zeeman's experiment, a photon emitted from one of the excited atomic m-substates is registered only if it passes a high-resolution optical spectrometer. This energy selection by the apparatus can be described by a *projection operator* P that filters out the state on which the spectrometer is tuned. Discussion of projection operators is postponed to Sect. 6.3, here we just want to point out that the definition of an expectation value $\psi^\dagger \mathrm{A}\psi$ in (3.28) requires some generalization when spectroscopic measurements intervene:

> For a spectroscopically resolved measurement, the expectation value $\psi'^\dagger \mathrm{A}\psi'$ must be taken with appropriately projected quantum states $\psi' = \mathrm{P}\psi$.

If energy eigenvalues E_+ or E_- are selected by the spectrometer, the projection operators transform ψ_D into

$$\mathrm{P}_+\psi_\mathrm{D} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a \\ 0 \end{pmatrix}, \tag{3.30a}$$

$$\mathrm{P}_-\psi_\mathrm{D} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ b \end{pmatrix}. \tag{3.30b}$$

In this spectroscopically fully resolved 2×2 case, the expectation values of H in (3.29b) project out the eigenvalues $E_\pm$ as $|a|^2 E_+$ or as $|b|^2 E_-$.

3.3.2 Expectation Value of Spin

Next we calculate the expectation value of the Pauli spin operator $\langle \boldsymbol{\sigma} \rangle = \psi^\dagger \boldsymbol{\sigma} \psi$. In (3.17a), we saw that any two-dimensional hermitian operator can be expressed in terms of the Pauli matrices. Therefore, for any two-state quantum system, the only thing we need to know are the expectation values of the Pauli matrices. We find from (3.9a)–(3.9c)

$$\langle \sigma_x \rangle = \psi^\dagger \sigma_x \psi = \begin{pmatrix} a^* & b^* \end{pmatrix} \begin{pmatrix} b \\ a \end{pmatrix} = a^* b + b^* a = 2\,\mathrm{Re}\big(a^* b\big), \tag{3.31a}$$

$$\langle \sigma_y \rangle = \psi^\dagger \sigma_y \psi = \begin{pmatrix} a^* & b^* \end{pmatrix} \begin{pmatrix} -\mathrm{i}b \\ \mathrm{i}a \end{pmatrix} = -\mathrm{i}\big(a^* b - b^* a\big) = 2\,\mathrm{Im}\big(a^* b\big), \tag{3.31b}$$

$$\langle\sigma_z\rangle = \psi^\dagger \sigma_z \psi = \begin{pmatrix} a^* & b^* \end{pmatrix} \begin{pmatrix} a \\ -b \end{pmatrix} = aa^* - bb^* = |a|^2 - |b|^2. \tag{3.31c}$$

Note that the expectation value $\langle\boldsymbol{\sigma}\rangle$ is an ordinary vector with real numbers $\langle\sigma_x\rangle, \langle\sigma_y\rangle, \langle\sigma_z\rangle$ as elements, in contrast to the vector $\boldsymbol{\sigma}$ with three Pauli matrices as components.

The state vector $\psi = a\alpha + b\beta$ used in (3.31a)–(3.31c), normalized to $|\psi|^2 = |a|^2 + |b|^2 = 1$, has, up to an unobservable arbitrary overall phase factor, the most general form

$$\psi = \alpha \mathrm{e}^{-\mathrm{i}\delta} \cos\gamma + \beta \mathrm{e}^{\mathrm{i}\delta} \sin\gamma, \tag{3.32a}$$

$$\psi = \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \mathrm{e}^{-\mathrm{i}\delta} \cos\gamma \\ \mathrm{e}^{\mathrm{i}\delta} \sin\gamma \end{pmatrix}, \tag{3.32b}$$

where angle γ should not be confused with the gyromagnetic ratio. With $|a|^2 = \cos^2\gamma$, $|b|^2 = \sin^2\gamma$, and $a^*b = \mathrm{e}^{2\mathrm{i}\delta} \sin\gamma\cos\gamma$ inserted into (3.31a)–(3.31c), we have

$$\langle\sigma_x\rangle = 2\,\mathrm{Re}\big(a^*b\big) = \cos 2\delta \sin 2\gamma, \tag{3.33a}$$

$$\langle\sigma_y\rangle = 2\,\mathrm{Im}\big(a^*b\big) = \sin 2\delta \sin 2\gamma, \tag{3.33b}$$

$$\langle\sigma_z\rangle = |a|^2 - |b|^2 = \cos 2\gamma. \tag{3.33c}$$

In this text, we repeatedly use the trigonometric double-angle formulas, as well as Euler's formulas, which we list for the convenience of the reader:

$$\cos 2\gamma = \cos^2\gamma - \sin^2\gamma, \qquad \sin 2\gamma = 2\sin\gamma\cos\gamma, \tag{3.34a}$$

$$\cos^2\gamma = \frac{1}{2}(1 + \cos 2\gamma), \qquad \sin^2\gamma = \frac{1}{2}(1 - \cos 2\gamma), \tag{3.34b}$$

$$\cos\varphi = \frac{1}{2}\big(\mathrm{e}^{-\mathrm{i}\varphi} + \mathrm{e}^{\mathrm{i}\varphi}\big), \qquad \sin\varphi = \frac{1}{2}\mathrm{i}\big(\mathrm{e}^{-\mathrm{i}\varphi} - \mathrm{e}^{\mathrm{i}\varphi}\big), \tag{3.34c}$$

$$\mathrm{e}^{\mathrm{i}\varphi} = \cos\varphi + \mathrm{i}\sin\varphi. \tag{3.34d}$$

As can be checked by direct computation, $\langle\boldsymbol{\sigma}\rangle = (\langle\sigma_x\rangle, \langle\sigma_y\rangle, \langle\sigma_z\rangle)$ is a unit vector

$$\langle\boldsymbol{\sigma}\rangle^2 = \langle\sigma_x\rangle^2 + \langle\sigma_y\rangle^2 + \langle\sigma_z\rangle^2 = 1. \tag{3.35}$$

The expectation value of the Pauli operator, the *spin polarization* vector $\boldsymbol{P} = \langle\boldsymbol{\sigma}\rangle$, is a unit vector of magnitude $|\boldsymbol{P}| = |\langle\boldsymbol{\sigma}\rangle| = 1$.

Being a unit vector, $\boldsymbol{P}$ can be written in polar coordinates as

$$\boldsymbol{P} = \langle\boldsymbol{\sigma}\rangle = \begin{pmatrix} \langle\sigma_x\rangle \\ \langle\sigma_y\rangle \\ \langle\sigma_z\rangle \end{pmatrix} = \begin{pmatrix} \cos\varphi \sin\theta \\ \sin\varphi \sin\theta \\ \cos\theta \end{pmatrix}, \tag{3.36}$$

with polar and azimuthal angles θ and φ, see Fig. 5.1(a).

Comparison of this with (3.33a)–(3.33c) gives

$$\delta = \frac{1}{2}\varphi, \qquad \gamma = \frac{1}{2}\theta. \tag{3.37}$$

Equations (3.37) show a remarkable result:

> For spin $s = \frac{1}{2}$ the phases angles γ and δ of the wave function ψ in Hilbert space change at "half speed" when we change the *polar* and *azimuthal* angles φ and θ of the polarization vector $\boldsymbol{P} = \langle \boldsymbol{\sigma} \rangle$ in ordinary three-dimensional space.

Equation (3.32b) for the most general two-state vector then turns into

$$\psi = \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \mathrm{e}^{-\mathrm{i}\varphi/2} \cos \frac{1}{2}\theta \\ \mathrm{e}^{\mathrm{i}\varphi/2} \sin \frac{1}{2}\theta \end{pmatrix}. \tag{3.38}$$

Again, if a spin measurement intervenes that filters out a certain substate of angular momentum, then $\langle \boldsymbol{\sigma} \rangle = \psi'^{\dagger} \boldsymbol{\sigma} \psi'$ must be taken, with $\psi' = \mathrm{P}\psi$ being some projection onto a certain spin axis in Hilbert space (Sect. 6.3).

What about other observables? All spin-$\frac{1}{2}$ operators, represented by hermitian 2×2 matrices, are expandable in Pauli spin matrices, such as $\mathrm{A} = \boldsymbol{\alpha} \cdot \boldsymbol{\sigma}$, with some coefficient vector $\boldsymbol{\alpha}$ as in (3.17b), and with expectation value $\langle \mathrm{A} \rangle = \boldsymbol{\alpha} \cdot \langle \boldsymbol{\sigma} \rangle$. For instance, an electric dipole moment $\langle \mathbf{d}_n \rangle$ of a neutron, if it exists, must be proportional to $\langle \boldsymbol{\sigma} \rangle$, just like the magnetic dipole moment $\langle \boldsymbol{\mu}_n \rangle$. In Chap. 17 and following on irreducible tensors, we shall generalize this to angular momenta $j > \frac{1}{2}$ and to interactions beyond the dipolar interaction.

3.4 Uncertainties

The *variance* of a random variable X is defined as the mean square deviation from its mean value $\langle X \rangle$, namely,

$$(\Delta X)^2 \equiv \big\langle \big(X - \langle X \rangle \big)^2 \big\rangle. \tag{3.39}$$

Multiplied out, this is $(\Delta X)^2 = \langle X^2 - 2X\langle X \rangle + \langle X \rangle^2 \rangle$. For uncorrelated events, the average of a sum is the sum of the averages, with the well known result

$$(\Delta X)^2 = \big\langle X^2 \big\rangle - \langle X \rangle^2. \tag{3.40}$$

Applied to the expectation value $\langle \mathrm{A} \rangle$ of a quantum operator A, (3.28), this gives

$$(\Delta \mathrm{A})^2 = \big\langle \mathrm{A}^2 \big\rangle - \langle \mathrm{A} \rangle^2 = \psi^{\dagger} \mathrm{A}^2 \psi - \big(\psi^{\dagger} \mathrm{A} \psi \big)^2. \tag{3.41}$$

If the quantum system is in the nth eigenstate of A with eigenvalue a_n, then $\psi^\dagger \mathrm{A}^2 \psi = a_n^2 \psi^\dagger \psi = a_n^2$, and the same result for $(\psi^\dagger \mathrm{A} \psi)^2 = a_n^2$, and the variance disappears, $(\Delta A)^2 = \langle \mathrm{A}^2 \rangle - \langle \mathrm{A} \rangle^2 = 0$:

> If a quantum system is in an eigenstate of an operator A, then the variance of A disappears, and the corresponding eigenvalue of A can, in principle, be measured with arbitrary precision.

This implies:

> If two or more observables and their operators have the same set of eigenstates, then they can be measured simultaneously with arbitrary precision.

Examples are the observables $\mathbf{J}^2$ and J_z of angular momentum which can be both measured precisely because they have the eigenstates ψ_{jm} in common. The individual components J_x, J_y, and J_z, on the other hand, do not have a common set of eigenstates and cannot simultaneously precisely be measured. The vector operator $\mathbf{J}$ therefore has a finite variance that we want to calculate for the spin-$\frac{1}{2}$ case.

3.4.1 Uncertainty of Spin

We know from (3.35) that $\langle \boldsymbol{\sigma} \rangle^2 = \langle \sigma_x \rangle^2 + \langle \sigma_y \rangle^2 + \langle \sigma_y \rangle^2 = 1$. We further know from (3.10) that $\boldsymbol{\sigma}^2 = 3\mathrm{I}$, hence

$$\left\langle \boldsymbol{\sigma}^2 \right\rangle = \left\langle \sigma_x^2 \right\rangle + \left\langle \sigma_x^2 \right\rangle + \left\langle \sigma_x^2 \right\rangle = 3. \tag{3.42}$$

As a consequence, the variance of $\boldsymbol{\sigma}$ is

$$(\Delta\sigma)^2 = \left\langle \boldsymbol{\sigma}^2 \right\rangle - \langle \boldsymbol{\sigma} \rangle^2 = 3 - 1 = 2. \tag{3.43}$$

The operators σ_z and $\boldsymbol{\sigma}^2$ have α and β as common eigenfunctions, with eigenvalues ± 1 and 3 from (3.9c) and (3.10). σ_z and $\boldsymbol{\sigma}^2$ hence can be measured simultaneously with no uncertainty. Therefore, all uncertainty of the spin operator $\boldsymbol{\sigma}$, (3.43), must reside in its transverse component $\boldsymbol{\sigma}_\perp$, with $\Delta\sigma_\perp = \Delta\sigma = \sqrt{2}$. If we assume rotational symmetry about the quantization axis z, then the expectation value of its transverse component must vanish, $\langle \boldsymbol{\sigma}_\perp \rangle = \langle \pm(\sigma_x^2 + \sigma_y^2)^{\frac{1}{2}} \rangle = 0$, hence:

> For a single-particle, the transverse component of its spin operator has the expectation value $\langle \boldsymbol{\sigma}_\perp \rangle \pm \Delta\sigma_\perp = 0 \pm \sqrt{2}$, which signals complete uncertainty for the azimuthal spin phase φ.

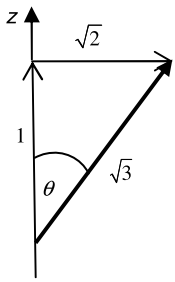


Fig. 3.1 The root mean square expectation values of the Pauli spin operator and its components are $\langle \boldsymbol{\sigma}^2 \rangle^{\frac{1}{2}} = \sqrt{3}$, $\langle \sigma_z^2 \rangle^{\frac{1}{2}} = 1$, and $\langle \boldsymbol{\sigma}_\perp^2 \rangle^{\frac{1}{2}} = \sqrt{2}$. In a semiclassical picture, these three quantities can be interpreted as the components of a spin vector

Semiclassically, one can define the magnitude of the spin vector $\boldsymbol{\sigma}$ and of its longitudinal component σ_z as the root mean square expectation values $\langle \boldsymbol{\sigma}^2 \rangle^{\frac{1}{2}} = \sqrt{3}$, and $\langle \sigma_z^2 \rangle^{\frac{1}{2}} = 1$, respectively. In the same way, one can define $\langle \boldsymbol{\sigma}_\perp^2 \rangle^{\frac{1}{2}} = (\langle \sigma_x^2 \rangle + \langle \sigma_y^2 \rangle)^{\frac{1}{2}} = \sqrt{2}$ as the transverse component $\sigma_\perp$ of the spin vector, which gives the semiclassical spin vector picture of Fig. 3.1.

3.4.2 Uncertainty of Energy

In the Zeeman effect, energy is measured in a spectroscopically resolved way which projects out the eigenstate E_+ or E_- with zero uncertainty, as in (3.30a), (3.30b). The same in principle is true for atomic energy levels with eigenvalues E_n. Still, the energies of excited atomic states are not perfectly sharp.

Let us take an atomic optical transition between an excited state E_2 and the stable ground state E_1 with $E_2 - E_1 = \hbar\Omega$. Let the energy be measured by using the wave train of light that is emitted during the excited state's lifetime τ, which typically is of the order of nanoseconds for the main allowed transitions between low-lying atomic states. Semiclassically, the frequency spectrum of the emitted light is the Fourier transform of the excited state's exponentially damped oscillation, which is the *Lorentzian* of full natural line width $\Delta\Omega = 1/\tau$,

$$g(\Omega) \propto \frac{1}{1 + 4\tau^2(\Omega - \Omega_0)^2}. \tag{3.44}$$

With photon energy $E_{\text{ph}} = \hbar\Omega$, this line width $\Delta\Omega$ translates into an energy uncertainty $\Delta E_{\text{ph}} = \hbar\Delta\Omega = \hbar/\tau$ of the emitted photons. In quantum mechanics, physical quantities are defined solely by measurement. Since the energy of the excited atomic state is measured via the photon energy, the photon's energy uncertainty translates into an energy uncertainty ΔE of the excited atomic state, while the stable ground state has no natural width.

With $\Delta E = \hbar\Delta\Omega$, the *time–energy uncertainty* is

$$\Delta E \cdot \tau \geq \hbar, \tag{3.45}$$

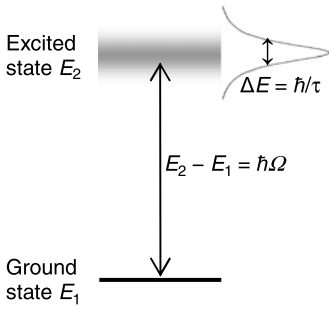


Fig. 3.2 The bandwidth $\Delta\Omega = 1/\tau$ of a radiative transition from a short-lived excited atomic state of lifetime τ to the stable ground state translates into an energy uncertainty $\Delta E = \hbar/\tau$ of the excited state

or $\Delta E \geq 7 \times 10^{-7}$ eV for $\tau = 1$ ns, see Fig. 3.2. For most experiments, semiclassical theory works surprisingly well.

3.5 Spin Precession

We next derive the time dependent response of our spin-$\frac{1}{2}$ system to a magnetic field and see how polarization $\boldsymbol{P}(t) = \langle\boldsymbol{\sigma}\rangle(t)$ evolves from an initial $\boldsymbol{P}(0) = \langle\boldsymbol{\sigma}\rangle(0)$ for two simple field configurations. We write the time dependence (t) outside of $\langle\boldsymbol{\sigma}\rangle$ because the operator $\boldsymbol{\sigma}$ is time independent. The general solution of the 2×2 problem will be postponed to Chap. 5, after we have studied a number of simple spin rotation experiments in Chap. 4.

3.5.1 Longitudinal Field

We first solve the Schrödinger equation for quantization axis z along the magnetic field $\boldsymbol{B} = (0, 0, B_z)$, which we call the *longitudinal* field configuration, with the diagonal Hamiltonian

$$\mathrm{H} = -\mu\sigma_z B_z, \tag{3.46a}$$

$$\mathrm{H} = -\mu \begin{pmatrix} B_z & 0 \\ 0 & -B_z \end{pmatrix}. \tag{3.46b}$$

This means that for $\boldsymbol{B}$ along z, the energy eigenfunctions simultaneously are spin eigenfunctions. The Schrödinger equation (3.13) is

$$\dot{\psi} = \frac{\mathrm{i}}{\hbar}\mu\sigma_z B_z \psi, \tag{3.47a}$$

$$\begin{pmatrix} \dot{a} \\ \dot{b} \end{pmatrix} = \frac{\mathrm{i}}{\hbar}\mu \begin{pmatrix} B_z & 0 \\ 0 & -B_z \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}. \tag{3.47b}$$

The equations are decoupled, $\dot{a} = -\frac{1}{2}\mathrm{i}\omega_\mathrm{L} a$ and $\dot{b} = \frac{1}{2}\mathrm{i}\omega_\mathrm{L} b$, with $\frac{1}{2}\hbar\omega_\mathrm{L} = -\mu B_z$ from (2.15) with (2.11) and (2.7). If initially the system is in a state $\psi(0) = a(0)\alpha + b(0)\beta$, after time t it is in the state $\psi(t) = a(t)\alpha + b(t)\beta$, with

$$a(t) = a(0)\mathrm{e}^{-\mathrm{i}\omega_\mathrm{L} t/2}, \qquad b(t) = b(0)\mathrm{e}^{\mathrm{i}\omega_\mathrm{L} t/2}. \tag{3.48}$$

The probability of being spin up or spin down does not change in the process,

$$p_\mathrm{up}(t) = \big|a(t)\big|^2 = \big|a(0)\big|^2, \qquad p_\mathrm{down}(t) = \big|b(t)\big|^2 = \big|b(0)\big|^2. \tag{3.49}$$

The particle's wave function ψ, (3.24) and (3.48), merely acquires an additional *dynamical phase* $\delta_\pm(t) = \mp\frac{1}{2}\omega_\mathrm{L} t$, or, for a time dependent field $B_z(t)$,

$$\delta_\pm(t) = \mp\frac{1}{2}\int_0^t \omega_\mathrm{L}\big(t'\big)\,\mathrm{d}t'. \tag{3.50}$$

We insert the solutions $a(t)$, $b(t)$ into (3.31a)–(3.31c) and obtain, for real $a(0)$ and $b(0)$, that is, for an initial $\langle\sigma_x\rangle(0) = 2a(0)b(0)$ and $\langle\sigma_y\rangle = 0$,

$$\langle\sigma_x\rangle(t) = 2\,\mathrm{Re}\big(a^* b\big) = 2a(0)b(0)\,\mathrm{Re}\big(\mathrm{e}^{\mathrm{i}\omega_\mathrm{L} t}\big) = \langle\sigma_x\rangle(0)\cos\omega_\mathrm{L} t, \tag{3.51a}$$

$$\langle\sigma_y\rangle(t) = 2\,\mathrm{Im}\big(a^* b\big) = 2a(0)b(0)\,\mathrm{Im}\big(\mathrm{e}^{\mathrm{i}\omega_\mathrm{L} t}\big) = \langle\sigma_x\rangle(0)\sin\omega_\mathrm{L} t, \tag{3.51b}$$

$$\langle\sigma_z\rangle(t) = \big|a(t)\big|^2 - \big|b(t)\big|^2 = \big|a(0)\big|^2 - \big|b(0)\big|^2 = \langle\sigma_z\rangle(0). \tag{3.51c}$$

While the longitudinal polarization $P_z = \langle\sigma_z\rangle(0) = |a(0)|^2 - |b(0)|^2$ is time independent, the transverse polarization of constant amplitude $\langle\sigma_y\rangle(0)$ rotates in the x–y plane about the field $\boldsymbol{B}$. Setting $P_{x0} = P_x(0)$ and $P_{z0} = P_z(0)$:

The polarization vector precesses about $\boldsymbol{B} = (0, 0, B_z)$ as

$$\boldsymbol{P}(t) = \langle\boldsymbol{\sigma}\rangle(t) = (P_{x0}\cos\omega_\mathrm{L} t,\, P_{x0}\sin\omega_\mathrm{L} t,\, P_{z0}). \tag{3.52}$$

3.5.2 Transverse Field

For further practice, we next choose a *transverse* field configuration, with the magnetic field $\boldsymbol{B} = (B_x, 0, 0)$ directed along x. The Schrödinger equation then reads

$$\dot{\psi} = \frac{\mathrm{i}}{\hbar}\mu\sigma_x B_x \psi, \tag{3.53a}$$

$$\begin{pmatrix}\dot{a}\\ \dot{b}\end{pmatrix} = \frac{\mathrm{i}}{\hbar}\mu\begin{pmatrix}0 & B_x\\ B_x & 0\end{pmatrix}\begin{pmatrix}a\\ b\end{pmatrix}, \tag{3.53b}$$

$$\dot{a} = -\frac{1}{2}\mathrm{i}\omega_\mathrm{L} b, \qquad \dot{b} = -\frac{1}{2}\mathrm{i}\omega_\mathrm{L} a, \tag{3.53c}$$

with $\frac{1}{2}\hbar\omega_\mathrm{L} = -\mu B_x$. Following the usual procedure for the classical coupled pendulum, we decouple the equations by forming the *symmetric* and *antisymmetric* combinations of a and b,

$$c = \sqrt{\frac{1}{2}}(a+b), \qquad d = \sqrt{\frac{1}{2}}(a-b). \tag{3.54}$$

This reduces the problem to the previous diagonal case, with

$$\dot{c} = -\frac{1}{2}\mathrm{i}\omega_\mathrm{L} c, \qquad \dot{d} = \frac{1}{2}\mathrm{i}\omega_\mathrm{L} d. \tag{3.55}$$

We solve this for the case that initially all spins are up, $a(0) = 1$ and $b(0) = 0$, that is, $c(0) = d(0) = \sqrt{\frac{1}{2}}$. With $c(t) = \sqrt{\frac{1}{2}}\mathrm{e}^{-\frac{1}{2}\mathrm{i}\omega_\mathrm{L} t}$, $d(t) = \sqrt{\frac{1}{2}}\mathrm{e}^{\frac{1}{2}\mathrm{i}\omega_\mathrm{L} t}$ from (3.48), and with

$$a = \sqrt{\frac{1}{2}}(c+d), \qquad b = \sqrt{\frac{1}{2}}(c-d), \tag{3.56}$$

we find the solutions

$$a(t) = \cos\left(\frac{1}{2}\omega_\mathrm{L} t\right), \qquad b(t) = -\mathrm{i}\sin\left(\frac{1}{2}\omega_\mathrm{L} t\right). \tag{3.57}$$

The probability of being spin up or spin down then oscillates as

$$p_\mathrm{up} = \left|a(t)\right|^2 = \cos^2\left(\frac{1}{2}\omega_\mathrm{L} t\right) = \frac{1}{2}(1+\cos\omega_\mathrm{L} t), \tag{3.58a}$$

$$p_\mathrm{down} = \left|b(t)\right|^2 = \sin^2\left(\frac{1}{2}\omega_\mathrm{L} t\right) = \frac{1}{2}(1-\cos\omega_\mathrm{L} t). \tag{3.58b}$$

Again, we study the polarization vector $\boldsymbol{P} = \langle\boldsymbol{\sigma}\rangle = \psi^\dagger\boldsymbol{\sigma}\psi$ and find

$$\langle\sigma_x\rangle(t) = 2\,\mathrm{Re}\big(a^* b\big) = 0, \tag{3.59a}$$

$$\langle\sigma_y\rangle(t) = 2\,\mathrm{Im}\big(a^* b\big) = -\sin\omega_\mathrm{L} t, \tag{3.59b}$$

$$\langle\sigma_z\rangle(t) = \left|a(t)\right|^2 - \left|b(t)\right|^2 = \cos\omega_\mathrm{L} t, \tag{3.59c}$$

or

$$\boldsymbol{P}(t) = P_{z0}(0, -\sin\omega_\mathrm{L} t, \cos\omega_\mathrm{L} t). \tag{3.60}$$

Starting with $P = P_z(0) = P_{z0}$, the polarization precesses in the z–y plane about the field B_x. This is what we expect from rotational symmetry because the present configuration differs from the previous configuration merely by a rotation of the

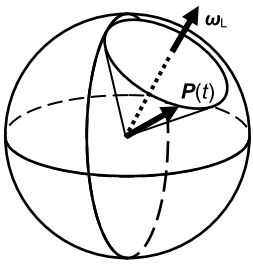


Fig. 3.3 The spin polarization $\boldsymbol{P}(t) = \langle \mathbf{S} \rangle / \hbar s$ precesses about a magnetic field $\boldsymbol{B}$ with Larmor frequency vector $\boldsymbol{\omega}_L = \gamma \boldsymbol{B}$

coordinate axes. For the electron with $\omega_L > 0$, both (3.60) and (3.52) describe a counterclockwise spin rotation.

Can we observe the rotating polarization $\boldsymbol{P}(t)$ from (3.52) or (3.60) experimentally? From Sect. 3.4.1 we know that after one spin measurement on a single ($N = 1$) particle, the phase φ_1 of the angular momentum vector is completely uncertain, with uncertainty $\Delta\varphi_1 = \pm\pi$. Let us do a measurement on an ensemble of $N \gg 1$ independent particles. According to the rules of error propagation, the measurement with N equal particles diminishes the error by a factor of $1/\sqrt{N}$, and the phase uncertainty diminishes to $\Delta\varphi_N = \Delta\varphi_1/\sqrt{N} = \pm\pi/\sqrt{N}$.

The uncertainty relation between *particle number and phase* is

$$\Delta\varphi\sqrt{N} \sim 1. \tag{3.61}$$

Therefore, in a sample with a large number N of magnetic particles, a transverse polarization $\langle \boldsymbol{\sigma}_\perp \rangle$ can have a very well defined phase.

Let us draw some conclusions. When an external magnetic field is applied to a magnetic spin-$\frac{1}{2}$ system, its polarization performs *Larmor precession* about this field, as shown in Fig. 3.3. We shall later find that this result holds for the dipole interaction of any angular momentum $s \geq \frac{1}{2}$. In the literature, a quantum result is often said to be "explained first in a hand-waving manner via Larmor precession", to continue: "now let us prove the same result quantum mechanically". The second step is unnecessary because, as we stressed before, only expectation values of observables have relevance in quantum physics. The only significant vector in a two-state system is the vector operator $\boldsymbol{\sigma}$, and its expectation value is the spin polarization $\boldsymbol{P} = \langle \boldsymbol{\sigma} \rangle$, which precesses as given by (3.52). Certainly, other mathematical objects of quantum physics like ψ or $\boldsymbol{\sigma}$ are not described by Larmor precession, but these quantities are not measurable expectation values.

Larmor precession is a genuine quantum result.

Larmor precession is semiclassical only in the sense that one does not quantize the external field $\boldsymbol{B}$, whose state is not significantly changed by the observation of

the spin system, but in this respect also the Schrödinger equation (3.2) is semiclassical. In Part IV of this treatise, we shall see that the Larmor precession picture is the irreducible representation of the problem, where *irreducible* designates the physically most adequate and pictorial description of a given problem.

References

Feynman, R.P., Vernon, F.L. Jr., Hellwarth, R.W.: Geometrical representation of the Schrödinger equation for solving maser problems. J. Appl. Phys. **28**, 49–52 (1957)

Chapter 4
Experiments on Spin Precession

Abstract This chapter presents several basic experiments on the precession of spins. While the results derived in the preceding chapters are valid for any quantum mechanical two-state system, whatever the origin of the matrix elements of the 2×2 Hamiltonian, we limit this first discussion to the precession of true angular momenta. We treat spin precession of muons, atoms, and neutrons. The Hanle effect is derived as a spin rotation effect detected via light scattering, and analogies with Thomson scattering are shown. We present experimental proof for the quantum prediction that a spin-$\frac{1}{2}$ wave function needs two full rotations in Hilbert space to come back to its original state.

4.1 Muon Spin Precession

To observe spin rotation, we must go through three steps:

Step 1: Prepare an ensemble of polarized particles;
Step 2: Apply a transverse field to disturb this ensemble;
Step 3: Detect the response of the polarized ensemble.

> The three steps of preparation, evolution, and analysis of a system are typical for a large class of problems in physics and other fields.

These steps can be realized in many different ways, as we shall see in the examples that follow. We start with an experiment on the spin rotation of polarized muons. Muons μ^- and their antiparticles μ^+ are members of the second of the three known *families of particles*, and differ from electrons e^- and their antiparticles, the positrons e^+, of the first family only by their mass $m_\mu = 106\ \mathrm{MeV}/c^2$, which is 207 times the electron mass. From (2.6), the magnetic moment of the muon then is 207 times smaller than that of the electron.

Step 1, muon polarization: Muons from pion decay are almost fully polarized, the spins of the μ^- pointing opposite to their direction of flight, and in the direction of flight for the μ^+.

D. Dubbers, H.-J. Stöckmann, *Quantum Physics: The Bottom-Up Approach*, Graduate Texts in Physics,
DOI 10.1007/978-3-642-31060-7_4, © Springer-Verlag Berlin Heidelberg 2013

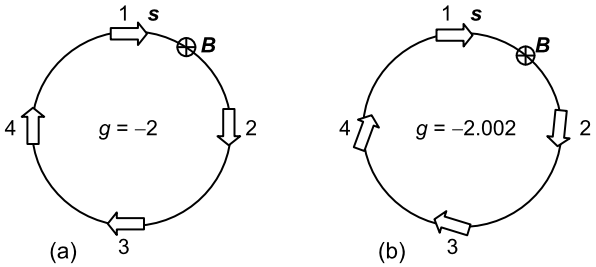


Fig. 4.1 Muon spin precession in-flight in the transverse magnetic field of 1.4 T of a storage ring of 14 m diameter. Shown are four positions of the muon on its first passage. (**a**) If the g-factor of the muon were exactly $g = -2$, then muon spin (*block arrow*) would always point along the muon line of flight. (**b**) Due to effects from quantum electrodynamics, g does not exactly equal -2. Therefore, the muons show a slow spin precession away from their momentary direction of flight

Muons are produced by bombarding a metal target with a multi-GeV proton beam, thus producing a large flux of spinless pions (mass 140 GeV/c^2), which for kinematic reasons fly all into near forward direction. Each pion in turn almost immediately decays in-flight into a muon plus the corresponding muon-(anti)neutrino, both also into forward beam direction. Pion decay is due to the *weak interaction*, one of the four known interactions in nature. Muon polarization from π-decay into direction z is $P_z = -v_\mu/c$, with muon velocity v_μ, i.e., muons are *left-handed*. This *helicity* of decay muons is due to the maximal *violation of parity* in weak interactions, which essentially means that the right-handed mirror image of a left-handed weak-interaction process does not exist in nature.

Step 2, muon spin precession: The polarized muons are guided into the transverse magnetic field of a storage ring, as shown in Fig. 4.1(a). Due to the Lorentz force, (2.19), the charged muons follow the ring's circular track, rotating at the *cyclotron angular frequency*

$$\omega_C = \frac{e}{m_\mu} B. \tag{4.1}$$

At the same time, the muon spins precess in-flight with Larmor angular frequency, see (2.15) with (2.11) and (2.6),

$$\omega_L = -\frac{1}{2} g \frac{e}{m_\mu} B. \tag{4.2}$$

Hence if the muon's spin g-factor was exactly $g = g_s = -2$, as predicted by Dirac theory, the Larmor precession frequency ω_L of the muon would equal its cyclotron frequency ω_C. In this case, the spin of the muon on its circular path in the storage ring would always point along the muon's momentary direction of flight.

In reality, the g-factor of the muon is not exactly -2. The theory of quantum electrodynamics predicts that the g-factor is, to first order, $|g| \approx 2 + \alpha/\pi$, with the fine-structure constant $\alpha \approx 1/137$. Therefore, muon spins rotate relative to their circular line of flight at angular frequency

$$\Delta\omega = \omega_L - \omega_C = \frac{1}{2}\big(|g| - 2\big)\omega_C, \tag{4.3}$$

and their frequency of rotation is proportional to this small deviation of $|g|$ from 2. Therefore, for each turn of the muon on its circular trajectory, its spin direction changes by another incremental angle, see Fig. 4.1(b).

In a vacuum, the muon not only sees the magnetic field $\boldsymbol{B}$ of the storage ring, but also *vacuum fluctuations*, the same that lead to the Lamb shift in the energy levels of the hydrogen atom. *Virtual* particle–antiparticle excitations lead to a (charge-)*polarization of the vacuum* near the muon. A precise measurement of the muon's g-factor enables us to check whether there are particles contributing to these fluctuations that have not been discovered yet in high-energy experiments. The muon is more sensitive to such heavy exotic particles than the electron by a factor of $(m_\mu/m_e)^2 = 4 \times 10^4$. Therefore, the electron's $|g| - 2$ is used to measure the precise value of α, while the muon's $|g| - 2$ is used to search for new physics.

Step 3, polarization analysis: Muons, like the pions, are unstable under the weak interaction. They transform via *beta decay* into an electron and two neutrinos of different kinds, $\mu^- \to e^- + \bar{\nu}_e + \nu_\mu$, with a lifetime $\tau_\mu = 2.2\,\mu s$. The precession of muon polarization in the storage ring can be monitored via the asymmetric electron emission pattern, which rotates synchronously with muon polarization at angular frequency $\Delta\omega$.

Let the muons enter the storage ring at time $t = 0$. At the time t of muon decay, the precessing muon polarization points into a direction given by the phase angle $\varphi = \Delta\omega t$. In the same way as the muons emitted in pion decay, the electrons emitted in muon decay carry their spin preferably against their direction of emission. Conservation of angular momentum then requires that the electrons, which carry away part of muon polarization, be preferentially emitted against the momentary direction of muon spin (more on this in Sect. 20.5.1). When the polarization of the muons rotates at angular frequency $\Delta\omega$, so does this parity-violating *β-decay asymmetry*. Electron detectors installed near the muon beam then see the electron counting rate modulated at the angular frequency $\Delta\omega$ of the precessing muons, in the same way as a bystander on a sidewalk sees the rotating flash-light of a passing ambulance car.

The result of this so-called $g - 2$ experiment (Bennett et al. 2006) was

$$a_\mu^{\text{exp}} = 1.16592080(63) \times 10^{-3}, \tag{4.4}$$

with $a_\mu = \frac{1}{2}(|g| - 2)$. The theoretical prediction (Jegerlehner and Nyffeler 2009), based on the present *standard model of particle physics*, is

$$a_\mu^{\text{theor}} = 1.16591790(65) \times 10^{-3}. \tag{4.5}$$

The numbers in parentheses are the standard error in the last two digits. The difference between both numbers, with the errors added in quadrature,

$$\Delta a_\mu = a_\mu^{\text{exp}} - a_\mu^{\text{theor}} = (290 \pm 90) \times 10^{-11}, \tag{4.6}$$

deviates from zero by more than three standard deviations. This is either due to unrecognized errors in a_μ^{exp} or a_μ^{theor}, or is a hint to new physics beyond the standard model.

4.2 Light Scattering

Our second example is from atomic physics. We study the *resonant scattering* of light on atoms, without and with an externally applied magnetic field.

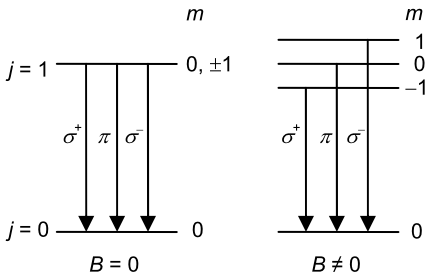


Fig. 4.2 Optical transitions between excited state and ground state of an atom, without and with an external B-field applied. The emitted σ^+ and σ^- light is *left* and *right* circularly polarized, π light is linearly polarized

Light scattering by atoms can be studied without sophisticated equipment. One places a piece of alkali metal into a glass bulb and, if necessary after moderate heating, irradiates the evaporated atoms with a beam of light, see Fig. 4.4(a). The resonantly scattered photons are registered in a photodiode or in a photomultiplier tube outside the cell. Modern experiments work with one or few atoms caught in an optical trap exposed to a laser beam.

Step 1, atomic polarization: Atoms can be polarized by the absorption of light. We choose the simplest case of an atom with a $j = 0$ ground state and a $j = 1$ excited state. A beam of light, incident along x and linearly polarized along the quantization axis z, will excite only the $j = 1$, $m = 0$ substate. Unpolarized light is a statistical mixture of σ^+ and σ^- photons, and absorption of unpolarized light incident along z excites only the $m = \pm 1$ substates, for details see Chap. 12. We first assume that the atoms are in zero magnetic field, and we go directly to Step 3.

Step 3, polarization analysis: Light from unpolarized atoms is emitted isotropically, due to rotational symmetry, while light from polarized atoms usually has an anisotropic angular distribution. In our example of a $j_i = 1$ to $j_f = 0$ decay shown in Fig. 4.2, the reemitted π light in a $m_i = 0$ to $m_f = 0$ transition has the angular distribution of a classical dipole antenna

$$I(\theta) = \frac{1}{2} I_0 \sin^2\theta, \quad \pi \text{ light.} \tag{4.7}$$

This distribution is rotationally symmetric about the axis z of polarization of the incoming light, as shown in Fig. 4.3(a). θ is the angle of emission of the outgoing light with respect to the direction z of linear polarization of the incoming light.

If only the $m = \pm 1$ substates are excited, the angular distribution of the reemitted σ light is that of a classical ring antenna

$$I(\theta) = \frac{1}{4} I_0 \big(1 + \cos^2\theta\big), \quad \sigma^+ \text{ or } \sigma^- \text{ light,} \tag{4.8}$$

rotationally symmetric about z, with the angle θ between in and outgoing light, see Fig. 4.3(b), and Sect. 20.5.2 for a derivation of these angular distributions.

Step 2, atomic spin rotation: After excitation of the $j = 1$, $m = 0$ substate, we apply a magnetic field $B \neq 0$ along the direction x of the incoming light, see Fig. 4.4(a). The polarization of the excited atoms will, during the lifetime of the excited state, precess about $\boldsymbol{B}$ with angular frequency $\omega_\mathrm{L} = \gamma B$. Like in the muon case discussed above, the light emission pattern precesses synchronously with the atom's polarization. In the case of π light emission, at a time t after the excitation of

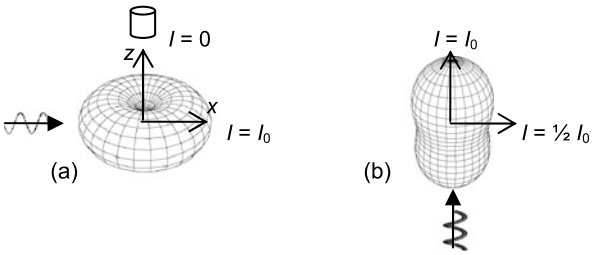


Fig. 4.3 Angular distribution of light scattered from atoms. (**a**) π light: If atoms are excited by a beam of linearly polarized light, the reemitted light has the angular distribution (4.7). (**b**) σ light: If the atoms are excited by a beam of unpolarized light or of circularly polarized light, the reemitted σ^+ plus σ^- light has the angular distribution (4.8)

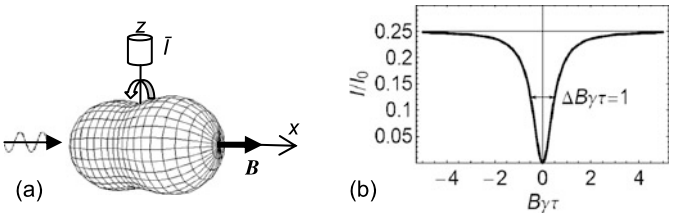


Fig. 4.4 Hanle effect. (**a**) Spin precession: In a strong magnetic field along the axis of the incoming light, the long-time average of the rotating π light distribution of Fig. 4.3(a) approaches a σ distribution, Fig. 4.3(b), with respect to the axis x of the incident beam of light. (**b**) Hanle signal as seen in the detector along axis z: For a zero magnetic field, the detector sees zero intensity, as in Fig. 4.3(a). In a strong magnetic field, the detector sees a mean intensity $\bar{I} = \frac{1}{4} I_0$. The full magnetic-field dependence of the time average light intensity follows (4.10). The width of the Hanle signal gives the lifetime τ of the excited state

the atom, the photon intensity in the detector under $\theta = 0$ will have increased from zero to $I_0 \sin^2 \omega_L t$, from (4.7). At the same time, the intensity of the emitted light will have dropped by another factor of $e^{-t/\tau}$, with atomic lifetime τ.

For large field values B, Fig. 4.4(a) shows the long-time averaged distribution of the scattered light, which is calculated from (4.7) to

$$I(\theta') = \frac{1}{4} I_0 (1 + \cos^2 \theta'), \quad \text{for } 4\gamma^2 B^2 \tau^2 \gg 1. \tag{4.9}$$

It has the same shape as (4.8), but with respect to axis x.

To show this, we regard the π light intensity $I(\theta) = \frac{1}{2} I_0 \sin^2 \theta$, rotationally symmetric about z, with polar angle θ towards z, emitted in direction of the unit vector $\boldsymbol{n} = (\cos\varphi \sin\theta, \sin\varphi \sin\theta, \cos\theta)$. We need to express this intensity in terms of two new angles, the angle θ' of $\boldsymbol{n}$ towards the precession axis x, and the precession angle $\varphi' = \omega_L t$ of $\boldsymbol{n}$ in the y–z plane. When we rotate coordinate axes from z to x, we find the relation $\cos\theta = \cos\varphi' \sin\theta'$ for the z-component of $\boldsymbol{n}$. Inserting this into (4.7) gives $I(\theta', t) = \frac{1}{2} I_0 (1 - \cos^2 \omega_L t \sin^2 \theta')$. Averaging over time gives $\bar{I}(\theta') = \frac{1}{2} I_0 (1 - \frac{1}{2} \sin^2 \theta')$, which equals (4.9).

We now turn to arbitrary values of B. As the sample is continuously irradiated with photons, and as the detector continuously counts the reemitted photons, we have to integrate over all prior times up to the present $t = 0$ to obtain the *Hanle signal* as a function of the magnetic field

$$\bar{I}(B) = \frac{1}{2} I_0 \frac{1}{\tau} \int_{-\infty}^{0} \mathrm{e}^{t/\tau} \sin^2(\gamma B t)\, \mathrm{d}t = \frac{\gamma^2 B^2 \tau^2}{1 + 4\gamma^2 B^2 \tau^2} I_0. \tag{4.10}$$

Intensity $\bar{I}$ vanishes for zero magnetic field. For large external fields, $\bar{I}$ rises to $\frac{1}{2} I_0$. The Hanle signal, shown in Fig. 4.4(b) as $\bar{I}(B)/I_0$ vs. $\gamma B \tau$, reaches half maximum at field values $B_{1/2} = \pm 1/(2\gamma\tau)$. With gyromagnetic ratios γ known from the Zeeman signal, atomic lifetimes τ can be derived from the Hanle effect. Our derivation of the Hanle signal is quantum mechanically correct, in the spirit of our remarks at the end of Chap. 3.

An angular distribution identical to that of (4.8) is found in *Thomson scattering* of unpolarized photons on electrons, which is Compton scattering in the elastic limit. Unpolarized light being a statistical mixture of σ^+ and σ^- photons, it is of no surprise that the differential Thomson scattering cross-section has the same angular distribution as $\sigma^\pm$ light, see (4.8),

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{1}{2}\sigma_0 \big(1 + \cos^2\theta\big), \tag{4.11}$$

with a prefactor $\sigma_0 = (\alpha\hbar c/m_e c^2)^2 = 7.9\ \mathrm{fm}^2$.

4.3 Spinor Rotation Through 720°

In a spin rotation experiment, the polarization vector $\boldsymbol{P}(\varphi)$ as a function of $\varphi = \omega_\mathrm{L} t$ comes back to its original value, $\boldsymbol{P}(2\pi) = \boldsymbol{P}(0)$ upon a full turn through $\omega_\mathrm{L} t = 2\pi$, as it should be. However, if we look at the time dependence of the spin-$\frac{1}{2}$ wave function $\psi(\varphi)$ in (3.57), we find something strange. Due to the factor $\frac{1}{2}$ in its argument, the wave function changes sign upon a full rotation $\psi(360°) = -\psi(0°)$,

$$\psi(0) = \begin{pmatrix} \cos(0) \\ -\mathrm{i}\sin(0) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \psi(2\pi) = \begin{pmatrix} \cos\pi \\ -\mathrm{i}\sin\pi \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix} = -\psi(0). \tag{4.12}$$

Only after two full rotations through 720° does the system come back to its original value, $\psi(4\pi) = \psi(0)$. The name spinor is reserved to wave functions with this property. In the early days of quantum physics, people reassured themselves that in the experimentally accessible probabilities $|\psi|^2 = \psi^\dagger \psi$ the phase factor $\mathrm{e}^{\pm \mathrm{i}\varphi}$ cancels, hence there seemed to be no reason to worry about this strange property of spinors.

This changed when neutron interference experiments for the first time measured the sign change of ψ (Rauch et al. 1975; Werner et al. 1975).

Neutrons for scientific use are produced either in a research reactor, where each *nuclear fission* event leads to the emission of two or three fast neutrons, or in a *nuclear spallation*

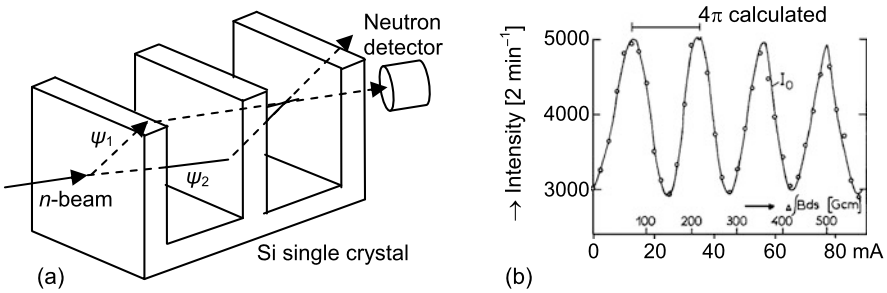


Fig. 4.5 Spinor spin rotation: (**a**) A neutron interferometer consists of three parallel plane slabs of single crystals of silicon. On each slab, the Bragg-reflected and the unreflected beams (*horizontal lines*) both have the same amplitude $\psi_2 = \psi_1$. In the third crystal, the amplitudes add coherently and interfere constructively or destructively, depending on the phase difference between the two neutron-optical paths. (**b**) If a magnetic field is switched on along one of the two paths, then the interference signal changes sign after one full rotation of the neutron spin in the field, and is back to the original intensity only after two full rotations, as shown *on the right of the figure*. Adapted from Rauch et al. (1975)

source, where high-energy protons are smashed onto heavy nuclei and produce about 30 neutrons per nuclear spallation event. The neutrons then are slowed down in a thermal or cold moderator, and leave the moderator through beam holes in the biological shield.
In a neutron interferometer, the two possible paths of the neutron can be spatially separated from each other by up to 10 cm. For stability the three crystal slabs are cut from one large single crystal of silicon. Thermal-neutron de Broglie wavelengths are $\lambda \sim 0.2$ nm, and the apparatus must be stable at this level while the neutron travels at rather low speed over macroscopic distances through the interferometer. Like in an optical interferometer, one can vary the relative phase δ of the two partial neutron beams by inserting a piece of matter into one neutron path, with a neutron-optical index of refraction n with $|n-1| \sim 10^{-3}$.

In neutron interferometry, Fig. 4.5, it is uncertain which path a neutron takes and therefore, as will be discussed in Sect. 6.6, the observed intensity is $|\psi|^2 = |\psi_1 + \psi_2|^2$, or

$$|\psi|^2 = |\psi_1|^2 + |\psi_2|^2 + 2|\psi_1||\psi_2| \sin\delta. \tag{4.13}$$

Here δ is the phase difference of the two partial de Broglie waves ψ_1 and ψ_2. The neutron intensity $I \propto |\psi|^2$ measured in the detector shown in Fig. 4.5(a) then is $I = I_1 + I_2 + 2(I_1 I_2)^{1/2} \sin\delta$, which for $I_1 = I_2 = \frac{1}{2} I_0$ becomes

$$I = I_0(1 + \sin\delta). \tag{4.14}$$

To measure the sign change of a spinor upon rotation through 2π, a transverse magnetic field of strength B is applied over a certain length of path 2 within the interferometer. The sign of the wave function ψ_2, as measured relative to the unperturbed ψ_1, should change when $\omega_L T = 2\pi$ for a neutron time-of-flight T through the field region. The measurements were done with a monochromatic beam of unpolarized neutrons, with results as shown in Fig 4.4(b). For each full rotation of neutron spin about B, the interference signal shows a change of sign, as predicted for a spinor wave function.

References

Bennett, G.W., et al.: Final report of the E821 muon anomalous magnetic moment measurement at BNL. Phys. Rev. D **73**, 072003(41) (2006)

Jegerlehner, F., Nyffeler, A.: The muon $g - 2$. Phys. Rep. **477**, 1–110 (2009)

Rauch, H., Zeilinger, A., Badurek, G., Wilfing, A., Bauspiess, W., Bonse, U.: Verification of coherent spinor rotation of fermions. Phys. Lett. **54A**, 425–427 (1975)

Werner, S.A., Colella, R., Overhauser, A.W., Eagen, C.F.: Observation of the phase shift of a neutron due to precession in a magnetic field. Phys. Rev. Lett. **35**, 1053–1055 (1975)

Chapter 5
General Solution for the Two-Level System

Abstract The essential steps of a quantum calculation are: the establishment of the matrix elements of the Hamiltonian, the derivation of its eigenvalues and eigenvectors, and of the time evolution of the observables. We slowly go through every step to arrive at the general response of the two-state quantum system to stationary fields.

In Sect. 3.5, we solved the Schrödinger equation for the simple case of a magnetic field $\boldsymbol{B}$ along z, with $\mathrm{H} \propto \sigma_z$. If $\boldsymbol{B}$ points towards an arbitrary direction, we must diagonalize the Hamiltonian to find the time dependent solution of the Schrödinger equation. The reader might wonder why we do not always choose the z-axis along $\boldsymbol{B}$ in order to avoid this complication. The answer is that, in general, the diagonal and the off-diagonal elements of the Hamiltonian need not belong to the same type of interaction, and we want to be able to vary them independently. We give some examples to be met in later chapters:

- The diagonal elements of the Hamiltonian have a static but externally variable field B_z, while the off-diagonal elements are due to a fixed crystal-magnetic field $B_\perp$.
- A constant field B_z acts again on the diagonal, but the off-diagonal elements are due to a radiofrequency field $B_\perp(t)$ that induces resonance transitions.
- A constant externally variable electric field E_z that acts on a polar molecule enters the diagonal, while the off-diagonal element is a transition amplitude between two configurations of the molecule.
- Two energies E_1 and E_2 of an atom's electronic level scheme are on the diagonal, while the electric-field component $E_\perp(t)$ of a light field that drives a laser-induced transition enters off the diagonal.
- For spin $s > \frac{1}{2}$, there may be both a magnetic dipole and an electric quadrupole interaction contributing to the off-diagonal components.

Hence there is ample reason to extend our study to non-diagonal Hamiltonians. Furthermore, our general treatment of the 2×2 case can serve as a model for the treatment of higher-dimensional Hamiltonians.

D. Dubbers, H.-J. Stöckmann, *Quantum Physics: The Bottom-Up Approach*, Graduate Texts in Physics,
DOI 10.1007/978-3-642-31060-7_5, © Springer-Verlag Berlin Heidelberg 2013

5.1 Matrix Diagonalization

As before, without loss of generality, we use magnetic field components as matrix elements. In the angular momentum basis, the most general 2×2 Schrödinger equation with H from (3.16c) and $B_\pm = B_x \pm \mathrm{i}B_y$ is

$$\dot{\psi} = -\frac{\mathrm{i}}{\hbar}\mathrm{H}\psi, \tag{5.1a}$$

$$\begin{pmatrix} \dot{a} \\ \dot{b} \end{pmatrix} = \frac{\mathrm{i}}{\hbar}\mu \begin{pmatrix} B_z & B_- \\ B_+ & -B_z \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}, \tag{5.1b}$$

$$\dot{a} = \frac{\mathrm{i}\mu}{\hbar}(B_z a + B_- b), \tag{5.1c}$$

$$\dot{b} = \frac{\mathrm{i}\mu}{\hbar}(B_+ a - B_z b). \tag{5.1d}$$

To decouple the two equations, we must convert H into a diagonal matrix $\mathrm{H_D}$, with eigenvalues $E_\pm = \pm\frac{1}{2}\hbar\omega_\mathrm{L}$ from (3.22). The subsequent treatment holds for Hamiltonians of any dimension. We first must find the energy eigenfunction ψ_D in terms of a given spin eigenfunction ψ. Both ψ and ψ_D are complex and normalized to one, therefore they are connected by a *unitary transformation* T

$$\psi = \mathrm{T}\psi_\mathrm{D}. \tag{5.2}$$

Unitarity requires that the inverse of the matrix T equals its conjugate transpose, $\mathrm{T}^{-1} = \mathrm{T}^\dagger$, or

$$\mathrm{T}\mathrm{T}^\dagger = 1, \quad \text{with } \big|\det(\mathrm{T})\big| = 1. \tag{5.3}$$

Operation T induces a complex rotation of the state ψ_D in Hilbert space. Multiplication of (5.2) from the left by $\mathrm{T}^{-1} = \mathrm{T}^\dagger$ gives

$$\psi_\mathrm{D} = \mathrm{T}^\dagger\psi. \tag{5.4}$$

Multiplication of the Schrödinger equation (1.3) by $\mathrm{T}^\dagger$ from the left, $\mathrm{T}^\dagger\dot{\psi} = -(\mathrm{i}/\hbar)\mathrm{T}^\dagger\mathrm{H}\psi$, and squeezing in $\mathrm{T}\mathrm{T}^\dagger = 1$ on the right-hand side gives

$$\mathrm{T}^\dagger\dot{\psi} = -(\mathrm{i}/\hbar)\big(\mathrm{T}^\dagger\mathrm{H}\mathrm{T}\big)\mathrm{T}^\dagger\psi. \tag{5.5}$$

We use the energy eigenfunction $\psi_\mathrm{D} = \mathrm{T}^\dagger\psi$ and the diagonalized energy matrix

$$\mathrm{H_D} = \mathrm{T}^\dagger\mathrm{H}\mathrm{T}, \tag{5.6}$$

and obtain, once we know the transformation T, a decoupled set of differential equations for the energy eigenstate ψ_D,

$$\dot{\psi}_\mathrm{D} = -\frac{\mathrm{i}}{\hbar}\mathrm{H_D}\psi_\mathrm{D}. \tag{5.7}$$

We are now in a position to verify our earlier statement that expectation values (i.e., measurement results) do not depend on representation. This is an immediate consequence of the definition of expectation values, (3.28), together with the above transformation laws, namely, $\psi_{\mathrm{D}} = \mathrm{T}^{\dagger}\psi$, its conjugate transpose $\psi_{\mathrm{D}}^{\dagger} = \psi^{\dagger}\mathrm{T}$, the diagonal energy matrix $\mathrm{H}_{\mathrm{D}} = \mathrm{T}^{\dagger}\mathrm{H}\mathrm{T}$, and unitarity $\mathrm{T}\mathrm{T}^{\dagger} = 1$, which together give

$$\langle \mathrm{H}_{\mathrm{D}} \rangle = \psi_{\mathrm{D}}^{\dagger}\mathrm{H}_{\mathrm{D}}\psi_{\mathrm{D}} = \psi^{\dagger}\mathrm{T}\big(\mathrm{T}^{\dagger}\mathrm{H}\mathrm{T}\big)\mathrm{T}^{\dagger}\psi = \psi^{\dagger}\mathrm{H}\psi = \langle \mathrm{H} \rangle. \tag{5.8}$$

5.2 Construction of the Eigenvectors

Our task then is to find the unitary matrix T that diagonalizes the Hamiltonian H. Equation (5.6) can be written as $\mathrm{H}\mathrm{T} = \mathrm{T}\mathrm{H}_{\mathrm{D}}$, which, for the nth column t_n of T reads $\mathrm{H}\mathrm{t}_n = E_n \mathrm{t}_n$. Hence t_n is the nth eigenstate (i.e., eigenvector) of H, with eigenvalue E_n:

> The columns of the matrix T that diagonalizes the Hamiltonian H are the energy eigenstates of H.

Thus we need to find the eigenvectors of H that build the transformation matrix T. We do this for the case $s = \frac{1}{2}$. The most general 2×2 unitary matrix is, up to an arbitrary phase factor in front of each column,

$$\mathrm{T} = \begin{pmatrix} \mathrm{e}^{-\mathrm{i}\eta}\cos\chi & -\mathrm{e}^{-\mathrm{i}\eta}\sin\chi \\ \mathrm{e}^{\mathrm{i}\eta}\sin\chi & \mathrm{e}^{\mathrm{i}\eta}\cos\chi \end{pmatrix}, \tag{5.9}$$

with phase angles χ and η, and we readily check that $\mathrm{T}\mathrm{T}^{\dagger} = 1$ and $\det \mathrm{T} = 1$ as required for our unitary T. The columns of T are normalized wave functions of type (3.32b). With the abbreviation $c = \mathrm{e}^{\mathrm{i}\eta}\cos\chi$ and $s = \mathrm{e}^{-\mathrm{i}\eta}\sin\chi$, we calculate the elements of T by writing (5.6) with H from (3.16c) as

$$\mathrm{H} = \mathrm{T}\mathrm{H}_{\mathrm{D}}\mathrm{T}^{\dagger}, \tag{5.10a}$$

$$\begin{aligned} -\mu \begin{pmatrix} B_z & B_- \\ B_+ & -B_z \end{pmatrix} &= -\mu \begin{pmatrix} c^* & -s \\ s^* & c \end{pmatrix} \begin{pmatrix} B & 0 \\ 0 & -B \end{pmatrix} \begin{pmatrix} c & s \\ -s^* & c^* \end{pmatrix} \\ &= -\mu B \begin{pmatrix} |c|^2 - |s|^2 & 2c^*s \\ 2cs^* & |s|^2 - |c|^2 \end{pmatrix}, \end{aligned} \tag{5.10b}$$

$$-\mu \begin{pmatrix} B_z & B_- \\ B_+ & -B_z \end{pmatrix} = -\mu B \begin{pmatrix} \cos 2\chi & \mathrm{e}^{-2\mathrm{i}\eta}\sin 2\chi \\ \mathrm{e}^{2\mathrm{i}\eta}\sin 2\chi & -\cos 2\chi \end{pmatrix}, \tag{5.10c}$$

where we used (3.34a). The upper left entry in the matrix equation (5.10c) gives $B_z/B = \cos 2\chi$. The real and imaginary parts of the upper right entries give $B_x/B = \cos 2\eta \sin 2\chi$ and $B_y/B = \sin 2\eta \sin 2\chi$, and their ratio is $B_y/B_x = \tan 2\eta$.

In polar coordinates we have, on the other hand,

$$B_z/B = \cos\theta, \qquad B_y/B_x = \tan\varphi; \tag{5.11}$$

see also Fig. 5.1(a) in the next section. Therefore, as in (3.37), we again arrive at $2\chi = \theta$, $2\eta = \varphi$. This result is comforting but not really surprising as we had already stated that, for $s = \frac{1}{2}$, a rotation in Hilbert space proceeds at half speed, as compared to a rotation in ordinary vector space.

Hence the transformation matrix T that diagonalizes the Hamiltonian H reads

$$\mathrm{T} = \begin{pmatrix} \mathrm{e}^{-\mathrm{i}\varphi/2}\cos\frac{1}{2}\theta & -\mathrm{e}^{-\mathrm{i}\varphi/2}\sin\frac{1}{2}\theta \\ \mathrm{e}^{\mathrm{i}\varphi/2}\sin\frac{1}{2}\theta & \mathrm{e}^{\mathrm{i}\varphi/2}\cos\frac{1}{2}\theta \end{pmatrix}, \tag{5.12}$$

and the abbreviations used in (5.10b) become

$$c = \mathrm{e}^{\mathrm{i}\varphi/2}\cos\frac{1}{2}\theta, \qquad s = \mathrm{e}^{-\mathrm{i}\varphi/2}\sin\frac{1}{2}\theta. \tag{5.13}$$

In terms of the field components, we have, using (5.11) and (3.34b),

$$\cos\frac{1}{2}\theta = \sqrt{\frac{1}{2}(1 + B_z/B)}, \tag{5.14a}$$

$$\sin\frac{1}{2}\theta = \sqrt{\frac{1}{2}(1 - B_z/B)}. \tag{5.14b}$$

In an energy eigenstate, the probabilities of being spin up or spin down are

$$p_{\mathrm{up}} = \cos^2\frac{1}{2}\theta = \frac{1}{2}(1 + B_z/B), \tag{5.15a}$$

$$p_{\mathrm{down}} = \sin^2\frac{1}{2}\theta = \frac{1}{2}(1 - B_z/B). \tag{5.15b}$$

From $\cos\varphi = B_x/B_\perp$, $\sin\varphi = B_y/B_\perp$ with $B_\perp = (B_x^2 + B_y^2)^{1/2}$ follows from (3.34d) and (5.14a), (5.14b)

$$\mathrm{e}^{\mathrm{i}\varphi/2} = \sqrt{\frac{1}{2}(1 + B_x/B_\perp)} + \mathrm{i}\sqrt{\frac{1}{2}(1 - B_x/B_\perp)}. \tag{5.16}$$

5.3 The Time Dependent Solution

5.3.1 Evolution of an Energy Eigenstate

With the transformation matrix T at hand, we can write down the general time dependent solution. With diagonal N-dimensional Hamiltonian $\mathrm{H_D}$, the Schrödinger equation consists of N decoupled equations $\dot{\psi}_{\mathrm{D}n} = -(\mathrm{i}/\hbar)E_n\psi_{\mathrm{D}n}$, with solutions $\psi_{\mathrm{D}n}(t) = \psi_{\mathrm{D}n}(0)\exp(-\mathrm{i}E_n t/\hbar)$, $n = 1, \ldots, N$, formally written in column vector notation as

$$\psi_{\mathrm{D}}(t) = \mathrm{e}^{-\mathrm{i}\mathrm{H_D}t/\hbar}\psi_{\mathrm{D}}(0), \tag{5.17}$$

where the exponential expression is shorthand for a diagonal matrix with elements $\mathrm{e}^{-\mathrm{i}E_n t/\hbar}$.

We begin with an $s = \frac{1}{2}$ system, which initially is in a pure energy eigenstate $\psi_{\mathrm{D}}(0)$, for instance, in the upper energy level $\psi_{\mathrm{D}}(0) = \alpha$, a case frequently met in experimental physics, see the ammonia maser, Sect. 12.4.2. With (5.2), (5.17), and with T from (5.12), the spin state ψ evolves as

$$\psi(t) = \mathrm{T}\psi_{\mathrm{D}}(t) = \mathrm{T}\mathrm{e}^{-\mathrm{i}\mathrm{H_D}t/\hbar}\psi_{\mathrm{D}}(0), \tag{5.18a}$$

$$\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \begin{pmatrix} c^* & -s \\ s^* & c \end{pmatrix} \begin{pmatrix} \mathrm{e}^{-\mathrm{i}E_+ t/\hbar} & 0 \\ 0 & \mathrm{e}^{-\mathrm{i}E_- t/\hbar} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} c^* \\ s^* \end{pmatrix} \mathrm{e}^{-\mathrm{i}E_+ t/\hbar}, \tag{5.18b}$$

$$a(t) = \mathrm{e}^{-\mathrm{i}\varphi/2}\cos\frac{1}{2}\theta\mathrm{e}^{-\mathrm{i}\omega_{\mathrm{L}}t/2}, \tag{5.18c}$$

$$b(t) = \mathrm{e}^{\mathrm{i}\varphi/2}\sin\frac{1}{2}\theta\mathrm{e}^{-\mathrm{i}\omega_{\mathrm{L}}t/2}, \tag{5.18d}$$

where we used $E_\pm = \pm\frac{1}{2}\hbar\omega_{\mathrm{L}}$ from (2.16), and c and s from (5.13).

We are not that interested in $\psi(t)$ but rather in the expectation values of the Pauli matrices $\langle\boldsymbol{\sigma}\rangle$ from (3.31a)–(3.31c), which become

$$\langle\sigma_x\rangle = 2\,\mathrm{Re}\big(a^* b\big) = \cos\varphi\sin\theta = B_x/B, \tag{5.19a}$$

$$\langle\sigma_y\rangle = 2\,\mathrm{Im}\big(a^* b\big) = \sin\varphi\sin\theta = B_y/B, \tag{5.19b}$$

$$\langle\sigma_z\rangle = |a|^2 - |b|^2 = \cos\theta = B_z/B, \tag{5.19c}$$

$$\boldsymbol{P} = \langle\boldsymbol{\sigma}\rangle = \boldsymbol{B}/B. \tag{5.20}$$

The polarization vector $\boldsymbol{P}$ of an energy eigenstate is simply oriented along the field $\boldsymbol{B}$, and no spin precession takes place.

This agrees with what we knew before because (3.52) gave the result $P_z(t) = P_z(0)$ for $B = B_z$, which should hold for any choice of the z-axis.

5.3.2 Evolution of an Angular-Momentum Eigenstate

Next we look at a system initially in a state $\psi(0) = a(0)\alpha + b(0)\beta$. The time dependence of the spin amplitude then is $\psi(t) = \mathrm{T}\psi_\mathrm{D}(t)$, with T from (5.12), $\psi_\mathrm{D}(t)$ from (5.17), and $\psi_\mathrm{D}(0)$ from (5.4).

$$\psi(t) = \mathrm{T}\mathrm{e}^{-\mathrm{i}H_\mathrm{D}t/\hbar}\mathrm{T}^\dagger\psi(0), \tag{5.21a}$$

$$\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \begin{pmatrix} c^* & -s \\ s^* & c \end{pmatrix} \begin{pmatrix} \mathrm{e}^{-\mathrm{i}E_+t/\hbar} & 0 \\ 0 & \mathrm{e}^{-\mathrm{i}E_-t/\hbar} \end{pmatrix} \begin{pmatrix} c & s \\ -s^* & c^* \end{pmatrix} \begin{pmatrix} a(0) \\ b(0) \end{pmatrix}, \tag{5.21b}$$

$$a(t) = U_{11}(t)a(0) + U_{12}(t)b(0), \tag{5.21c}$$

$$b(t) = U_{21}(t)a(0) + U_{22}(t)b(0). \tag{5.21d}$$

The time evolution coefficients therein are

$$U_{11}(t) = \cos\left(\frac{1}{2}\omega_\mathrm{L}t\right) - \mathrm{i}\cos\theta\,\sin\left(\frac{1}{2}\omega_\mathrm{L}t\right), \tag{5.22a}$$

$$U_{12}(t) = -\mathrm{i}\mathrm{e}^{-\mathrm{i}\varphi}\sin\theta\,\sin\left(\frac{1}{2}\omega_\mathrm{L}t\right), \tag{5.22b}$$

and $U_{22}(t) = U_{11}^*(t)$, $U_{21}(t) = -U_{12}^*(t)$. The U_{ik} are the elements of the *time evolution operator* $U(t)$ that will be discussed in Sect. 6.1. With these elements, the general evolution of the polarization vector $\boldsymbol{P}(t) = \langle\boldsymbol{\sigma}\rangle(t)$ can be derived, as we did for special cases in (3.52) and (3.60).

We here limit discussion to the general evolution of the longitudinal polarization component $P_z(t)$. If, for example, the atom starts in a state spin up, with the initial condition $a(0) = 1$, $b(0) = 0$, the amplitude that the atom will be in the state spin down after a time t is

$$b(t) = -U_{21}(t)a(0) = -\mathrm{i}\mathrm{e}^{-\mathrm{i}\varphi}\sin\theta\,\sin\left(\frac{1}{2}\omega_\mathrm{L}t\right). \tag{5.23}$$

The probability that the atom is flipped into its spin down state then oscillates as

$$p_\mathrm{down}(t) = \left|b(t)\right|^2 = \frac{1}{2}\sin^2\theta(1 - \cos\omega_\mathrm{L}t), \tag{5.24}$$

where we used (3.34b). The probability of being in the spin up state is $p_\mathrm{up}(t) = 1 - p_\mathrm{down}(t)$. From this follows for $P_z = p_\mathrm{up} - p_\mathrm{down} = 1 - 2p_\mathrm{down}$:

The general time evolution of the longitudinal polarization is

$$P_z(t) = \cos^2\theta + \sin^2\theta\cos\omega_\mathrm{L}t. \tag{5.25}$$

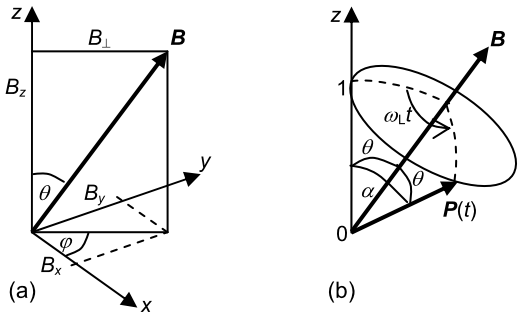


Fig. 5.1 (**a**) Components of the magnetic field vector, and its polar and azimuthal angles θ and φ. (**b**) Spin rotation of atomic polarization $\boldsymbol{P}(t)$ about $\boldsymbol{B}$: For a spherical triangle, a corner angle δ is related to the angles α, β, γ at the origin of the unit sphere by the spherical law of cosines $\cos\alpha = \cos\beta\sin\gamma + \sin\beta\sin\gamma\cos\delta$. In our case, with angle α between $\boldsymbol{P}$ and z, and with equal angles $\beta = \gamma = \theta$ between $\boldsymbol{B}$ and z and between $\boldsymbol{B}$ and $\boldsymbol{P}$, and with phase angle $\delta = \omega_{\mathrm{L}}t$, (5.25) is obtained

We immediately find the same result in the spin precession picture. We set $P_z(t) = P_{z0}\cos\alpha$, with $P_{z0} = 1$, and with the momentary angle α between $\boldsymbol{P}(t)$ and the z-axis, Indeed, the *spherical law of cosines* states that the three angles α, θ, and $\omega_{\mathrm{L}}t$ are related just as given by (5.25), see Fig. 5.1(b).

Instead of writing the equations above in terms of $\theta, \varphi, \omega_{\mathrm{L}}$, we can as well write them in terms of field amplitudes B_x, B_y, B_z, B, or in terms of the general matrix elements H_{ik}. The relations between these quantities are listed for the convenience of the reader:

$$\hbar\omega_{\mathrm{L}} = -2\mu B = -\left[(H_{11} - H_{22})^2 + 4|H_{12}|^2\right]^{1/2}, \tag{5.26a}$$

$$\cos\theta = B_z/B = (H_{11} - H_{22})/\left[(H_{11} - H_{22})^2 + 4|H_{12}|^2\right]^{1/2}, \tag{5.26b}$$

$$\sin\theta = \left(B_x^2 + B_y^2\right)^{1/2}/B = 2|H_{12}|/\left[(H_{11} - H_{22})^2 + 4|H_{12}|^2\right]^{1/2}, \tag{5.26c}$$

$$\mathrm{e}^{\mathrm{i}\varphi} = (B_x + \mathrm{i}B_y)/\left(B_x^2 + B_y^2\right)^{1/2} = H_{12}^*/|H_{12}|. \tag{5.26d}$$

Expressed in terms of magnetic field components, $P_z(t)$ from (5.25) reads

$$P_z(t) = \frac{B_z^2}{B^2} + \frac{B_x^2 + B_y^2}{B^2}\cos(2\mu Bt/\hbar). \tag{5.27}$$

Figure 5.2 shows $P_z(t)$ for $B_y = 0$ and various values of B_z/B_x. For $B_z = 0$ we have the full spin precession signal. If B_z is much larger than B_x, then only small but fast *spin nutations* are visible, whose amplitudes decrease quadratically with B, while the oscillation frequency increases linearly with B.

For time intervals Δt so short that the interaction is blurred by time–energy uncertainty, all solutions of the Schrödinger equation start out quadratically in time, which is known to be true even for exponential decay of a quantum state. This general law here is obeyed by the cosine in (5.27), for short times t such that $2|\mu|Bt/\hbar \ll 1$, see the dashed curve in the figure.

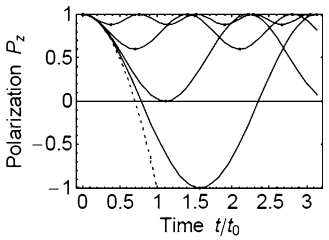


Fig. 5.2 Sensitivity of a spin rotation signal about a transverse field to "unwanted" longitudinal field components: Let the full signal be given by $P_z(t)$, with spin rotation about B_x. Additional field components along B_z diminish the signal, as shown for $B_z = 0$, $B_z = B_x$, $B_z = 2B_x$, and $B_z = 4B_x$. Time is given in units of the precession period $t_0 = 1/|\gamma B_x|$ about B_x. The *dashed curve* shows the universal quantum short-time behavior $P_z \propto t^2$

To summarize, any quantum calculation with time independent interactions proceeds in the same way. After selection of a Hamiltonian operator H ("building a model"), one writes down its matrix elements in an adequate representation. Diagonalization of the Hamiltonian matrix then gives the matrix T of energy eigenvectors and the diagonal matrix $\mathrm{H_D}$ of energy eigenvalues. The initial condition $\psi(0)$ in the diagonal representation is $\psi_\mathrm{D}(0) = \mathrm{T}^\dagger\psi(0)$. The time dependent solution in the diagonal representation $\psi_\mathrm{D}(t) = \exp(-\mathrm{i}\mathrm{H_D}t/\hbar)\psi_\mathrm{D}(0)$ then is transformed back to the chosen representation to give (5.21a), which describes the diagonalization procedure for general $N \times N$ Hamiltonians.

Chapter 6
Other Tools and Concepts

Abstract Quantum theory abounds with matrices of different kinds, and often it is not easy to keep track of all this. In this chapter, we introduce some further operators and their matrices, namely the time evolution, rotation, and projection operators. After a digression on pure and mixed quantum states, the density matrix is introduced in an elementary manner and used to derive again some previous results. It is followed by a section on coherences in quantum physics, and a discussion on Dirac's bra-ket notation.

6.1 Time Evolution Operator

Often the general time evolution of a state vector ψ is simply written as

$$\psi(t) = \mathrm{U}(t)\psi(0). \tag{6.1}$$

For spin $s = \frac{1}{2}$, we know the time evolution operator $\mathrm{U}(t)$ from (5.22a), (5.22b) in terms of the angular variables $\theta, \varphi, \omega_\mathrm{L}$. Using (5.26b)–(5.26d), we write $\mathrm{U}(t)$ in terms of the $\boldsymbol{B}$-field components in the Hamiltonian as

$$\mathrm{U}(t) = \begin{pmatrix} \cos\left(\frac{1}{2}\omega_\mathrm{L}t\right) - \mathrm{i}(B_z/B)\sin\left(\frac{1}{2}\omega_\mathrm{L}t\right) & -\mathrm{i}(B_-/B)\sin\left(\frac{1}{2}\omega_\mathrm{L}t\right) \\ \mathrm{i}(B_+/B)\sin\left(\frac{1}{2}\omega_\mathrm{L}t\right) & \cos\left(\frac{1}{2}\omega_\mathrm{L}t\right) + \mathrm{i}(B_z/B)\sin\left(\frac{1}{2}\omega_\mathrm{L}t\right) \end{pmatrix}. \tag{6.2}$$

In (5.17), with diagonal $\mathrm{H_D}$, we wrote the energy eigenstate as $\psi_\mathrm{D}(t) = \exp(-\mathrm{i}\mathrm{H_D}t/\hbar)\psi_\mathrm{D}(0)$. The same expression holds for any time independent Hamiltonian H, independent of a representation.

The time evolution operator is

$$\mathrm{U}(t) = \mathrm{e}^{-\mathrm{i}\mathrm{H}t/\hbar}. \tag{6.3}$$

In other words, when going from (5.17) to (6.1), the unitary transformation of ψ can be directly applied to the Hamiltonian in the exponent, as will be shown in the following for $s = \frac{1}{2}$.

D. Dubbers, H.-J. Stöckmann, *Quantum Physics: The Bottom-Up Approach*, Graduate Texts in Physics,
DOI 10.1007/978-3-642-31060-7_6, © Springer-Verlag Berlin Heidelberg 2013

To prove this statement for the two-dimensional $\mathrm{H} = -\mu\boldsymbol{\sigma}\cdot\boldsymbol{B}$, we write $\mathrm{H}^2 = (\mu B)^2\mathrm{I}$, using $\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = \mathrm{I}$ and $\sigma_x\sigma_y + \sigma_y\sigma_x = 0$, etc. (with null matrix 0). This allows us to expand $\mathrm{U}(t)$ as

$$\begin{aligned} \mathrm{e}^{\mathrm{i}\mu\boldsymbol{\sigma}\cdot\boldsymbol{B}t/\hbar} &= \big[1 - (\mu Bt/\hbar)^2/2! + (\mu Bt/\hbar)^4/4! - \cdots\big]\mathrm{I} \\ &\quad - \mathrm{i}\big[(\mu Bt/\hbar)/1! - (\mu Bt/\hbar)^3/3! + \cdots\big]\boldsymbol{\sigma}\cdot\boldsymbol{B}/B \\ &= \mathrm{I}\cos(\mu Bt/\hbar) - \mathrm{i}(\boldsymbol{\sigma}\cdot\boldsymbol{B})/B\sin(\mu Bt/\hbar), \end{aligned} \tag{6.4}$$

or

$$\mathrm{e}^{-\mathrm{i}\mathrm{H}t/\hbar} = \mathrm{I}\cos(Et/\hbar) - \mathrm{i}(\mathrm{H}/E)\sin(Et/\hbar), \tag{6.5}$$

with matrices H and I. This is nothing but Euler's formula $\mathrm{e}^{\mathrm{i}x} = \cos x + \mathrm{i}\sin x$, with x replaced by the operator $-\mathrm{i}\mathrm{H}t/\hbar$. Therefore, a transformation $\mathrm{T}^\dagger\boldsymbol{\sigma}\mathrm{T}$ of the matrix $\boldsymbol{\sigma}$ on the right of (6.4) directly transforms $\boldsymbol{\sigma}$ in the exponent on the left. This proves (6.3) for $s = \frac{1}{2}$. This also proves the unitarity of the evolution operator $\mathrm{U}(t)$, as we know that for any hermitian operator A (here $\mathrm{A} = -\mathrm{H}t/\hbar$ with hermitian H), the exponential operator $\mathrm{U} = \mathrm{e}^{\mathrm{i}\mathrm{A}}$ is unitary.

6.2 Rotation Matrices

In classical physics, a rotation of the coordinate system in a three-dimensional space changes a classical vector $\boldsymbol{v}$ to $\boldsymbol{v}' = \mathrm{R}\boldsymbol{v}$ with a 3×3 rotation matrix R. For instance, a rotation about the x-axis by an angle θ produces the new vector

$$\boldsymbol{v}' = \begin{pmatrix} v_x' \\ v_y' \\ v_z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} = \mathrm{R}_x(\theta)\boldsymbol{v}. \tag{6.6}$$

An example is the polarization vector $\boldsymbol{P} = \langle\boldsymbol{\sigma}\rangle$ of (3.36), which transforms under rotation like any other classical vector. Successive rotations are known to depend on the order of their execution, that is, rotation operations in general do not *commute*. You can check this by rotating the book you hold in your hands successively about two different axes, and the outcome will be different if you change the sequence of operations.

What happens if, instead of rotating in ordinary space, we rotate the coordinates of Hilbert space in which the quantum operators and states vectors are defined? In the general case, this is a tricky question that will be addressed in Sect. 16.5. For our simple spin-$\frac{1}{2}$ case, however, we know already how a 2×2 rotation matrix in Hilbert space must look like. The reason is that, in the two-dimensional case (and only there), any unitary operator describes a rotation in Hilbert space and as such is a rotation operator. Hence for effective $s = \frac{1}{2}$, any interaction, and in particular any magnetic dipole interaction, leads to a rotation of the system and to nothing else.

For example, let us perform a rotation of a spin-$\frac{1}{2}$ state through an angle φ about the z-axis. From the time evolution operator (6.2), setting $B_+ = B_- = 0$, $B = B_z$, $\omega_L t = \varphi$, and using Euler's formula, we find the rotated state $\psi' = \mathrm{d}_z^{(1/2)}\psi$, with the rotation matrix

$$\mathrm{d}_z^{(1/2)}(\varphi) = \begin{pmatrix} \mathrm{e}^{-\mathrm{i}\varphi/2} & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\varphi/2} \end{pmatrix}. \tag{6.7}$$

For a rotation through an angle θ about the y-axis, we set in (6.2) $B_y = B_z = 0$ and obtain the rotation matrix

$$\mathrm{d}_y^{(1/2)}(\theta) = \begin{pmatrix} \cos\frac{1}{2}\theta & -\sin\frac{1}{2}\theta \\ \sin\frac{1}{2}\theta & \cos\frac{1}{2}\theta \end{pmatrix}. \tag{6.8}$$

Rotations of a state vector ψ in Hilbert space are achieved by rotation matrices, as given in Appendix A.5.

The general 2×2 unitary transformation matrix T from (5.12) is

$$\mathrm{T}(\theta, \varphi) = \mathrm{d}_z^{(1/2)}(-\varphi)\mathrm{d}_y^{(1/2)}(\theta). \tag{6.9}$$

6.3 Projection Operators

In Sect. 3.3, we pointed out that the concept of expectation values must be complemented by adequate projection operations in order to include all types of spectroscopic or polarization measurements. Equations (3.30a), (3.30b) gave an example of a projection operation acting on an energy eigenstate, and we want to expand on this.

Let a beam of atoms, flying along $+y$ in an angular momentum state $\psi = a\alpha + b\beta$, enter a Stern–Gerlach magnet oriented along quantization axis z. The apparatus splits the incoming beam into two partial beams $\psi_{z+} = a\alpha$ and $\psi_{z-} = b\beta$. These states are obtained by projection operations

$$\psi_{z+} = \mathrm{P}_{z+}\psi = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a \\ 0 \end{pmatrix}, \tag{6.10a}$$

$$\psi_{z-} = \mathrm{P}_{z-}\psi = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ b \end{pmatrix}, \tag{6.10b}$$

as given in (3.30a), (3.30b), because in this simple case, energy and angular momentum eigenstates coincide. Their complex conjugates are $\psi_{z+}^\dagger = \psi^\dagger \mathrm{P}_{z+} = (a^* \quad 0)$ and $\psi_{z-}^\dagger = \psi^\dagger \mathrm{P}_{z-} = (0 \quad b^*)$.

The probabilities of finding an atom in one or the other partial beam of the Stern–Gerlach device then are the usual

$$p_{\text{up}} = \psi_{z+}^{\dagger}\psi_{z+} = |a|^2, \qquad p_{\text{down}} = \psi_{z-}^{\dagger}\psi_{z-} = |b|^2. \tag{6.11}$$

The projection operators $\mathrm{P}_{z\pm}$ are written in terms of the Pauli matrix σ_z as

$$\mathrm{P}_{z+} = \frac{1}{2}(\mathrm{I} + \sigma_z), \qquad \mathrm{P}_{z-} = \frac{1}{2}(\mathrm{I} - \sigma_z). \tag{6.12}$$

If, in the course of a measurement, ψ is projected onto an axis in Hilbert space, this process is called the *collapse of the wave function*. For a chosen quantization axis z, only collapse onto $+z$ or $-z$ is possible, hence from (6.12) we have the *closure relation*

$$\mathrm{P}_{z+} + \mathrm{P}_{z-} = \mathrm{I}. \tag{6.13}$$

Projection operators are *idempotent*, $\mathrm{P}^2 = \mathrm{P}$, because a projection is made only once, a second projection being the identity operation I.

Next we look at a Stern–Gerlach magnet oriented along x, while keeping the quantization axis along z. The x- and the z-axes are equivalent, therefore we can substitute x for z in (6.12), and obtain for the projections of ψ onto the x-axis the operators

$$\psi_{x\pm} = \frac{1}{2}(\mathrm{I} \pm \sigma_x)\psi, \tag{6.14a}$$

$$\psi_{x\pm} = \frac{1}{2}\begin{pmatrix} 1 & \pm 1 \\ \pm 1 & 1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \frac{1}{2}\begin{pmatrix} a \pm b \\ \pm a + b \end{pmatrix}. \tag{6.14b}$$

Let the incoming beam be fully polarized "spin up", $\psi = \alpha$, that is, $a = 1$, $b = 0$, with respect to the quantization axis z. After passage of the Stern–Gerlach device oriented along x, the outgoing partial beams have spin up or spin down probabilities, with respect to x, given by the scalar product

$$|\psi_{x\pm}|^2 = \frac{1}{2}(1 \quad \pm 1)\frac{1}{2}\begin{pmatrix} 1 \\ \pm 1 \end{pmatrix} = \frac{1}{2}, \tag{6.15a}$$

$$p_{\text{up}} = p_{\text{down}} = \frac{1}{2}. \tag{6.15b}$$

A measurement on a 100 % pure spin up state with a Stern–Gerlach magnet, oriented at a right angle to the polarization of the incoming beam, results in a state that is 50 % spin up and 50 % spin down, with respect to the Stern–Gerlach axis.

In ordinary space, the projection of the polarization vector $\boldsymbol{P} = \langle \boldsymbol{\sigma} \rangle$ along z onto a unit vector $\boldsymbol{n}$ along z', with angle θ between both vectors, is

$$P_{z'} = \boldsymbol{P} \cdot \boldsymbol{n} = P_z \cos\theta. \tag{6.16}$$

A Stern–Gerlach filter performs exactly this projection. In Sect. 5.2, we had investigated the case of a spin polarization $\boldsymbol{P}$ along the z-axis, exposed to a magnetic field $\boldsymbol{B}$ along axis z' under angle θ, and found the probabilities (5.15a), (5.15b) to be spin up or spin down with respect to the z' axis $p_{\text{up}} = \cos^2 \frac{1}{2}\theta$, $p_{\text{down}} = \sin^2 \frac{1}{2}\theta$. With no surprise we find that $P_{z'} = p_{\text{up}} - p_{\text{down}}$ via (3.34a) coincides with (6.16).

In the general case, the projection of a wavefunction ψ onto an axis defined by a unit vector $\boldsymbol{n}$ is accomplished by the projection operator

$$P_{\boldsymbol{n}} = \frac{1}{2}(\mathrm{I} + \boldsymbol{\sigma} \cdot \boldsymbol{n}). \tag{6.17}$$

What is the practical use of projection operators? In modern high-resolution spectroscopy, observables A usually are not measured as the average $\langle \mathrm{A} \rangle = \psi^\dagger \mathrm{A} \psi$ over all possible eigenstates of A. Instead,

If a spectrometer filters out a subset of states, then a projection of ψ onto this subset of states $\psi' = \mathrm{P}\psi$ is required to find the *projected expectation values*

$$\langle \mathrm{A} \rangle = \frac{\psi'^\dagger \mathrm{A} \psi'}{\psi'^\dagger \psi'}, \quad \psi' = \mathrm{P}\psi. \tag{6.18}$$

For a sufficiently high resolution, expectation values are reduced to eigenvalues and projection operations are not needed.

6.4 Pure States and Mixed States

We noticed in Sect. 3.3.2 that the length of a polarization vector $\boldsymbol{P} = \langle \boldsymbol{\sigma} \rangle$ is always unity, $|\boldsymbol{P}| = 1$. This is surprising, as most ensembles of particles in nature are unpolarized, and if they are polarized, their degree of polarization seldom reaches 100 %. Hence something important is missing in our treatment. In fact, our description of a quantum state ψ is grossly incomplete because we must carefully distinguish between what is called a pure state and a mixed state.

If ψ is a basis state like α, or a linear superposition state like $\psi = a\alpha + b\beta$, then ψ is called a *pure state* and polarization is unity. Examples of pure states are (mostly taken from Part III of this text):

- The pure spin state $\psi_{\frac{1}{2},-\frac{1}{2}}$ of neutrinos (when assumed massless) with respect to their line of flight z with helicity $P_z = -1$.

- Linear superpositions of spin states, for instance, linearly polarized photons, when regarded as a coherent superposition of "spin up" and "spin down" photons, or qubits in quantum computing.
- A pure energy state ψ_D, for instance, polar molecules in the upper energy state in a molecular beam that feeds a maser.
- The pure pseudospin state of two-level atoms all in their low-energy state.
- The pure isospin state of an ensemble of protons.
- The pure momentum state of particles in a monoenergetic unidirectional beam.
- The pure macroscopic quantum state of a Bose–Einstein condensate, or of a superconducting Josephson junction.

In most cases, a pure state is an idealization because the polarization of particles never reaches the full 100 %, as every experimentalist knows. In real life, an ensemble of particles is at best a statistical mixture of such pure states, called a *mixed state*.

Let some fraction p_1 of spin-$\frac{1}{2}$ particles be in the spin up state, and the remaining fraction $p_2 = 1 - p_1$ in the spin down state, with both fractions completely independent of each other. In a beam of slow neutrons, at the time when one neutron passes a Stern–Gerlach magnet, the next neutron often is not even born yet; therefore, the polarization state of the first neutron cannot depend on the state of the second. Examples of mixed states are:

- Unpolarized particles, nuclei, or atoms in our surroundings.
- Molecules N_2 and O_2 in the ambient air that populate low-lying rotational and vibrational molecular levels.
- Photons from an imperfect optical polarizer.
- Particles in a beam of finite divergence.

We regard two cases of mixed states. In the first case, the statistical mixture is made up of a fraction p_1 in the pure spin up state α, and a fraction p_2 in the pure spin down state β. These two independent fractions have $\langle\sigma_z\rangle = \langle\sigma_{z1}\rangle + \langle\sigma_{z2}\rangle = p_1(\alpha^\dagger\sigma_z\alpha) + p_2(\beta^\dagger\sigma_z\beta)$. With (3.9c) and (3.4a), this gives the longitudinal polarization $\langle\sigma_z\rangle = p_1 - p_2$. With (3.9a) and (3.4b), we see that the transverse polarization vanishes, $\langle\sigma_x\rangle = p_1(\alpha^\dagger\sigma_x\alpha) + p_2(\beta^\dagger\sigma_x\beta) = 0$ and $\langle\sigma_y\rangle = 0$. The ensemble average of the polarization $\boldsymbol{P}_{\text{mixed}}$ then is along z and has amplitude

$$P_{\text{mixed}} = p_1 - p_2. \tag{6.19}$$

In the second case, the statistical mixture is made up of two orthogonal pure superposition states. Let a fraction p_1 be in state $\psi_1 = e^{-i\varphi/2}(\alpha\cos\frac{1}{2}\theta - \beta\sin\frac{1}{2}\theta)$, and a fraction p_2 in the orthogonal state $\psi_2 = e^{i\varphi/2}(\alpha\sin\frac{1}{2}\theta + \beta\cos\frac{1}{2}\theta)$, cf. (5.12). The fraction p_1 has polarization $\langle\boldsymbol{\sigma}_1\rangle = \psi_1^\dagger\boldsymbol{\sigma}\psi_1$, and the fraction p_2 has opposite polarization $\langle\boldsymbol{\sigma}_2\rangle = \psi_2^\dagger\boldsymbol{\sigma}\psi_2 = -\langle\boldsymbol{\sigma}_1\rangle$, as can be verified by direct calculation. Hence $\boldsymbol{P}_{\text{mixed}} = p_1\langle\boldsymbol{\sigma}_1\rangle + p_2\langle\boldsymbol{\sigma}_2\rangle$, or:

The polarization in the general mixed case is

$$\boldsymbol{P}_{\text{mixed}} = (p_1 - p_2)\boldsymbol{P} \equiv P_{\text{mixed}}\boldsymbol{P}. \tag{6.20}$$

Pure states have polarization $\boldsymbol{P} = \langle \boldsymbol{\sigma} \rangle$ of magnitude $|\boldsymbol{P}| = 1$. Mixed states have polarization of magnitude $P_{\text{mixed}} = |p_1 - p_2| \leq 1$. Talking about a *degree of polarization* always means P_{mixed}.

More generally, for a two-dimensional operator A with eigenstates ψ_1 and ψ_2, the ensemble average is

$$\langle \mathrm{A} \rangle = p_1 \langle \mathrm{A}_1 \rangle + p_2 \langle \mathrm{A}_2 \rangle = p_1 \big(\psi_1^\dagger \mathrm{A} \psi_1 \big) + p_2 \big(\psi_2^\dagger \mathrm{A} \psi_2 \big), \tag{6.21}$$

suitably projected as in (6.18) if necessary.

To summarize, we have two extreme cases of expectation values of an operator A, with eigenvalues A_1 and A_2:

1. For a pure ensemble, the expectation value $\langle \mathrm{A} \rangle = \psi^\dagger \mathrm{A} \psi$ derives from the quantum probability amplitudes ψ.
2. For a mixed ensemble, the expectation value $\langle \mathrm{A} \rangle = p_1 A_1 + p_2 A_2$ derives from the classical occupation probabilities p_i.

Let us have a closer look at the second case of a mixed ensemble of spin up α and spin down β states, with no transverse polarization. If one counts a number $N_\uparrow = p_1 N$ of particles with spin up, and a number $N_\downarrow = p_2 N$ with spin down, then the longitudinal spin polarization is (omitting the subscript "mixed", and writing with $p_{\frac{1}{2}}$ for p_1 and $p_{-\frac{1}{2}}$ for p_2)

$$P_z = \langle \sigma_z \rangle = \frac{N_\uparrow - N_\downarrow}{N_\uparrow + N_\downarrow} = p_{\frac{1}{2}} - p_{-\frac{1}{2}}, \tag{6.22}$$

and $P_x = P_y = 0$. For $s > \frac{1}{2}$, this expression is replaced by a weighted sum over magnetic quantum numbers m with weights p_m, in which case $P_z = \langle \mathrm{S}_z \rangle / \hbar s$ becomes

$$P_z = \frac{1}{s} \frac{\sum_m m p_m}{\sum_m p_m}, \tag{6.23}$$

with summation index m running from $-s$ to $+s$. The same formula is used to calculate the average grading $\langle m \rangle$ of a class of students of number $\sum_m p_m$.

So far, the results hold for any kind of two-level system with effective spin **S**, whatever the interaction under study. For the special case of a sample of n paramagnetic atoms per volume element, the macroscopic *magnetization* is, with (2.5),

$\boldsymbol{M} = n\langle\boldsymbol{\mu}\rangle = n\mu\langle\mathbf{S}\rangle/\hbar s$ or, using $\boldsymbol{P} = \langle\mathbf{S}\rangle/\hbar s$,

$$\boldsymbol{M} = n\mu\boldsymbol{P}. \tag{6.24}$$

A sizable *brute force* magnetization can be achieved by exposing a paramagnet at low temperature T to a strong magnetic field B. Then, in thermal equilibrium, the (unnormalized) *Boltzmann population*

$$p_m = \mathrm{e}^{-E_m/k_\mathrm{B}T} = \mathrm{e}^{mg\mu_\mathrm{B}B/k_\mathrm{B}T}, \tag{6.25}$$

with E_m taken from (2.12), must be inserted into (6.23), with magnetic quantum number m and the appropriate g-factor. For $s = \frac{1}{2}$, (6.22) then gives what usually is called a magnetization curve, namely the polarization

$$P_z = \frac{\mathrm{e}^x - \mathrm{e}^{-x}}{\mathrm{e}^x + \mathrm{e}^{-x}} = \tanh x, \quad x = \mu B/k_\mathrm{B}T, \tag{6.26}$$

which saturates at ± 1 for high fields with $|x| \gg 1$. For low field and/or high temperature, with $|x| \ll 1$ and $\tanh x \approx x$, we recover *Curie's law*

$$P_z \approx \mu B/k_\mathrm{B}T \quad \text{and} \quad |M_z| \approx n\mu^2 B/k_\mathrm{B}T. \tag{6.27}$$

For paramagnetic atoms of spin quantum number s in a probe of condensed matter, Curie's law gives the atomic spin polarization,

$$P_\mathrm{atom} \approx 0.67 gs\, B(\mathrm{T})/T(\mathrm{K}), \quad \text{for } |P_\mathrm{atom}| \ll 1. \tag{6.28}$$

One needs temperatures of liquid helium (i.e., $T \leq 4.2$ K) to achieve a sizeable polarization even in the strong field of a superconducting magnet. Free electrons have $gs \approx -2 \times \frac{1}{2} = -1$.

Nuclei with spin quantum number I inside diamagnetic atoms are polarized to

$$P_\mathrm{nucl} \approx 0.36 \times 10^{-4} g_I I\, B(\mathrm{T})/T(\mathrm{K}), \quad \text{for } |P_\mathrm{nucl}| \ll 1. \tag{6.29}$$

One needs millikelvin temperatures to polarize nuclei in a strong magnetic field. Protons have spin $I = \frac{1}{2}$ and $g_I I = 2.8$.

6.5 The Density Matrix

Quantum mechanics makes predictions exclusively on observable quantities, given by their expectation values (6.21). Wave functions ψ are not observables, and the concept of density matrices is introduced to eliminate these nonobservables from the basic equations. The density matrix formalism displays its full power mainly beyond the two-state systems, as will be seen in Chap. 19 and later. For a two-state system, the use of the density matrix is somewhat redundant, as we can only derive again our earlier results on spin precession. However, this new derivation will show in a transparent way how the density matrix ρ is used in general.

We start with a general N-state system. Let us take a mixed ensemble of independent atoms or molecules, distributed over quantized energy levels E_n, $n =$

$1, 2, 3, \ldots, N$ with population numbers p_n normalized to $\sum_n p_n = 1$. We begin our discussion in the energy representation, with diagonal energy matrix $\mathrm{H_D}$, which is without loss of generality when we look at quantities that are independent of representation. If we can choose any representation, we choose the one that is easiest to handle.

The average energy of this mixed ensemble is $\langle \mathrm{H} \rangle = \sum_n p_n E_n$. We introduce the density matrix ρ_D as the diagonal matrix that has the occupation numbers p_n arranged along its diagonal, hence

$$\rho_\mathrm{D} = \begin{pmatrix} p_1 & 0 & \ldots \\ 0 & p_2 & \ldots \\ \ldots & \ldots & \ldots \end{pmatrix}, \qquad \mathrm{H_D} = \begin{pmatrix} E_1 & 0 & \ldots \\ 0 & E_2 & \ldots \\ \ldots & \ldots & \ldots \end{pmatrix}, \tag{6.30}$$
$$\rho_\mathrm{D}\mathrm{H_D} = \begin{pmatrix} p_1 E_1 & 0 & \ldots \\ 0 & p_2 E_2 & \ldots \\ \ldots & \ldots & \ldots \end{pmatrix}.$$

From these diagonal matrices we can derive the following properties of the density matrix, all independent of representation, and therefore valid for general density operators ρ and general hermitian operators like H.

The trace of an operator $\mathrm{Tr}(\rho) = \sum_n p_n$ is the sum of its diagonal elements. Normalization $\sum_n p_n = 1$ implies:

> The density operator has trace one, $\mathrm{Tr}(\rho) = 1$.

The trace of the square of the density operator generally is smaller that unity, $\mathrm{Tr}(\rho^2) = \sum_n p_n^2 \leq 1$, because for $p_n \geq 0$ we have $(\sum_n p_n)^2 \geq \sum_n p_n^2$, thus $1 = [\mathrm{Tr}(\rho)]^2 \geq \mathrm{Tr}(\rho^2)$. The equal sign holds for a pure state with one of the occupation numbers equaling one, and with all other occupation numbers vanishing:

> For a pure state $\mathrm{Tr}(\rho^2) = 1$. For a mixed state $\mathrm{Tr}(\rho^2) < 1$.

The elements of the diagonal density matrix ρ_D are real, therefore ρ is hermitian, with eigenvalues p_n.

A very important property of the density operator ρ is that the mean energy $\langle \mathrm{H} \rangle$ is the trace of the product matrix $\rho_\mathrm{D}\mathrm{H_D}$ in (6.30), $\langle \mathrm{H_D} \rangle = \mathrm{Tr}(\rho_\mathrm{D}\mathrm{H_D}) = \sum_n p_n E_n$, which holds regardless of representation:

> The expectation value of the Hamiltonian is
>
> $$\langle \mathrm{H} \rangle = \mathrm{Tr}(\rho \mathrm{H}). \tag{6.31}$$

Equation (6.31) replaces our earlier definition (3.28) of an expectation value. Later on, we shall generalize this result to the expectation value of arbitrary hermitian operators, see Sect. 19.1, which implies that knowledge of the density operator ρ is sufficient to predict the outcome of any experiment.

So far we have studied the properties of the density operator that are independent of representation and dimension, and used for this the diagonal representation. We now specialize to the spin-$\frac{1}{2}$ case, and apply a unitary transformation T, (5.12), to the density operator ρ, while keeping the field vector $\boldsymbol{B}$ along z. The rotated density matrix is

$$\rho = \mathrm{T}\rho_D\mathrm{T}^\dagger, \tag{6.32a}$$

$$\rho = \begin{pmatrix} c^* & -s \\ s^* & c \end{pmatrix} \begin{pmatrix} p_1 & 0 \\ 0 & p_2 \end{pmatrix} \begin{pmatrix} c & s \\ -s^* & c^* \end{pmatrix}, \tag{6.32b}$$

$$\rho = \begin{pmatrix} cc^* p_1 + ss^* p_2 & sc^*(p_1 - p_2) \\ s^*c(p_1 - p_2) & ss^* p_1 + cc^* p_2 \end{pmatrix}. \tag{6.32c}$$

In the special case of a pure state, say, $p_1 = 1$, this reduces to

$$\rho = \begin{pmatrix} cc^* & sc^* \\ s^*c & ss^* \end{pmatrix}. \tag{6.33}$$

We insert $c = \mathrm{e}^{\mathrm{i}\varphi/2}\cos\frac{1}{2}\theta$ and $s = \mathrm{e}^{-\mathrm{i}\varphi/2}\sin\frac{1}{2}\theta$ from (5.13) into (6.32c) and arrive at

$$\rho = \frac{1}{2}\begin{pmatrix} 1 + (p_1 - p_2)\cos\theta & (p_1 - p_2)\mathrm{e}^{-\mathrm{i}\varphi}\sin\theta \\ (p_1 - p_2)\mathrm{e}^{\mathrm{i}\varphi}\sin\theta & 1 - (p_1 - p_2)\cos\theta \end{pmatrix}. \tag{6.34}$$

Next we are going to prove the following statement:

The 2×2 density matrix ρ and the polarization $\boldsymbol{P} = \langle\boldsymbol{\sigma}\rangle$ are related as

$$\rho = \frac{1}{2}\big(\mathrm{I} + \langle\boldsymbol{\sigma}\rangle \cdot \boldsymbol{\sigma}\big), \tag{6.35a}$$

$$\rho = \frac{1}{2}\begin{pmatrix} 1 + P_z & P_x - \mathrm{i}P_y \\ P_x + \mathrm{i}P_y & 1 - P_z \end{pmatrix}, \tag{6.35b}$$

which also holds for the mixed case. In the pure case, the density operator (6.35a) coincides with the projection operator (6.17).

Recall that $\langle\boldsymbol{\sigma}\rangle$ in (6.35a) is an ordinary vector with real coefficients P_x, P_y, and P_z, while I and $\boldsymbol{\sigma}$ operate in Hilbert space. When we compare (6.35b) with (6.34), with $\boldsymbol{P} = \langle\boldsymbol{\sigma}\rangle$ from (3.36), we are back to $\boldsymbol{P}_{\text{mixed}} = (p_1 - p_2)\boldsymbol{P}$, (6.20), with $|\boldsymbol{P}_{\text{mixed}}| \le 1$, and $|\boldsymbol{P}| = 1$.

To prove (6.35a), we use the following properties of the Pauli matrices σ_α, $\alpha = x, y, z$: $\mathrm{Tr}(\sigma_\alpha) = 0$, $\mathrm{Tr}(\sigma_\alpha^2) = 2$, $\mathrm{Tr}(\sigma_\alpha\sigma_\beta) = 0$ for $\alpha \neq \beta$, from (3.8b) and (3.11). Since ρ is hermitian, it can be expanded in terms of the Pauli matrices as $\rho = c_1\mathrm{I} + c_x\sigma_x + c_y\sigma_y + c_z\sigma_z$, as was done for H in (3.17b). Multiplying this with σ_x and taking the trace $\mathrm{Tr}(\rho\sigma_x)$, we find that $\langle\sigma_x\rangle = \mathrm{Tr}(\rho\sigma_x) = 2c_x$, or $c_x = \frac{1}{2}\langle\sigma_x\rangle$, and similarly $c_y = \frac{1}{2}\langle\sigma_y\rangle$, $c_z = \frac{1}{2}\langle\sigma_z\rangle$. Hence $\rho = \frac{1}{2}(\mathrm{I} + \langle\sigma_x\rangle\sigma_x + \langle\sigma_y\rangle\sigma_y + \langle\sigma_z\rangle\sigma_z)$, with $\rho = \frac{1}{2}\mathrm{I}$ for an unpolarized system, which is (6.35a).

With the help of the density matrix, we are able to derive again our earlier spin precession results from $\boldsymbol{P}(t) = \langle\boldsymbol{\sigma}\rangle(t) = \mathrm{Tr}[\rho(t)\boldsymbol{\sigma}]$. Knowing the elements of the density operator, we can study its time evolution $\rho(t) = \mathrm{U}(t)\rho(0)\mathrm{U}^\dagger(t)$, with time evolution operator $\mathrm{U}(t) = \mathrm{e}^{-\mathrm{i}Ht/\hbar}$. As an example, we take an initial transverse polarization $P_x(0)$ and insert the angles $\theta = \frac{1}{2}\pi$ and $\varphi = 0$ into (6.34) to obtain the initial density matrix

$$\rho(0) = \frac{1}{2}\begin{pmatrix} 1 & p_1 - p_2 \\ p_1 - p_2 & 1 \end{pmatrix}. \tag{6.36}$$

With $B = B_z$ the evolution operator $\mathrm{U}(t)$ from (6.2) is diagonal with elements $\exp(\mp\frac{1}{2}\mathrm{i}\omega_\mathrm{L}t)$. With this the density matrix evolves as

$$\rho(t) = \mathrm{U}(t)\rho(0)\mathrm{U}^\dagger(t) = \frac{1}{2}\begin{pmatrix} 1 & (p_1 - p_2)\mathrm{e}^{-\mathrm{i}\omega_\mathrm{L}t} \\ (p_1 - p_2)\mathrm{e}^{\mathrm{i}\omega_\mathrm{L}t} & 1 \end{pmatrix}. \tag{6.37}$$

The time evolution of the x-component of polarization then is

$$P_x(t) = \langle\sigma_x\rangle(t) = \mathrm{Tr}\big[\rho(t)\sigma_x\big] = (p_1 - p_2)\cos\omega_\mathrm{L}t, \tag{6.38}$$

and similarly $P_y(t) = (p_1 - p_2)\sin\omega_\mathrm{L}t$. Altogether the Larmor precession in the x–y plane becomes $P_\perp(t) = (p_1 - p_2)(\cos\omega_\mathrm{L}t, \sin\omega_\mathrm{L}t, 0)$, in agreement with the earlier spin rotation result (3.52). For general values of the angles θ and φ that rotate $\boldsymbol{P}$ away from $\boldsymbol{B}$, we can in a similar way recover the formula (5.25) for the longitudinal polarization $P_z(t) = \cos^2\theta + \sin^2\theta\cos\omega_\mathrm{L}t$.

For the two-state system, the density matrix approach is not simpler than direct calculation either from the Schrödinger equation or from the spin precession picture that both lead us to $P_z(t)$. Still, the 2×2 density matrix is helpful to get an insight into the general procedure. For pure states, the hermitian $N \times N$ density matrix with $\mathrm{Tr}(\rho) = 1$ has $N^2 - 1$ real parameters, while the normalized wave function of the Schrödinger equation has $2N - 1$ real parameters, unless these numbers are reduced by symmetries. Therefore, the density matrix approach generally requires more numerical effort, but has considerable advantages in quantum statistics, as will become evident in Chap. 19 and later.

6.6 Coherence and Interference

In classical optics, light waves are called *coherent* if they can give rise to interference signals. To be able to interfere, partial wave amplitudes must have stable phase relations with each other.

The components of the density matrix are called coherent if they can give rise to spin rotation signals. According to (3.31a)–(3.31c), with $\boldsymbol{B}$ along z, the phase $\varphi = \omega_L t$ that leads to spin rotation enters $\langle\sigma_x\rangle$ and $\langle\sigma_y\rangle$ via terms a^*b and b^*a. The component $\langle\sigma_z\rangle = |a|^2 - |b|^2$ is not phase sensitive and does not contribute to the spin rotation signal. Therefore, in (6.32c), only the off-diagonal elements constitute the coherent part of the density matrix ρ and contribute to the signal via the transverse polarization terms $\langle\sigma_x\rangle\sigma_x$ and $\langle\sigma_y\rangle\sigma_y$. The diagonal elements of ρ depend on $\langle\sigma_z\rangle\sigma_z$ and constitute the *incoherent* part of ρ. The decay of coherence with time will be discussed in Chap. 22 on relaxation and decoherence.

This separation of ρ into coherent and incoherent parts is independent of a representation. In our derivation, the field $\boldsymbol{B}$ stayed along the quantization axis z when ρ and with it the polarization $\boldsymbol{P}$ were rotated away from z via (6.32a). If the z-axis is changed, the results obtained stay the same because the angles θ and φ between $\boldsymbol{B}$ and $\boldsymbol{P}$ remain the same.

Let us return to the discussion of coherence and interference in wave optics that we had started in Chap. 1, see Fig. 1.1. First, we note that, although the name coherent is applied to both interference and spin rotation signals, there is a notable difference: interference signals, for instance, in the double slit experiments, are generated by partial waves of equal frequency, while spin rotation signals are beat signals from the superposition of waves of different frequencies.

An important result of quantum statistics links self-interference of a matter wave to uncertainty, via its properties in *phase space*, which is the six-dimensional space made up of the particle's position $\boldsymbol{x}$ and momentum $\boldsymbol{p}$ coordinates.

Self-interference signals are visible only if the matter waves of the particles are confined within the same *phase space cell* of volume

$$\Delta\boldsymbol{p} \cdot \Delta\boldsymbol{x} \leq h^3. \tag{6.39}$$

Therein, h is Planck's constant, $\Delta\boldsymbol{p} \equiv \Delta p_x \Delta p_y \Delta p_z$ is the cell volume in momentum space, and $\Delta\boldsymbol{x} \equiv \Delta x \Delta y \Delta z$ is called the *coherence volume*. Equation (6.39) states that interference is possible only if position and momentum of the particles are blurred by the uncertainty relation. Uncertainty therefore goes hand in hand with the ability to interfere, with the two limiting cases:

1. If, due to uncertainty, we cannot know which path a particle takes, then interference signals become visible, which means that we must first add partial amplitudes and then take the absolute square,

$$|\Psi|^2 = \left|\sum_i \psi_i\right|^2. \tag{6.40}$$

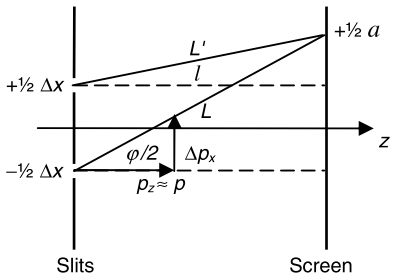


Fig. 6.1 Optical paths in a two-slit experiment. The distance between the slits is Δx, and the width of the interference pattern on the screen is a. The light is emitted under small angles $\varphi \approx a/l \approx \Delta p_x/p \ll 1$

On the other hand,

2. If we can know which path a particle takes, then no interference is visible, which means that we must add the probabilities for the different possible paths of the particles

$$|\Psi|^2 = \sum_i |\psi_i|^2. \tag{6.41}$$

In this way, the uncertainty principle prevents that anyone can uncover the secrets of particle–wave duality. Equations (6.40) and (6.41) are generalizations of (1.13) and (1.14), holding for the two-slit experiment. We want to show the validity of the interference condition (6.39) for the special case of an optical two-slit experiment. A beam of light can give rise to interference under the following conditions:

1. Temporal or *longitudinal coherence*: This assures that the two partial waves passing the slits arrive at approximately the same time at a point within the screen (Fig. 6.1).
2. Spatial or *transverse coherence*: This assures that the path difference from the two slits to the point on the screen is of the order of the wavelength λ or smaller.

1. Longitudinal coherence: From Fourier's theorem, the duration of a wave train of spectral width $\Delta\omega$ is $\Delta t = 1/\Delta\omega$, and its longitudinal coherence length is $\Delta z = c\Delta t = c/\Delta\omega$, where Δt is shorter than or equal to the lifetime τ of the light-emitting atomic state. A lifetime $\tau = 1$ ns implies $\Delta z \leq 0.3$ m. The momentum of the photons in the narrow beam is $p = h/\lambda = \hbar\omega/c \approx p_z$, and the uncertainty in longitudinal momentum is $\Delta p_z \approx \Delta p = \hbar\Delta\omega/c$. Hence $\Delta p_z \Delta z \approx \hbar$, in agreement with the uncertainty relation, but here deduced from the requirement of interference.
2. Transverse coherence: Let the two slits be arranged in two dimensions, x and z, as shown in Fig. 6.1. The distances L and L' of the two slits to the point on the screen are given by $L^2 = l^2 + \frac{1}{4}(a + \Delta x)^2$ and $L'^2 = l^2 + \frac{1}{4}(a - \Delta x)^2$, thus $L^2 - L'^2 = (L + L')(L - L') = a\Delta x$. For $l \gg \Delta x$ and $l \gg a$ (far field) we have $L + L' \approx 2l$, and the path difference is $L - L' \approx a\Delta x/2l$. For lateral coherence, this path difference must be of the order of the wavelength, $L - L' \approx \lambda$, or $\Delta x \approx (2l/a)\lambda$. This means that the width a of the interference pattern is large for a small slit width Δx and small for a large slit width. To derive the uncertainty Δp_x in lateral momentum we note that the small angle of

light emission at the slits is $\varphi \approx a/l \approx \Delta p_x/p$. Hence the uncertainty in lateral momentum is $\Delta p_x \approx (a/l)p = (a/l)(h/\lambda)$, and the appearance of interference signals requires $\Delta p_x \Delta x \approx 2h$.

In our crude two-dimensional model, interference is possible when the waves are inside a phase space volume $V = \Delta \boldsymbol{p} \Delta \boldsymbol{x} \approx h^2/\pi$. The model can be extended to three dimensions with the result (6.39).

For the visibility of spin rotation signals, important is not the momentum-space uncertainty limit, but the energy-time uncertainty limit, as will be seen in the discussion of atomic quantum beats in Sect. 12.3.

6.7 Dirac's Bra-Ket Notation

For the study of more complicated quantum systems, the *Dirac bra-ket notation* is very useful. A classical two-dimensional unit vector $\boldsymbol{n}$ in real space under angle φ to the x-axis is written in terms of its orthonormal basis vectors $\boldsymbol{e}_x$ and $\boldsymbol{e}_y$, the unit vectors in x and y directions, as

$$\boldsymbol{n} = \begin{pmatrix} n_x \\ n_y \end{pmatrix} = \begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix} = \boldsymbol{e}_x \cos\varphi + \boldsymbol{e}_y \sin\varphi, \tag{6.42a}$$

$$\boldsymbol{n} = \boldsymbol{e}_x(\boldsymbol{e}_x \cdot \boldsymbol{n}) + \boldsymbol{e}_y(\boldsymbol{e}_y \cdot \boldsymbol{n}). \tag{6.42b}$$

The expansion coefficients n_x, n_y are the scalar products $\boldsymbol{e}_x \cdot \boldsymbol{n}$ and $\boldsymbol{e}_y \cdot \boldsymbol{n}$, i.e., the projections of $\boldsymbol{n}$ onto $\boldsymbol{e}_x$ and $\boldsymbol{e}_y$, and are normalized to $n_x^2 + n_y^2 = 1$.

In quantum mechanics, we can do the same for the state vector ψ. Instead of $\psi = a\alpha + b\beta$, we write, in analogy to (6.42b), in Dirac notation

$$|\psi\rangle = |\alpha\rangle\langle\alpha|\psi\rangle + |\beta\rangle\langle\beta|\psi\rangle. \tag{6.43}$$

The expansion coefficients are the scalar products $\langle\alpha|\psi\rangle = a$ and $\langle\beta|\psi\rangle = b$ of (3.24), which are the projections of ψ onto α and β, normalized to $|a|^2 + |b|^2 = 1$. Equation (6.43) can be interpreted as the multiplication of $|\psi\rangle$ from the left, $|\psi\rangle = \mathrm{I}|\psi\rangle$, with the unit operator

$$\mathrm{I} = |\alpha\rangle\langle\alpha| + |\beta\rangle\langle\beta|. \tag{6.44}$$

To prove this statement we recall the definition of the *direct product* of two vectors

$$\begin{pmatrix} a \\ b \end{pmatrix} (c \quad d) \equiv \begin{pmatrix} ac & ad \\ bc & bd \end{pmatrix}, \tag{6.45}$$

the two terms in (6.44) are written as

$$|\alpha\rangle\langle\alpha| = \begin{pmatrix} 1 \\ 0 \end{pmatrix} (1 \quad 0) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad |\beta\rangle\langle\beta| = \begin{pmatrix} 0 \\ 1 \end{pmatrix} (0 \quad 1) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \tag{6.46}$$

We see that the individual terms in (6.44) are the projection operators from (6.12), $P_{z+} = |\alpha\rangle\langle\alpha|$ and $P_{z-} = |\beta\rangle\langle\beta|$, whose sum is the unit operator I. Equation (6.44) is the closure relation met before in (6.13), which expresses the fact that the basis states $|\alpha\rangle$ and $|\beta\rangle$ span the entire Hilbert space.

The direct product should not to be confused with the scalar product

$$(c \quad d)\begin{pmatrix} a \\ b \end{pmatrix} = ac + bd. \tag{6.47}$$

Notice that, due to the trace property Tr(AB) = Tr(BA), it does not make a difference if one takes the trace of a scalar or of a direct product. Therefore, the scalar product (6.47) equals the trace of the direct product, (6.45).

The Dirac notation is useful whenever the state $|\psi\rangle$ is characterized by a larger set of quantum numbers. Instead of decorating ψ with many subscripts, like $\psi_{nlm_l sm_s}$ for the hydrogen atom, one drops the letter ψ and writes instead the ket $|n\ l\ m_l\ s\ m_s\rangle$.

The matrix elements $H_{mm'}$ of the Hamiltonian then are written as $\langle m|\mathrm{H}|m'\rangle$, and its diagonalization $\mathrm{T}^\dagger\mathrm{HT} = \mathrm{H_D}$ as

$$\sum_{mm'} \langle E_n|m\rangle\langle m|\mathrm{H}|m'\rangle\langle m'|E_{n'}\rangle = E_n\delta_{nn'}. \tag{6.48}$$

The expansion coefficients $\langle E_n|m\rangle$ in this equation are the elements of the unitary transformation matrix T, i.e., the coefficients $\langle E_n|m\rangle$ are the components of the eigenvector ψ_n to the eigenvalue E_n. We return to the diagonal representation $\mathrm{H_D}$ with matrix elements $\langle E_n|\mathrm{H_D}|E_{n'}\rangle = E_n\delta_{nn'}$, simply by cutting out the terms $\sum_m |m\rangle\langle m| = \mathrm{I}$ and $\sum_{m'} |m'\rangle\langle m'| = \mathrm{I}$ from (6.48), where $\delta_{nn'}$ is the Kronecker symbol.

Any operator A of an observable can be expressed in terms of its eigenvalues a_n and eigenstates $|n\rangle$ as

$$\mathrm{A} = \sum_n a_n|n\rangle\langle n|. \tag{6.49}$$

In particular, the density operator with the population numbers p_n as eigenvalues becomes $\rho = \sum_n p_n|n\rangle\langle n|$. In the representation diagonal with respect to the eigenstates of A, the elements of matrix $\mathrm{A_D}$ are $\langle n|A_\mathrm{D}|n'\rangle = a_n\delta_{nn'}$. With respect to any other basis, A must be transformed accordingly.

For extended calculations, the Dirac notation is highly recommended, and we shall make use of it in later chapters. However, for the less experienced readers, compact equations like $\mathrm{T}^\dagger\mathrm{HT} = \mathrm{H_D}$ are easier to grasp than their equivalent (6.48) in Dirac's notation.

This page intentionally left blank

Chapter 7
Diabolic Points, Geometric Phases, and Quantum Chaos

Abstract Quantum systems have a number of peculiar features that can all be traced back to one common origin, namely the universal "diabolic" shapes of energy surfaces in the vicinity of a degeneracy. The first such feature is the avoided-crossing effect, when energy levels do not cross but seem to repel each other when plotted in dependence of an external parameter. The next feature is the occurrence of geometric phases, also known as Berry phases, which accumulate during excursions in parameter space. We discuss experimental examples of such phases, both in the space of magnetic fields and in what we call the space of shapes. The Aharonov–Bohm phase is closely related to such geometric phases. Finally, in a short section we show that the hallmarks of quantum chaos can be understood in terms of locally defined 2×2 matrices and their avoided level crossings.

7.1 Level Crossings and Level Repulsions

In quantum physics, energy level diagrams $E_n(B)$, when plotted as a function of some external parameter B, often show a peculiar behavior. If in such a plot two neighboring energy levels approach each other with different slopes, then one expects them to cross at some parameter value. Instead, often the levels seem to repel each other near the expected crossing point. This so-called *anticrossing* or *avoided-crossing* effect is understood best when we first look at a two-level spin-$\frac{1}{2}$ system. This will be followed by an example of a three-level spin-1 system.

7.1.1 The Field Dependence of Energy and of Polarization

In a 2×2 Hamiltonian, diagonal and off-diagonal elements can often be varied independently, some examples were given at the start of Chap. 5. In the case of a magnetic interaction, we can, for instance, keep the off-diagonal transverse component fixed at $B_x = B_{x0}$, $B_y = 0$, and vary the external field component B_z on the diagonal. The energy eigenvalues

$$E_{\pm}(B_z) = \mp\mu\sqrt{B_{x0}^2 + B_z^2} \tag{7.1}$$

D. Dubbers, H.-J. Stöckmann, *Quantum Physics: The Bottom-Up Approach*, Graduate Texts in Physics,
DOI 10.1007/978-3-642-31060-7_7,

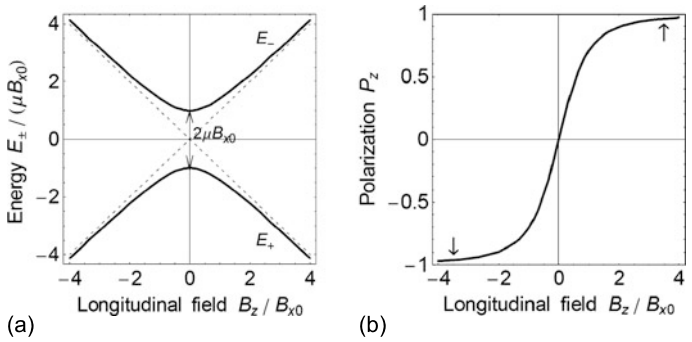


Fig. 7.1 (**a**) An avoided level crossing appears if, for instance, the field component $B_x = B_{x0}$ is kept constant, and the energy eigenvalues $E_\pm$ are plotted as a function of B_z. The gap at $B_z = 0$ then has the size $2\mu B_{x0}$ of twice the off-diagonal matrix element. For $B_{x0} = 0$, E_+ and E_- do cross, and we have the usual Zeeman splitting $E_\pm = \mp\mu B_z$ of Fig. 2.2, as indicated by the *dashed line*. (**b**) Due to state mixing near an avoided crossing, the polarization P_z is also field dependent. It turns out to be the derivative $\partial E_-/\partial B_z$ of the *upper curve* in (**a**)

from (3.21) then depend hyperbolically on B_z as shown in Fig. 7.1(a), with an avoided crossing near $B_z = 0$.

The probabilities $p_{\text{up,down}} = \frac{1}{2}[1 \pm B_z/(B_{x0}^2 + B_z^2)^{1/2}]$ for being spin up or spin down, (5.15a), (5.15b), also depend on B_z. For large positive $B_z \gg B_{x0}$, the probability p_{up} approaches unity, and the upper energy state is approximately a pure spin up state, while for large negative B_z, p_{up} vanishes and the upper energy state is approximately a pure spin down state, as indicated by the arrows in Fig. 7.1(b). For $B_z = 0$, that is, for a purely transverse field B_{x0}, the spin up and spin down states enter with equal weights, with amplitudes $\cos\frac{1}{4}\pi = \sqrt{\frac{1}{2}}$ and $\sin\frac{1}{4}\pi = \sqrt{\frac{1}{2}}$. This is called a *state mixing*, not to be confused with the mixed state of Sect. 6.4.

The longitudinal polarization $P_z = p_{\text{up}} - p_{\text{down}}$, shown in Fig. 7.1(b), then varies as

$$P_z(B_z) = \frac{B_z}{\sqrt{B_{x0}^2 + B_z^2}}. \tag{7.2}$$

From classical electrodynamics, we know that the macroscopic magnetization of a system is the derivative of magnetic energy with respect to the magnetic field. We find the same in the quantum case if we take the B_z-derivative of the energy eigenvalue of E_- from (7.1)

$$\frac{\partial E_-}{\partial B_z} = \frac{\mu B_z}{\sqrt{B_{x0}^2 + B_z^2}} = \mu P_z(B_z). \tag{7.3}$$

The field dependence of longitudinal polarization $P_z(B_z)$ is given by the derivative $\partial E_-/\partial B_z$ of the energy eigenvalue $E_-(B_z)$.

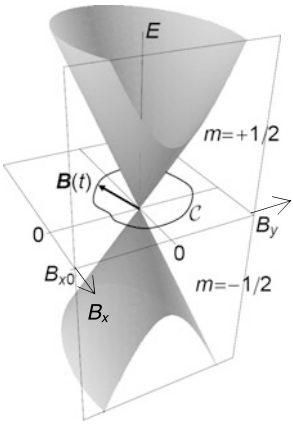


Fig. 7.2 The Zeeman energies $E_\pm(B_x, B_y)$, see (7.4), define a double cone or "diabolo". If the x-component of the field is kept constant at $B_x = B_{x0}$, the hyperbolic level scheme $E(B_{x0}, B_y)$ of Fig. 7.1(a) (there in the x–z plane) reappears on the *front face of the figure*. The trajectory $\mathcal{C}$ shows an example of an excursion $\boldsymbol{B}(t)$ in B_x–B_y parameter space

More generally, we may write (7.3) as $\nabla_{\boldsymbol{B}} E_- = \mu \boldsymbol{P}$, with $\nabla_{\boldsymbol{B}} = (\partial/\partial B_x, \partial/\partial B_x, \partial/\partial B_x)^{\mathrm{T}}$. Indeed, taking $E_- = \mu(B_x^2 + B_x^2 + B_x^2)^{1/2}$ from (3.21), we have $\nabla_{\boldsymbol{B}} E_- = \mu \boldsymbol{B}/B$, and for an energy eigenstate we had found before that $\boldsymbol{P} = \langle \boldsymbol{\sigma} \rangle = \boldsymbol{B}/B$, see (5.20). A level crossing then occurs if one varies one component of $\boldsymbol{B}$, no matter whether it enters via a diagonal or an off-diagonal matrix element, while keeping the other components fixed. N.B.: Do not confuse $P_z(B_z)$, Fig. 7.1(b), of a pure superposition state ψ with the similar-looking magnetization curve $P_z(B)$, see (6.26), of a mixed state with Boltzmann population.

The avoided-crossing effect is a general property of the eigenvalues of any Hamiltonian with nonzero off-diagonal matrix elements.

7.1.2 Level Repulsion in a Spin-$\frac{1}{2}$ Systems

To introduce the notion of a *diabolic point*, first we go to the B_x–B_y plane and plot the energy $E_\pm(B_x, B_y) = \mp\mu(B_x^2 + B_y^2)^{1/2}$ as a function of the two independent parameters B_x and B_y. The energy surface shown in Fig. 7.2 then has the form of a double cone with circular horizontal cross-sections. If, again, one field component B_x is fixed at a value $B_{x0} \neq 0$ and $E_\pm$ is plotted as a function of the other field component B_y, then we are back to an equivalent of (7.1), namely

$$\frac{E^2}{\mu^2 B_{x0}^2} - \frac{B_y^2}{B_{x0}^2} = 1. \tag{7.4}$$

Seen as a function $E_\pm(B_{x0}, B_y)$, this defines the pair of hyperbolae that appears on the front face of Fig. 7.2, and which corresponds to the curves in Fig. 7.1(a). The two curves again are separated by a gap of size $2\mu B_{x0}$, and only cross if $B_{x0} = 0$.

The conical shape of the energy surfaces near a degeneracy is a general feature of quantum physics. The point of degeneracy where the two cones touch is called

a diabolic point. These surfaces are called diabolic because they are shaped like a diabolo, an ancient juggling toy. To show how this comes about, let us take a real $N \times N$ Hamiltonian matrix that depends on two external parameters p and q.

The energy eigenvalues $E_n(p, q)$, $n = 1, \ldots, N$, plotted as functions of two parameters p and q of the Hamiltonian, form surfaces that do no cross but at most touch each other for certain diabolic points (p_0, q_0) in the p–q plane.

To prove this statement, we investigate the immediate neighborhood of such a degeneracy in the p–q plane. We need only to consider the two energy levels E_i and E_k involved in the degeneracy (see Fig. 7.4(b) for a more general case). Hence these two energy surfaces of an $N \times N$ Hamiltonian can be regarded as belonging to an $s = \frac{1}{2}$ effective-spin system, locally described by a 2×2 matrix. We develop the matrix elements H_{ik} linearly as functions of the distances $x = p - p_0$ and $y = q - q_0$ to the diabolic point (p_0, q_0). Then

$$\mathrm{H}_{\mathrm{loc}} \approx \begin{pmatrix} ax + by & cy \\ cy & -ax - by \end{pmatrix}, \tag{7.5}$$

with $a = \partial H_{11}/\partial x$, etc. Without loss of generality, we chose the zero of energy at the vertex of the double cones, and the coordinates x and y such that H is diagonal for $y = 0$. The characteristic equation $\det(\mathrm{H} - E\mathrm{I}) = 0$ then gives $E_{\pm} = \pm[a^2x^2 + 2abxy + (b^2 + c^2)y^2]^{1/2}$. This is the formula of a double cone with elliptical cross-section in any plane at right angles to the E-axis, and with hyperbolic avoided crossings in any plane parallel to the E-axis. In short, we have an (elliptically deformed) diabolo centered at $p = p_0$ and $q = q_0$, as postulated.

In the Hanle effect, described in Sect. 4.2, the diabolic point of Fig. 7.2 is crossed along a line through the origin of $\boldsymbol{B}$-space. Via state mixing, this crossing leads to a resonance-like polarization signal, whose width is proportional to the lifetime of the short-lived atomic state. One can detect such polarization signals for general level crossings in parameter space, which is the basis of *level crossing spectroscopy*. Such level crossing signals are helpful for the construction of complicated atomic energy level diagrams.

Sometimes energy levels do cross at nonzero field values and do not repel each other. This happens if the Hamiltonian is *block-diagonal* with several submatrices arranged along the matrix diagonal (an example of block diagonal matrices is shown in Fig. 8.1). The energy level schemes from different submatrices then are independent of each other and are free to cross.

7.1.3 Level Repulsion in a Spin-1 System

To find a diabolic point beyond the trivial case of Fig. 7.2, we shortly enter the topic of electric quadrupole interactions. This next term in the multipole expansion of the electromagnetic interaction will be treated in detail in Sect. 18.2.2. The deuteron, the bound state of a proton and a neutron, is the mass $A = 2$ isotope of hydrogen

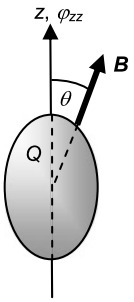


Fig. 7.3 Schematic view of an electric quadrupole Q in a crystalline electric field gradient of amplitude φ_{zz}, in the presence of a magnetic field $\boldsymbol{B}$ applied under angle θ with respect to the symmetry axis z of the crystal

with nuclear spin $I = 1$. The deuteron has a nonspherical charge distribution $\rho(r)$, characterized by an *electric quadrupole moment*

$$eQ = \int \left(3z^2 - r^2\right)\rho(\boldsymbol{r})\,\mathrm{d}^3 r \tag{7.6}$$

of size 0.286 $e\mathrm{fm}^2$. For a spherical charge distribution $\rho(r)$, the quadrupole moment vanishes because for each spherical shell with radius r, the integral of $r^2\rho(r)$ is three times the integral of $z^2\rho(r)$. If the deuteron is located inside a crystal of not too high symmetry (i.e., neither cubic nor tetrahedral), then in the case of rotational symmetry about z, the interaction between the quadrupole moment and the gradient of the crystalline electric field is described by the Hamilton operator

$$\mathrm{H_Q} = \hbar\omega_\mathrm{Q}\left[3(\mathrm{I}_z/\hbar)^2 - (\mathbf{I}/\hbar)^2\right]. \tag{7.7}$$

The *quadrupole interaction frequency* is

$$\omega_\mathrm{Q} = \frac{e^2 q\, Q}{4I\,(2I-1)}, \quad I \geq 1, \tag{7.8}$$

with the component $eq = \partial^2\varphi/\partial z^2$ of the electric field gradient, with electric potential φ; see Sect. 18.2.2. For $I = \frac{1}{2}$, the quadrupole interaction vanishes, $\mathrm{H_Q} = 0$, because $\frac{3}{4}\sigma_z^2 - \frac{3}{4}\mathrm{I} = 0$. In angular momentum representation, the Hamiltonian $\mathrm{H_Q}$ of (7.7) is diagonal, and for $I = 1$ has energies $-2\hbar\omega_\mathrm{Q}$ for $m = 0$ and $+\hbar\omega_\mathrm{Q}$ for $m = \pm 1$. The quadrupole interaction is not sensitive to the sign of m because angular momentum operators enter $\mathrm{H_Q}$ only quadratically.

If, in addition to eq, a magnetic field $\boldsymbol{B}$ is applied in the z–x plane, under an angle θ to the z-axis (see Fig. 7.3), then we must add the magnetic Hamiltonian, which is obtained from $\mathrm{H_M} = \hbar\omega_\mathrm{L}(\mathrm{I}_z/\hbar)$ as

$$\mathrm{H_M} = \hbar\omega_\mathrm{L}\left[(\mathrm{I}_x/\hbar)\sin\theta + (\mathrm{I}_z/\hbar)\cos\theta\right] \tag{7.9}$$

because angular momentum $\mathbf{I}$ transforms under rotation like a vector. For $I = 1$, spin matrices I_x and I_z are given in Sect. 16.3. The Hamiltonian for a simultaneous

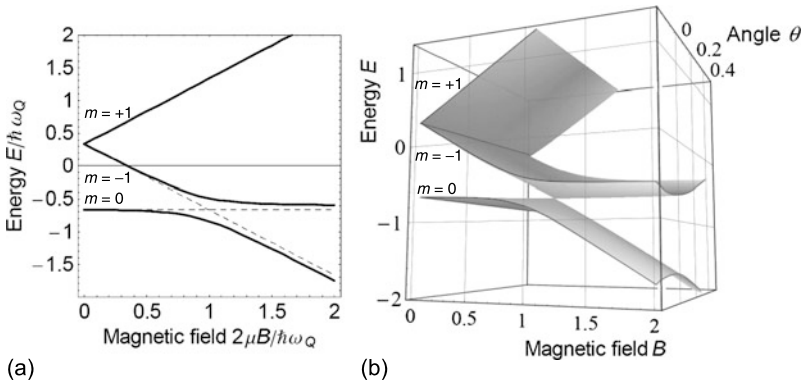


Fig. 7.4 (**a**) Magnetic-field dependence of the energy levels of a spin-1 nucleus under simultaneous magnetic dipole and electric quadrupole interactions, with angle $\theta = 10°$ (see Fig. 7.3). The two lower levels show an avoided crossing at $\omega_L = \omega_Q$. The *dashed lines* show the case $\theta = 0$ with diagonal Hamiltonian. (**b**) Eigenenergies $E(B, \theta)$ of a spin-1 nucleus in the same units as in (**a**), but with angle θ as a second variable. A diabolic point is visible at the *point where the two energy sheets touch each other*. For $\theta = 10°$, the energy levels of figure (**a**) reappear *on the front face of the box*

Zeeman and quadrupole interaction then is

$$\mathrm{H} = \mathrm{H_M} + \mathrm{H_Q},$$

$$\mathrm{H} = \hbar \begin{pmatrix} \omega_L \cos\theta + \omega_Q & \omega_L \sin\theta & 0 \\ \omega_L \sin\theta & -2\omega_Q & \omega_L \sin\theta \\ 0 & \omega_L \sin\theta & -\omega_L \cos\theta + \omega_Q \end{pmatrix}. \tag{7.10}$$

Figure 7.4(a) shows the three energy eigenvalues of H as a function of the field amplitude B, for an angle $\theta = 10°$. This level scheme is obtained by numerical diagonalization of (7.10). Such energy diagrams are needed, for instance, to calculate the frequencies of nuclear magnetic resonance (NMR) transitions between the levels (see Sect. 11.3).

Figure 7.4(b) shows the same eigenvalues as Fig. 7.4(a), but as a function of both the field amplitude B and the angle θ. This angle can easily be varied by rotating the crystal in the magnetic field. One sees a beautiful diabolic point at $\theta = 0$ and $\omega_L = \omega_Q$. At $\theta = 10°$, the avoided level crossing of Fig. 7.4(a) reappears on the front face E vs. B of the box of Fig. 7.4(b). Another avoided crossing is seen on the right face E vs. θ of the box.

7.2 The Adiabatic Theorem

Before we enter our next topic of the geometric phases, we discuss an important result of quantum theory regarding slow and fast changes of a system. In quantum

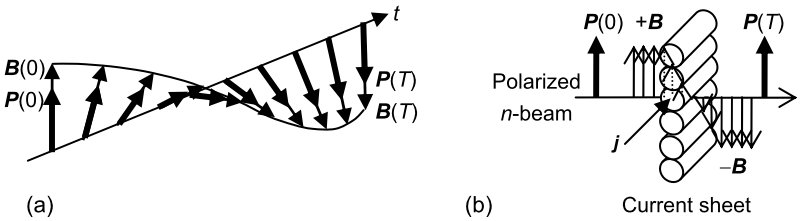


Fig. 7.5 (**a**) Adiabatic following of the polarization $\boldsymbol{P}(t)$ in a slowly varying magnetic field $\boldsymbol{B}(t)$. (**b**) Nonadiabatic transition of $\boldsymbol{P}(t)$ in a field $\boldsymbol{B}(t)$ that varies rapidly, as compared to the Larmor frequency $\omega_{\mathrm{L}} = \gamma B$. The polarization at first is unaffected, $\boldsymbol{P}(t) = \boldsymbol{P}(0)$, and then responds to the new field configuration. Shown is the case of a beam of polarized neutrons passing a thin current sheet of parallel wires, all with the same electric current density $\boldsymbol{j}$

mechanics, like in ordinary mechanics and in everyday life, there is a law of inertia. There are two extreme cases. If external conditions change very rapidly, the system under study at first does not react, and then slowly adapts to the new environment. If external conditions change very slowly, then the system adapts continuously to the evolving conditions. Rapid or slow is meant in relation to the time constants of the system.

An example from classical physics is the temperature T of an external heat bath that is changed at a rate $1/\tau_{\mathrm{ext}} = (1/T)\partial T/\partial t$. The system has a response time $\tau_{\mathrm{syst}} = RC$. The thermal resistance R of the coupling to the heat bath is the inverse of thermal conductivity, the latter being measured in W/K. The heat capacity C of the system is measured in J/K. The change of external conditions is rapid if $1/\tau_{\mathrm{ext}} \gg 1/\tau_{\mathrm{syst}}$, and is slow if $1/\tau_{\mathrm{ext}} \ll 1/\tau_{\mathrm{syst}}$.

In quantum mechanics, these two limiting cases are the subject of the *adiabatic theorem* which we discuss for the example of a beam of polarized neutrons flying along z through a spatially varying magnetic field $\boldsymbol{B}(z)$, seen by the moving neutron as a time dependent field $\boldsymbol{B}(t)$.

For the first case of slowly varying external conditions, Fig. 7.5(a) shows the *adiabatic following* of neutron polarization in a spatially rotating field. Neutron polarization $\boldsymbol{P}$ follows the field direction $\boldsymbol{B}/B$ adiabatically, if the frequency of rotation of $\boldsymbol{B}(t)$ is slow compared to the frequency of Larmor precession of $\boldsymbol{P}(t)$ about $\boldsymbol{B}$, $\omega_{\mathrm{ext}} \ll \omega_{\mathrm{L}}$. In other words, in the adiabatic limit, the neutron polarization $\boldsymbol{P}(t)$ precesses through many cycles before the external field $\boldsymbol{B}(t)$ has changed significantly. For a neutron velocity $\upsilon_n \sim 1000$ m/s, and for one full turn of $\boldsymbol{B}$ over a distance of $L = 0.3$ m, the requirement for such adiabatic following is calculated to $B \gg 2\pi\upsilon_n/\gamma_n L \approx 0.1$ mT, with the neutron gyromagnetic ratio $\gamma_n = 29 \times 2\pi$ MHz/T. To compare, the earth's magnetic field has a typical amplitude of ∼0.05 mT.

In the general case, the theorem states that for a very slow change of the Hamiltonian $\mathrm{H}(t)$, from an initial H_i at time $t = 0$ to a final H_f at time $t = T$, called a *quantum adiabatic process*, the eigenstates of the system at every moment adapt to the momentary Hamiltonian and remain eigenstates of the changing $\mathrm{H}(t)$:

If a system is initially in an eigenstate of the Hamiltonian $\mathrm{H} = \mathrm{H}_i$, after an *adiabatic change* of H it will be in an eigenstate of the final Hamiltonian H_f.

The other extreme is a *nonadiabatic transition* that rapidly changes the Hamiltonian, such that $\omega_{\mathrm{ext}} \gg \omega_{\mathrm{L}}$. In this case, the wave function of the system at first remains unaffected.

If a system is initially in an eigenstate of $\mathrm{H} = \mathrm{H}_i$, then, after a *nonadiabatic change* of H, it will still be in the same eigenstate, which enters the new initial conditions for the system's further evolution under H_f.

This second case is illustrated in Fig. 7.5(b) which shows neutron polarization in a field $\boldsymbol{B}(z)$ that rapidly changes direction from $\boldsymbol{B}$ to $-\boldsymbol{B}$ over a very short distance. This is realized with a current sheet across the neutron beam, with field $+\boldsymbol{B}$ on one side of the sheet, $-\boldsymbol{B}$ on the other. We use the same parameters as before, but for a 300 times shorter flight distance of 1 mm, the thickness of the wires that make up the current sheet. Hence the process is nonadiabatic as long as $B \ll 30$ mT. In the process, the initial polarization $\boldsymbol{P}(0)$ does not follow $\boldsymbol{B}$, but after passage of the current sheet starts its slow precession about the new field direction $-\boldsymbol{B}$. If downstream the field goes over adiabatically into a longitudinal holding field, then the current sheet operates as a neutron spin flipper: if the current is turned on, the final neutron polarization points opposite to the holding field, if the current is turned off, polarization points into the direction of the holding field.

7.3 Geometric Phases

Geometric, or Berry, phases make their appearance in quantum systems that depend on a set of external parameters assembled into a vector $\boldsymbol{B}(t)$, not necessarily a magnetic field, which in the course of time adiabatically changes amplitude and direction. Geometric phases do not depend on the dynamics of the problem, i.e., on the elements of the Hamiltonian, as do the conventional dynamical phases, but only on the topology of the places visited in parameter space. After a sketch of Berry's general derivation, we turn to the case of a spin-$\frac{1}{2}$ particle in a slowly varying magnetic field and show that in the simplest case Berry's phase can be derived from Rabi's spin resonance formula.

7.3.1 Derivation of the Berry Phase

We have seen how a quantum system reacts to adiabatic changes of external parameters. Let, for instance, the tip of the magnetic field vector $\boldsymbol{B}$ move slowly along

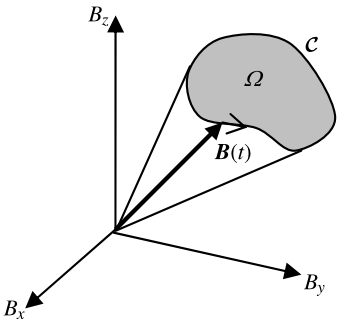


Fig. 7.6 Geometric or Berry phases show up when a quantum system depends on several parameters, here the three components of a magnetic field $\boldsymbol{B}$. Let $\boldsymbol{B}$ slowly change direction and amplitude. We call this an adiabatic excursion of the system along a closed trajectory $\mathcal{C}$ in parameter space. For a spin-$\frac{1}{2}$ system, Berry phases turn out to depend only on the solid angle Ω under which $\mathcal{C}$ is seen from the origin of parameter space

a closed curve $\mathcal{C}$ in $\boldsymbol{B}$-space such that after a time T it comes back to its starting point, as shown in Fig. 7.6. The polarization $\boldsymbol{P}(t)$ of the system follows $\boldsymbol{B}(t)$ adiabatically, as it did in Fig. 7.5(a). Adiabatic again means that the frequency of revolution through the cycle $\mathcal{C}$ is slow compared to the Larmor frequencies of spin precession along $\mathcal{C}$.

Along such a trajectory in parameter space, the eigenstates of the system will develop their usual dynamical phases $\delta_{\pm}(t)$ from (3.50). However, an additional phase appears that does not depend on the strength of the interaction, as does the dynamical phase, but only on the geometric history of the excursion (Berry 1984). To derive Berry's phase law, we start our circuit in a specific eigenstate $\psi_{\mathrm{D}}(0)$ of energy. From the adiabatic theorem, we know that for adiabatic changes of the parameters, the system will stay in the momentary eigenstate $\psi_{\mathrm{D}}(t)$ of the stationary Schrödinger equation,

$$\mathrm{H}(t)\psi_{\mathrm{D}}(t) = E(t)\psi_{\mathrm{D}}(t), \tag{7.11}$$

up to a phase factor $\mathrm{e}^{\mathrm{i}\Phi}$, which evolves in time such that the solution of the time dependent Schrödinger equation is $\psi(t) = \mathrm{e}^{\mathrm{i}\Phi(t)}\psi_{\mathrm{D}}(t)$. Berry assumed that the phase therein is the sum

$$\Phi = \delta + \gamma_{\mathrm{g}} \tag{7.12}$$

of the well known dynamical phase δ and a geometrical phase γ_{g}, resulting from the adiabatic change of $\mathrm{H}(t)$, hence

$$\psi(t) = \mathrm{e}^{\mathrm{i}(\delta+\gamma_{\mathrm{g}})}\psi_{\mathrm{D}}(t). \tag{7.13}$$

The excursion must be adiabatic, and therefore many revolutions of the dynamical phase must take place before a geometric phase develops along path $\mathcal{C}$. Therefore, the dynamical phase is always much larger than the geometric phase,

$$|\delta| \gg |\gamma_{\mathrm{g}}|, \tag{7.14}$$

which makes the detection of a Berry phase look very challenging. We shall see later on that there are ways to isolate γ_g from Φ.

We first retrace the first steps of Berry's general proof before quoting a simple special result that covers a frequently encountered situation, and finally proving this simple result for the particular case of an adiabatic excursion along a circular path $\mathcal{C}$.

When deriving Berry's phase, we have to keep in mind that all quantities ψ, ψ_D, δ, γ_g involved depend on time via the slowly varying n-tuple of parameters $\boldsymbol{B}(t)$. We insert the wave function $\psi(t)$ from (7.13) into the time dependent Schrödinger equation under adiabatic conditions and use (7.11),

$$\dot{\boldsymbol{\psi}}(t) = -(\mathrm{i}/\hbar)\mathrm{H}(t)\psi(t) = -(\mathrm{i}/\hbar)E(t)\psi(t). \tag{7.15}$$

This becomes

$$\dot{\psi} = \big[\mathrm{i}(\dot{\delta} + \dot{\gamma}_g)\psi_D + \dot{\psi}_D\big]\exp\big[\mathrm{i}(\delta + \gamma_g)\big] = -(\mathrm{i}/\hbar)E\psi_D\exp\big(\mathrm{i}[\delta + \gamma_g)\big]. \tag{7.16}$$

Therein, the dynamical phase with $\dot{\delta} = -E/\hbar$ from (3.50) on the left cancels with the term $-E/\hbar$ on the right, and we are left with

$$\mathrm{i}\dot{\gamma}_g\psi_D + \dot{\psi}_D = 0. \tag{7.17}$$

The chain rule of differentiation applied to $\psi_D(\boldsymbol{B}(t))$ gives $\dot{\psi}_D = \dot{\boldsymbol{B}}\cdot\nabla_{\boldsymbol{B}}\psi_D$, hence

$$\mathrm{i}\dot{\gamma}_g\boldsymbol{\psi}_D + \dot{\boldsymbol{B}}\cdot\nabla_{\boldsymbol{B}}\psi_D = 0. \tag{7.18}$$

Multiplication with orthonormal $\psi_D^\dagger$ from the left gives Berry's equation of motion for the geometric phase

$$\dot{\gamma}_g = \mathrm{i}\dot{\boldsymbol{B}}\cdot\big(\psi_D^\dagger\nabla_{\boldsymbol{B}}\psi_D\big). \tag{7.19}$$

We do not pursue this equation any longer but turn right away to an important special result for γ_g of remarkable simplicity.

7.3.2 Excursions in Magnetic-Field Space

For a Hamiltonian of the magnetic dipole type, with matrix elements linear in the components of a vector $\boldsymbol{B}$, Berry obtained a very simple general result.

If the parameter vector $\boldsymbol{B}$ moves on a closed trajectory $\mathcal{C}$ in $\boldsymbol{B}$-space, and if this trajectory is seen from the origin $\boldsymbol{B} = 0$ under a solid angle Ω (see Fig. 7.6), then the geometric phase acquired by a substate with quantum number m is

$$\gamma_g = -m\Omega. \tag{7.20}$$

In our example of a two-dimensional parameter space B_x–B_y, Fig. 7.2, we had a diabolic point at $\boldsymbol{B} = 0$. During an adiabatic excursion along a trajectory $\mathcal{C}$ in the B_x–B_y plane, one sees a field $\boldsymbol{B}_\perp$ of slowly varying direction and amplitude that encircles this diabolic point. When one comes back to the starting point, the solid

angle under which the trajectory as seen from the origin is $\Omega = 2\pi$. From (7.20), for a particle with spin $s = \frac{1}{2}$ moving adiabatically along $\mathcal{C}$, the geometric phase accumulated by the $m = \pm\frac{1}{2}$ states is $\gamma_g = \mp\frac{1}{2}2\pi = \mp\pi$. With $e^{\pm i\pi} = -1$ we have the rule:

> In two dimensions, if the trajectory $\mathcal{C}$ encloses a diabolic point, then the wave functions $\psi_{\pm\frac{1}{2}}$ change sign for each full turn. If $\mathcal{C}$ does not enclose a diabolic point, then the wave functions do not change sign.

This result is independent of the shape of $\mathcal{C}$. Instead of the B_x–B_y parameter space of Fig. 7.2, we could as well have taken the B–θ-space of Fig. 7.4.

Next we treat the special case that $\mathcal{C}$ is a circle in three-dimensional $\boldsymbol{B}$-space such that a particle on its trajectory along $\mathcal{C}$ sees a rotating $\boldsymbol{B}_\perp$-field of constant amplitude $B_\perp$. The cone in Fig. 7.6, swept out by $\boldsymbol{B}(t)$ then has circular cross-sections and opening angle 2θ. In this case, $\mathcal{C}$ is seen from the origin $\boldsymbol{B} = 0$ under a solid angle $\Omega = 2\pi(1 - \cos\theta)$. With $\cos\theta = B_z/B$, this is $\Omega = 2\pi(1 - B_z/B)$. We want to prove that the geometric phase for this simple case indeed is $\gamma_g = \mp\frac{1}{2}\Omega$, as predicted by (7.20).

As we are going to test Berry's phase law via spin precession, we use the well established fact (see Sect. 4.3) that phase shifts of polarization $\boldsymbol{P}$ are twice the phase shifts of wave function ψ. The particle sees the field component $\boldsymbol{B}_\perp$ rotating slowly with an angular frequency ω_{ext}. In a reference frame rotating with this frequency about z, the field component $\boldsymbol{B}_\perp$ is at rest. In this rotating frame, the precession frequency $\omega_0 = \gamma B_z$ about B_z is slightly modified to $\omega_0 - \omega_{ext}$, with $\omega_0 \gg \omega_{ext}$ from the adiabatic condition. Hence B_z, the source of this precession, is reduced to $B_z - \omega_{ext}/\gamma$, as will be discussed in Sect. 11.1, and treated in detail in Sect. 21.1.2.

Altogether, in the rotating frame the magnetic field is transformed to $\boldsymbol{B}_{eff} = (\boldsymbol{B}_\perp, B_z - \omega_{ext}/\gamma)$, with gyromagnetic ratio γ, and the polarization $\boldsymbol{P}(t)$ rotates about this field $\boldsymbol{B}_{eff}$ with a slightly modified Larmor frequency $\omega'_L = \gamma B_{eff}$.

At time T, after one full cycle around the circle $\mathcal{C}$, the total phase accumulated by the wave functions in Hilbert space is

$$2\Phi(T) = -\gamma\sqrt{B_\perp^2 + (B_z - \omega_{ext}/\gamma)^2}\,T - 2\pi, \tag{7.21}$$

where we have subtracted the phase $\omega_{ext}T = 2\pi$ after return to the laboratory frame. As $\omega_{ext} \ll \omega_L$, we can drop the quadratic term ω_{ext}^2/γ^2 and expand the remaining square root to

$$2\Phi(T) \approx -\gamma BT + (B_z/B)\omega_{ext}T - 2\pi. \tag{7.22}$$

The first term is twice the dynamical phase $2\delta = \delta_+ - \delta_- = -\gamma BT$, therefore the second term must be the geometric phase $2\gamma_g$. Hence, with $\omega_{ext}T = 2\pi$ for one complete turn, the geometric phase is given by

$$2\gamma_g = -2\pi(1 - B_z/B) = -\Omega. \tag{7.23}$$

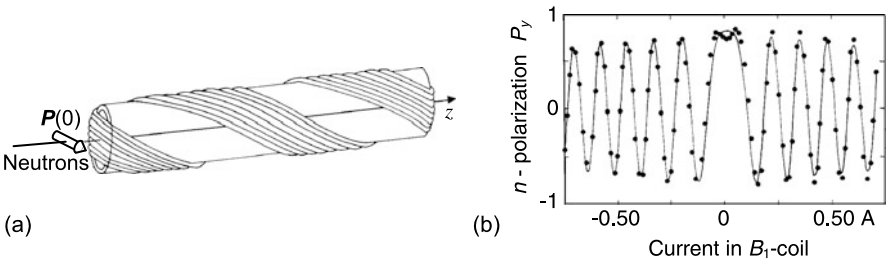


Fig. 7.7 (**a**) The geometric phase can be disentangled from the much larger dynamical phase by letting a beam of polarized neutrons move in the magnetic field of a twisted coil, with the field $\boldsymbol{B}_\perp$ varying in space as in Fig. 7.5(a). (**b**) The spin rotation curve measured with this setup (with a neutron spin analyzer and a neutron detector added) is the neutron polarization $P_y(B) = P_{y0}\cos[2\Phi(B)]$. For sufficiently large field values B, the adiabatic condition is met such that $2\Phi = \gamma BT + 2\gamma_g$, with gyromagnetic ratio γ and flight time T through the field region. The deviation from a pure cosine curve is due to the geometric phase $2\gamma_g$; for details, see text. From Bitter and Dubbers (1987)

This proves (7.20) for a circular trajectory in parameter space. It is noteworthy that (7.21), and with it Berry's result for this simple example, coincides with the phase 2Φ in the oscillating term of Rabi's formula that we shall meet in (11.5), a formula well known in nuclear magnetic resonance.

We describe an experimental study of Berry's phase. We take a magnetic field $\boldsymbol{B}(z)$ that spatially varies along an axis z, in a manner as was shown in Fig. 7.5(a), but for a full turn of $\boldsymbol{B}$. The field is produced by the twisted coil shown in Fig. 7.7(a). A solenoid with axis along z (not shown) generates the longitudinal field component B_z. We let a beam of transversally polarized neutrons of velocity υ_n enter this field region of length L, nonadiabatically by using the current sheet method of Fig. 7.5(b). Seen in the frame of the slowly moving neutron, the tip of the field vector $\boldsymbol{B}(t)$ describes a closed circular path.

When leaving this field region at time $T = L/\upsilon_n$, the neutron's polarization will have acquired a phase factor $P_y = P_{y0}\cos(2\Phi)$, with $\Phi = \delta + \gamma_g$ from (7.13), or

$$P_y = P_{y0}\cos(\gamma BT + \Omega). \tag{7.24}$$

In an experiment with a non-twisted coil, one would obtain a field dependent spin rotation curve with purely dynamical phase 2δ. With the twisted coil of Fig. 7.7(a), with increasing magnitude B of the magnetic field, the phase of the polarization asymptotically reaches $2\delta + 2\gamma_g$ and the wings of the spin rotation curve are shifted outwards by $\approx \pm 2\gamma_g$ at field values B strong enough to meet the adiabatic condition. The resulting spin rotation curve is shown in Fig. 7.7(b). The measurements were repeated for different values of B_z, and with it for different solid angles Ω, and in this way Berry's law $\gamma_g = -m\Omega$ from (7.20) was verified for spin $s = \frac{1}{2}$ particles.

There is a close link between geometric phases and what mathematicians call *parallel transport* (Simon 1983) that we touch upon in the following classical analogue. Take a wooden pointer, a "spear", and successively go through the following exercise.

(i) Start with your arm stretched upwards along z such that the spear, in the horizontal above your head, is on the north pole, pointing forward along y.
(ii) Keep your arm stretched without twisting your wrist, and rotate it forward in southward direction, until you land on the equator, the spear then pointing downwards.
(iii) Sweep your arm horizontally through 90° westwards to the right, following the equator until your arm points sideways along x, the spear's head still pointing down.
(iv) Move your stretched arm upwards to the north pole, again without twisting your wrist. The spear, which at the start pointed forward along y, after this excursion points sideways along x, it has turned clockwise by an angle $\gamma_g = -\frac{1}{2}\pi$. The solid angle Ω enclosed by the movement of your arm is one eighth of 4π, or $\Omega = -\frac{1}{2}\pi = \gamma_g$, as expected from (7.20) for a spin-1 object. You may choose other variants of this exercise and will always find $\Omega = -\gamma_g$.

7.3.3 *Excursions in the Space of Shapes*

Geometric phases make their appearance in many different contexts. We turn to another type of parameter space, which we call the *space of shapes*. Let us regard a standing wave in a resonator. This may be a matter wave, an electromagnetic wave, a sound wave, a water wave, or other. A classical wave of amplitude ψ, constrained by some time independent boundary conditions, obeys the Helmholtz equation $(\nabla^2 + k^2)\psi = 0$. In the absence of a potential, the Helmholtz equation is formally identical to the time independent Schrödinger equation $(\nabla^2 + 2mE/\hbar^2)\psi = 0$. This allows us to limit discussion to classical waves, which automatically includes the corresponding quantum case.

If we slowly change the shape of the resonator along a trajectory $\mathcal{C}$ in its space of shapes, the original pattern $\psi_0(\boldsymbol{x})$ of the standing wave within the resonator will also change. When, after such an excursion, the resonator comes back its original shape, then the wave function should also come back to the original standing-wave pattern $\psi_0(\boldsymbol{x})$, multiplied by some phase factor $e^{i\Phi}$.

In a resonator, many different modes with different eigenfrequencies can usually be excited. With changing shape of the resonator, its frequency spectrum will also change. At certain positions in the resonator's parameter space one should find diabolic points where two frequency surfaces touch each other, in the same way as depicted in Fig. 7.4(b) for two energy surfaces. If the trajectory $\mathcal{C}$ encircles such a degeneracy n times, then its wave-function should acquire a geometric phase $n\gamma_g$ with the *winding number* n.

In a flat microwave resonator made of two parallel horizontal metallic plates enclosed by vertical side walls of arbitrary shape, one can excite the TM or transverse magnetic modes and measure the electric field intensity $|E_z(\boldsymbol{x})|^2$ of the vertical field component with a small antenna without significantly disturbing the field distribution. In this way, one can scan the pattern $|E_z(x, y)|^2$ for various shapes of the resonator (Stöckmann 1999).

As an example, take a flat triangular resonator. Free parameters are height H and segment length P of the triangle, shown in Fig. 7.9. The parameter space (H, P) has two dimensions, and the wave function $E_z(x, y)$ should change sign if $\mathcal{C}$ encloses a

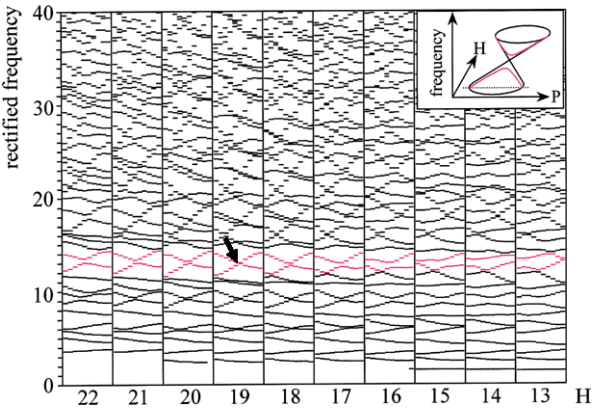


Fig. 7.8 Measured frequency levels for 100 different shapes (H, P) of the triangular resonator shown in the center-left of Fig. 7.9. For each value of the triangle's height H, given in cm on the abscissa of the figure, the segment P ranges from 8 to 17 cm. The *arrow* shows a crossing of two levels, *colored red* near the value 14 of the rectified frequency on the ordinate. When looking at the same levels for different values of H, to *the left and the right of the arrow*, one sees that these display the double-cone structure indicated in *inset*, with a diabolic point near $H_0 \approx 19$ cm, and $P_0 \approx 13$ cm. The resonances are measured in the frequency range from 2 to 7.5 GHz. From Lauber et al. (1994)

diabolic point, and otherwise should keep its sign. First, the microwave spectra were measured at a fixed position of the antenna in the resonator, for 10×10 different values of H and P, with the results shown in Fig. 7.8. If one looks closely, a frequency level crossing is seen between the states no. 13 and no. 14, indicated by the arrow, for H fixed at 19 cm and P varied between 8 and 17 cm. In the neighboring level schemes with lower or higher values of H, this level crossing changes into avoided crossings, as is expected when one moves away from a diabolic point with its typical double cone structure sketched in Fig. 7.8 (inset).

Then space dependent standing-wave patterns $|E_z(x, y)|^2$ were measured for different sets of H and P. Such a mapping takes about one hour for each triangular shape (H, P). Furthermore, these patterns carry no direct phase information. How then can a sign change of field component E_z, oscillating at GHz frequency, ever be detected after successive changes of the values (H, P) along a closed circuit around a diabolic point at (H_0, P_0)?

This can be done in the following way. In the wave patterns shown in Fig. 7.9, regions of high intensity are separated from each other by nodal lines of zero intensity, which can be clearly identified in each pattern. On one side of a nodal line, the sign of E_z must differ from the sign on the other side of the line. This allows us to define the relative phases of the pattern at a given starting point (H_1, P_1) in the vicinity of the diabolic point (H_0, P_0). We arbitrarily paint in yellow color the regions for one sign, and in red color for the other, with the nodal lines remaining black. If we change (H, P) only by small increments, then the phase relation indicated by this coloring scheme can continuously be followed up through the different triangular shapes 1 to 12 of Fig. 7.9. These where chosen to lie on a small ellipse around the

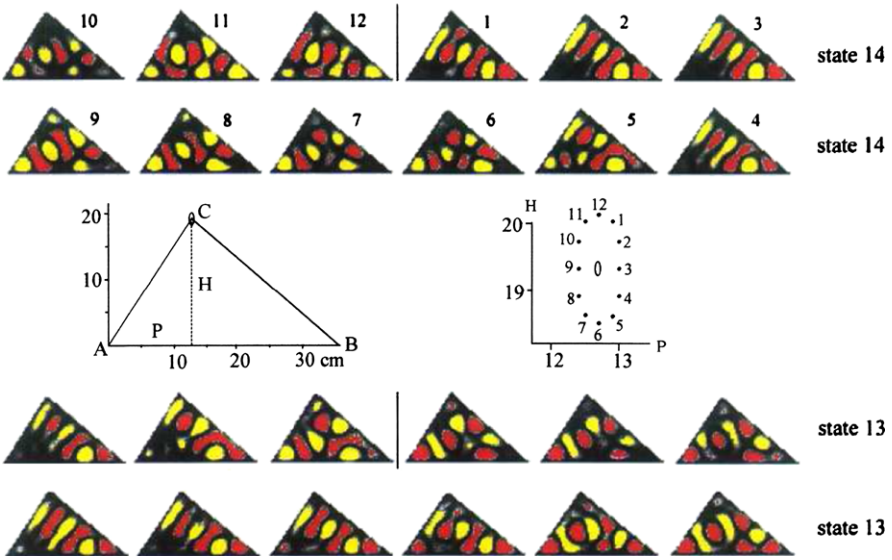


Fig. 7.9 The shape of the microwave resonator used is indicated by the *triangle in the center left*. The *tiny ellipse near corner* C of the triangle is the circuit traversed in $H-P$ space. This circuit, with its 12 different resonator shapes used in the experiment, is shown enlarged in the *center right*, with the position of the diabolic point indicated by the *small shaded area in the center*. The corresponding standing-wave patterns nos. 1 to 12 shown are measured at the two eigenfrequencies nos. 13 and 14 of Fig. 7.8. *Yellow and red colors* indicate different signs of the electric field amplitude, separated by nodal *lines in black*. After one full revolution in parameter space, the amplitude has changed sign, compare patterns 1 and 12. From Lauber et al. (1994)

diabolic point in the space of shapes, see the middle part of Fig. 7.9. Indeed, when we come back to the starting point, near pattern 12 in the figure, the sign of E_z has changed, in accordance with Berry's prediction. Note that in the same experiment, an effective spin-1 diabolic point was studied that did not reproduce Berry's prediction (7.20). For completeness we mention that meanwhile this problem has been resolved by Manini and Pistolesi (2000).

7.3.4 The Aharonov–Bohm Effect

The phase measured in the Aharonov–Bohm effect has some similarities with a geometric phase. Take a two-slit setup as in Fig. 1.1(a), but with the central wire replaced by a long tiny current-carrying solenoid, whose axis is perpendicular to the paper plane of the figure, with a magnetic field inside, but essentially zero magnetic B-field outside the solenoid. An electron's matter wave passes the double slit arrangement created in this way on both sides of the solenoid. There the magnetic field is zero, and therefore one would expect that the B-field inside the impermeable

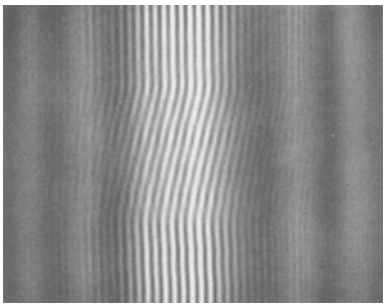


Fig. 7.10 Measurement of the Aharonov–Bohm effect, with a photographic plate used for electron detection. This plate slowly moved vertically behind a horizontal slit in the interference region, such that each horizontal line on the plate is an electron interference pattern. In the *lower and upper parts of the figure*, the electric current in the solenoid was kept constant, and it was linearly varied in the *middle part*. While the interference pattern is seen to shift linearly with the field, its envelope remains unchanged. From Bayh (1962)

solenoid will have no influence on the electron. The electron still senses the presence of the magnetic field $\boldsymbol{B}$ inside the solenoid via its magnetic potential $\boldsymbol{A}$, with $\boldsymbol{B} = \nabla \times \boldsymbol{A}$, which is nonzero also outside the solenoid. The partial electron waves ψ_1 and ψ_2 develop a phase difference when they pass through the two slits on the left and on the right of the solenoid, with $\psi_1 = \mathrm{e}^{\mathrm{i}\Delta\varphi}\psi_2$. This phase difference turns out to be (Aharonov and Bohm 1959)

$$\Delta\varphi = \frac{e}{\hbar}\oint_{\mathcal{C}} \mathbf{A}\cdot\mathrm{d}\mathbf{s}. \tag{7.25}$$

From Stoke's law, this phase shift is proportional to the magnetic flux $\Phi = \int \boldsymbol{B}\cdot \mathrm{d}\boldsymbol{a}$ through the solenoid of area a and can be observed via the B-dependent shift of the interference pattern on the screen shown in Fig. 7.10. In the experiment, the change in magnetic flux needed to create a phase shift $\Delta\varphi = 2\pi$ equaled the predicted value $h/e = 2\Phi_0$ within 4 %, with Planck's constant h. The flux quantum Φ_0 will be discussed in Sect. 13.2.2.

The Aharonov–Bohm phase does not require adiabatic transport as does Berry's phase, but in certain cases both phases coincide, as discussed by Aharonov and Anandan (1987).

7.4 Quantum Chaos

Imagine you have to deal with the spectrum of a system with no clear idea on the relevant underlying interactions governing it. This was the situation physicists met with the spectra of nuclei in the midst of the twentieth century, when the details of the strong interaction still were unknown. For the low-lying states, there were still some regularities in the spectra that could be interpreted in terms of collective

rotations or vibrations. For higher energies, however, one soon arrived at a regime where the spectrum exhibited a chaotic mess without any discernible structure. It was Wigner who proposed the use of the distance of neighboring eigenvalues, $s_n = E_n - E_{n-1}$, as a tool to classify the spectra. He conjectured the following:

For chaotic spectra the level spacings should obey the distribution function

$$p(s) = \frac{1}{2}\pi s \mathrm{e}^{-\pi s^2/4}. \tag{7.26}$$

This is now known as the *Wigner distribution*. In the following years, it became clear that the same distribution is typical for all chaotic systems. Figure 7.11 shows a number of examples, only part of them quantum mechanical in origin.

The distribution shows that there must be a level repulsion in chaotic systems since $p(s)$ has its maximum close to $s = 1$ and disappears for $s \to 0$. This is in clear contrast to a purely random spectrum, with no correlation between the eigenvalues and only constraining the mean level spacing to one. For such a spectrum, a *Poissonian distribution* is expected,

$$p(s) = \mathrm{e}^{-s}, \tag{7.27}$$

showing its maximum for $s = 0$, see the dashed lines in Fig. 7.11. This is the same as, e.g., the distribution of time delays between successive random events like nuclear decay, or rain drops on your head.

To understand these findings, it is helpful to look into the spectral level dynamics in dependence of a suitable external parameter. We met already such a level dynamics in Fig. 7.8, where the spectrum of a triangular microwave billiard was studied in dependence of the position of the triangle's top corner. Another example is shown in Fig. 7.12, where the spectra of a circular and an elliptical billiard of electron matter waves are plotted as a function of the strength of a perpendicularly applied magnetic field B. For $B = 0$ both billiards are integrable, but for $B \neq 0$ the elliptical billiard becomes nonintegrable whereas the circular billiards remains integrable.

Looking into the spectra, one observes an overall similar structure for both systems, apart from one conspicuous difference: whereas for the circular billiard there are a large number of crossings of eigenvalues, for the elliptical case many crossings are converted into avoided crossings, thus shifting the maximum of the level spacing distribution function to nonzero values.

The main fingerprints of chaos in the spectra of chaotic systems, for brevity termed quantum chaos, are thus these avoided crossings as it seems, and this suggests again a reduction to a two-level approximation in the vicinity of a crossing. The Hamiltonian of the system may then be represented by a 2×2 matrix. For the general case, the matrix is hermitian, but for systems with time reversal symmetry, which shall be assumed now, it is always possible to take the matrix symmetric,

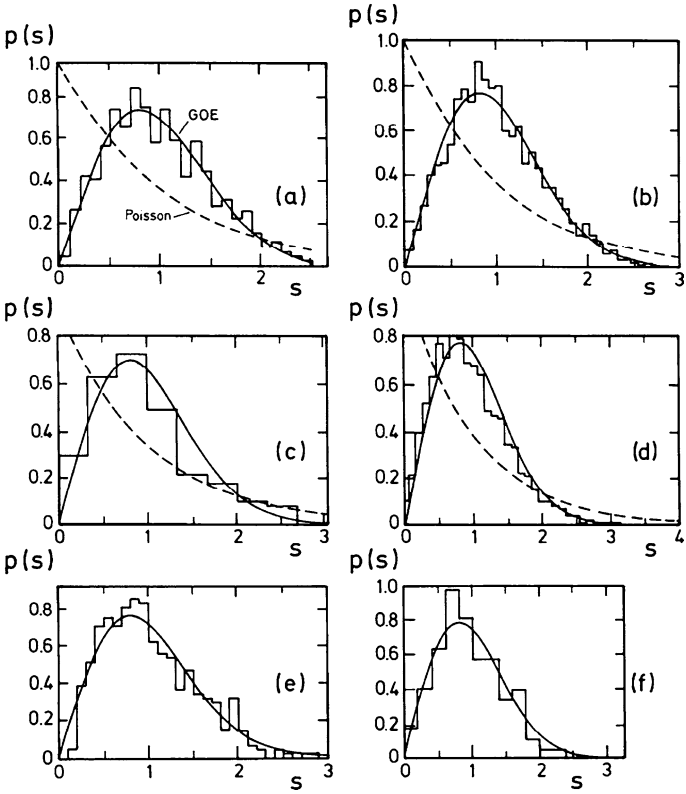


Fig. 7.11 Level spacing distribution for (**a**) a Sinai billiard (a quadratic billiard table with a circular disk inside); (**b**) a hydrogen atom in a strong magnetic field; (**c**) the excitation spectrum of a NO_2 molecule; (**d**) the acoustic resonance spectrum of a Sinai-shaped quartz block; (**e**) the microwave spectrum of a three-dimensional chaotic cavity; (**f**) the vibration spectrum of a quarter-stadium shaped plate. In all cases, a Wigner distribution is found though only in the first three cases the spectra are quantum mechanical in origin. From Stöckmann (1999)

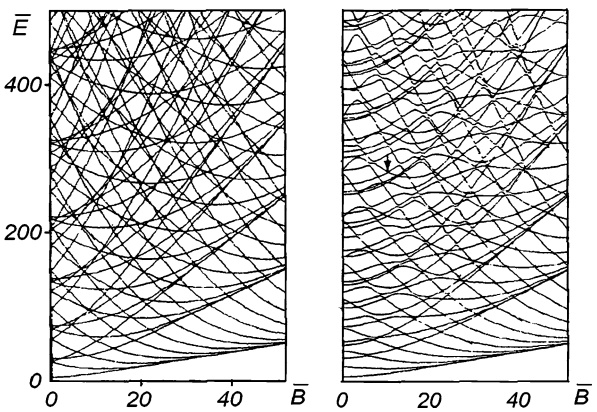


Fig. 7.12 Spectra of the even parity states for (**a**) a circular and (**b**) an elliptic billiard, in dependence of the strength of a perpendicularly applied magnetic field B. In (**b**), most crossings are turned into avoided crossings. From Nakamura and Thomas (1988)

$\mathrm{M}^{\mathrm{T}} = \mathrm{M}$, with real elements,

$$\mathrm{M} = \begin{pmatrix} a & b \\ b & d \end{pmatrix}; \tag{7.28}$$

see, for instance, Haake (2010) and Stöckmann (1999). The eigenvalues of M are

$$E_{\pm} = \frac{1}{2}(a+d) \pm \sqrt{\frac{1}{4}(a-d)^2 + b^2}, \tag{7.29}$$

hence the distance between the two eigenvalues is

$$s = 2\sqrt{\frac{1}{4}(a-d)^2 + b^2}. \tag{7.30}$$

We are interested in the distribution function of s. Here the random matrix approach has proven to be extremely useful in cases where nothing is known about the details of the interaction, or if the system is too complicated to understand its details. One simply takes the matrix elements of M as random numbers. However, one is not completely free in the choice of the distribution function of the matrix elements. First, it is natural to take the matrix elements as uncorrelated, i.e., to assume that the distribution function for the matrix elements factorizes. This assumption can be, and is, occasionally relaxed. However, a second assumption is indispensable: in chaotic systems with no constants of motion apart from the total energy, there is no preferred basis, meaning that the distribution function of the matrix elements of M and TMT^{-1}, where T is an arbitrary orthogonal matrix, are the same. This constraint means that the distribution of matrix elements must be a function of invariants like the trace and determinant. In fact, there is only one distribution function obeying both constraints at the same time, namely,

$$p(a,b,d) = \text{const.} \times \mathrm{e}^{-\mathrm{Tr}(\mathrm{M}^2)/2\sigma} = \text{const.} \times \mathrm{e}^{-(a^2+2b^2+d^2)/2\sigma}. \tag{7.31}$$

The unspecified constant is fixed by the requirement that $p(a,b,d)$ must be normalized. With this distribution function for the matrix elements and with (7.30) for the eigenvalue distance, the distribution of the spacings of eigenvalues can be calculated. The result is, after normalizing the mean level spacing to one, exactly the Wigner distribution (7.26)!

The exact level spacing distribution, obtained for $N \times N$ symmetric random matrices in the limit $N \to \infty$, deviates from the Wigner distribution only by several percent (Haake 2010), showing that, as anticipated, the relevant things really happen at close encounters of the eigenvalues.

The overwhelming evidence for the universality of the level spacing distribution in chaotic systems, as exhibited, e.g., in Fig. 7.11, was the motivation for Bohigas, Giannoni, Schmit to formulate what is known today as the BGS conjecture (Bohigas et al. 1984). The technical jargon skipped, the conjecture states that the universal features of the spectra of chaotic systems, including the level spacing distribution,

should be identical with the corresponding quantities found for ensembles of random matrices. It needed more than 25 years and the efforts of several groups, until the conjecture was finally proven. Details are found in Chap. 10 of Haake (2010).

If this were the whole truth, then the information contained in a chaotic spectrum would be close to zero: according to the BGS conjecture one can, after normalizing the mean level spacing to one, no longer distinguish between the spectrum of a nucleus, a microwave resonator, a vibrating plate, or a matrix embodying just random numbers! Fortunately, there is a complementary approach relating the spectra to the classical dynamics of the system. It goes back to Gutzwiller, who established a relation between the spectra of chaotic systems and its periodic orbits (Gutzwiller 1967). As a consequence, there are long-range correlations in the spectra, which cannot be accounted for by random matrix theory, and which comprise the information on the individual properties of the system. Though this is another story; see, e.g., Stöckmann (1999) for more details.

References

Aharonov, Y., Anandan, J.: Phase change during a cyclic quantum evolution. Phys. Rev. Lett. **58**, 1593–1596 (1987)

Aharonov, Y., Bohm, D.: Significance of electromagnetic potentials in quantum theory. Phys. Rev. **115**, 485–491 (1959)

Bayh, W.: Messung der kontinuierlichen Phasenschiebung von Elektronenwellen im kraftfeldfreien Raum durch das magnetische Vektorpotential einer Wolfram-Wendel. Z. Phys. **169**, 492–510 (1962)

Berry, M.V.: Quantal phase factors accompanying adiabatic changes. Proc. R. Soc. Lond. A **392**, 45–57 (1984)

Bitter, T., Dubbers, D.: Manifestation of Berry's topological phase in neutron spin rotation. Phys. Rev. Lett. **59**, 251–254 (1987)

Bohigas, O., Giannoni, M.J., Schmit, C.: Characterization of chaotic quantum spectra and universality of level fluctuation laws. Phys. Rev. Lett. **52**, 1–4 (1984)

Gutzwiller, M.C.: Phase-integral approximation in momentum space and the bound states of an atom. J. Math. Phys. **8**, 1979–2000 (1967)

Haake, F.: Quantum Signatures of Chaos, 3rd edn. Springer, Berlin (2010)

Lauber, H.-M., Weidenhammer, P., Dubbers, D.: Geometric phases and hidden symmetries in simple resonators. Phys. Rev. Lett. **72**, 1004–1007 (1994)

Manini, N., Pistolesi, F.: Off-diagonal Berry phases. Phys. Rev. Lett. **85**, 3067–3071 (2000)

Nakamura, K., Thomas, H.: Quantum billiard in a magnetic field: chaos and diamagnetism. Phys. Rev. Lett. **61**, 247–250 (1988)

Simon, B.: Holonomy, the quantum adiabatic theorem, and Berry's phase. Phys. Rev. Lett. **51**, 2167–2170 (1983)

Stöckmann, H.-J.: Quantum Chaos—An Introduction. University Press, Cambridge (1999)

Chapter 8
The Coupling of Particles

Abstract To understand matter, we must know how particles interact with each other. When we have a bound quantum system of identical particles, the question whether these particles are bosons or fermions is often more important than the details of their interaction. We therefore first discuss the distinction between these two types of particles. Next the coupling of two spin-$\frac{1}{2}$ systems is described. An illustrative example is the coupling of a proton and an electron, for which we derive the Zeeman effect in the hyperfine splitting of the hydrogen atom.

8.1 Bosons and Fermions

In *classical* statistical mechanics, the *Maxwell–Boltzmann law* gives the distribution of an ensemble of identical but distinguishable particles over the accessible energy states in the case of thermal equilibrium, see (6.25). To arrive at this law, the number Ω_{class} of *microstates* must be calculated, which is the number of different possible microscopic particle configurations corresponding to the observed thermodynamic macrostate of the ensemble. A given configuration of two identical particles, plus the same configuration, but with both particles exchanged, together count as two configurations.

In *quantum* physics, identical particles are indistinguishable from each other, therefore two identical particles and their exchanged configuration count as one single configuration. In a large system of particles, there are many ways of exchanging particles within pairs. Therefore, the number Ω of different particle configurations can be dramatically diminished, as compared to a classical system.

Another important law of quantum physics is the *spin-statistics theorem* whose roots are deeply buried in relativistic quantum field theory. This theorem distinguishes between particles with half-integer spin $s = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \ldots$ called *fermions*, and particles with integer spin $s = 0, 1, 2, \ldots$ called *bosons*.

For an ensemble of fermions, the spin-statistics theorem requires that its total wave function change sign when two particles of the system are exchanged.

D. Dubbers, H.-J. Stöckmann, *Quantum Physics: The Bottom-Up Approach*, Graduate Texts in Physics,
DOI 10.1007/978-3-642-31060-7_8, © Springer-Verlag Berlin Heidelberg 2013

For an ensemble of bosons, the total wave function remains the same when two particles are exchanged.

We say that for fermions the wave function is *antisymmetric* under particle exchange, and for bosons it is *symmetric*. These quantum prescriptions change the classical Maxwell–Boltzmann distribution law into the *Fermi–Dirac* and *Bose–Einstein* distribution laws, respectively. A total system made up of fermions may itself be fermionic or bosonic, depending on whether it contains an odd or even number of fermions. A system of bosons is always bosonic. All quantum states are either bosonic or fermionic, that is, symmetric or antisymmetric under exchange of two particles.

Assume that two identical particles numbered 1 and 2 can each occupy two different states α and β of two-dimensional Hilbert space. In general, α and β can represent a whole set of quantum numbers like $|njm_jlm_lsm_s\rangle$. If particle 1 is in state α, we call it α_1, if it is in state β we call it β_1, with orthonormal α_1, β_1, and the same for α_2, β_2 of particle 2. The basis states of the (noninteracting) two-particle system then are

$$\alpha_1\alpha_2,\ \alpha_1\beta_2,\ \beta_1\alpha_2,\ \beta_1\beta_2. \tag{8.1}$$

The states (8.1) are the four elements of the direct product matrix $\mathrm{A} \equiv (\alpha_1, \beta_1)^{\mathrm{T}}(\alpha_2, \beta_2)$, see (6.45). They are orthonormal basis states of the four-dimensional Hilbert space of pairs of particles because $\mathrm{A}^\dagger\mathrm{A} = 1$ for orthonormal α_i, β_i. For the moment, we treat these states merely as symbols to designate the basis states of the direct-product Hilbert space.

The sequence of the products of states does not matter, that is, $\alpha_2\alpha_1 = \alpha_1\alpha_2$, $\alpha_1\beta_2 = \beta_2\alpha_1$, etc. Therefore, the first basis state $\alpha_1\alpha_2$ stays the same under particle exchange and is symmetric, as is the fourth basis state $\beta_1\beta_2 = \beta_2\beta_1$. The second basis state $\alpha_1\beta_2$ becomes the third state $\alpha_2\beta_1$ under particle exchange, and vice versa. As these two states $\alpha_1\beta_2$ and $\alpha_2\beta_1$ are orthogonal to each other, they cannot be the positive or negative of each other, $\alpha_1\beta_2 \neq \pm\alpha_2\beta_1$, hence they are neither symmetric nor antisymmetric under exchange.

Since quantum states must be either bosonic or fermionic, the basis states in (8.1) must be *symmetrized* or *antisymmetrized*. This can be accomplished by a simple change of basis in the four-dimensional product space. To this end, the states $\alpha_1\beta_2$ and $\beta_1\alpha_2$ are replaced by the antisymmetric and symmetric combinations of basis states

$$\Psi_{\mathrm{A}}(1,2) = \alpha_1\beta_2 - \alpha_2\beta_1 = -\Psi_{\mathrm{A}}(2,1), \tag{8.2a}$$

$$\Psi_{\mathrm{S}}(1,2) = \alpha_1\beta_2 + \alpha_2\beta_1 = +\Psi_{\mathrm{S}}(2,1), \tag{8.2b}$$

while the symmetric basis states $\alpha_1\alpha_2$ and $\beta_1\beta_2$ are left unchanged.

These equations have far reaching consequences. For two identical but distinguishable particles, classical statistics states that the number of configurations for one particle being in state α, the other in state β, is

$$\Omega_{\mathrm{class}} = |\alpha_1\beta_2|^2 + |\alpha_2\beta_1|^2 = 2. \tag{8.3}$$

The same result is obtained for the case that both particles are in the same state α,

$$\Omega_{\text{class}} = |\alpha_1\alpha_2|^2 + |\alpha_2\alpha_1|^2 = 2|\alpha_1\alpha_2|^2 = 2. \tag{8.4}$$

In quantum statistics, the result depends on whether we have two fermions or two bosons. For two identical fermions, (8.2a) gives for $\alpha = \beta$

$$\Omega_{\text{fermion}} = |\alpha_1\alpha_2 - \alpha_2\alpha_1|^2 = 0, \tag{8.5}$$

that is, two fermions cannot occupy the same quantum state. This is known as

Pauli's principle: Two fermions cannot have all quantum numbers the same.

This principle governs the building of the periodic table of elements, by requiring that the electrons in the atomic shell, being fermions, fill the available atomic states one by one, and the same for the building of the atomic nuclei.

For two identical bosons, (8.2a) gives for $\alpha = \beta$

$$\Omega_{\text{boson}} = |\alpha_1\alpha_2 + \alpha_2\alpha_1|^2 = 4|\alpha_1\alpha_2|^2 = 4 = 2\Omega_{\text{class}}, \tag{8.6}$$

which is twice the number of classical configurations from (8.4). Due to this, for a large system of bosons, the number of bosons that occupy a given state, in particular, the energetic ground state of the system, can be enormous.

While fermions strictly segregate, bosons like to congregate.

This tendency for aggregation can be so strong that the bosons condense into the ground state of the quantum system; examples for this will be given in Sect. 12.4.3 on Bose–Einstein condensates, and in Sect. 13.2.1 on superconductivity.

8.2 The Coupling of Spins

The Universe is populated with fermionic elementary particles with spin $s = \frac{1}{2}$, mainly the up and down quarks, the electrons, and the fugitive neutrinos. To understand the bound quantum systems that we are made of, it is important to know how spins couple to each other. We show this for the case of two spin-$\frac{1}{2}$ particles, examples being the magnetic coupling of the two electron spins in an atom, or the strong-interaction coupling of a proton and a neutron in a nucleus. Two spins $\mathbf{S}_1$ and $\mathbf{S}_2$ with spin quantum numbers $s_1 = s_2 = \frac{1}{2}$ couple to a total spin

$$\mathbf{S} = \mathbf{S}_1 + \mathbf{S}_2 \tag{8.7}$$

with total spin quantum number $s = s_1 - s_2 = 0$ or $s = s_1 + s_2 = 1$, from the triangle rule (1.23), while $S_z = S_{1z} + S_{2z}$ has total magnetic quantum numbers $m = m_1 + m_2 = 0$ or ± 1, see (1.24). Couplings of higher angular momenta or couplings of more than two angular momenta will be discussed in Sect. 17.3.1.

We have four two-particle states of type (8.1) for the still uncoupled spins. In bra-ket notation, we write these states as $|s_1 m_1\rangle|s_2 m_2\rangle \equiv |s_1 m_1 s_2 m_2\rangle$, or, knowing that s_1 and s_2 both equal $\frac{1}{2}$, shorthand as $|m_1 m_2\rangle$, or

$$\left|+\tfrac{1}{2}, +\tfrac{1}{2}\right\rangle, \quad \left|+\tfrac{1}{2}, -\tfrac{1}{2}\right\rangle, \quad \left|-\tfrac{1}{2}, +\tfrac{1}{2}\right\rangle, \quad \left|-\tfrac{1}{2}, -\tfrac{1}{2}\right\rangle. \tag{8.8}$$

We then couple the two spins together, for instance, by the magnetic dipole interaction. The operator of total spin $\mathbf{S} = \mathbf{S}_1 + \mathbf{S}_2$ must be an angular momentum operator, too. This means that $\mathbf{S}^2$ and S_z must have simultaneous eigenfunctions $|s, m\rangle$ with eigenvalues $s(s+1)\hbar^2$ and $m\hbar$, respectively, with m going in unit steps from $-s$ to $+s$. Generally, the direct-product basis states $|m_1 m_2\rangle$ are not simultaneous eigenstates of $\mathbf{S}^2$ and S_z, but we can always find linear combinations of them that are. The derivation of the adequate expansion coefficients, called the *Clebsch–Gordan coefficient*s, is difficult for arbitrary spins and will be discussed in Sect. 17.3. For two spin-$\frac{1}{2}$ particles, however, these expansion coefficients are already fixed by the symmetry requirement under the particle exchange operations, (8.2a), (8.2b).

The possible two-particle spin eigenstates $|sm\rangle$ are one *symmetric spin triplet* and one *antisymmetric spin singlet*.

$$|1, +1\rangle = \left|+\tfrac{1}{2}, +\tfrac{1}{2}\right\rangle, \tag{8.9a}$$

$$|1, 0\rangle = \left(\left|+\tfrac{1}{2}, -\tfrac{1}{2}\right\rangle + \left|-\tfrac{1}{2}, +\tfrac{1}{2}\right\rangle\right)/\sqrt{2}, \tag{8.9b}$$

$$|1, -1\rangle = \left|-\tfrac{1}{2}, -\tfrac{1}{2}\right\rangle, \tag{8.9c}$$

$$|0, 0\rangle = \left(\left|+\tfrac{1}{2}, -\tfrac{1}{2}\right\rangle - \left|-\tfrac{1}{2}, +\tfrac{1}{2}\right\rangle\right)/\sqrt{2}. \tag{8.9d}$$

These states $|sm\rangle$ are normalized such that they all have equal weights $\langle sm|sm\rangle = 1$. It remains to be shown that the $|sm\rangle$ are eigenfunctions of total spin, i.e., of S_z and of $\mathbf{S}^2$. The $|sm\rangle$ are eigenfunctions of S_z with eigenvalues $m\hbar$, because $S_z|sm\rangle = (S_{1z} + S_{2z})|sm\rangle = (m_1 + m_2)\hbar|sm\rangle = m\hbar|sm\rangle$. Note that S_{1z} acts only on the states $|s_1 m_1\rangle$ and S_{2z} acts only on the states $|s_2 m_2\rangle$. At the same time, the $|sm\rangle$ are eigenfunctions of $\mathbf{S}^2 = (\mathbf{S}_1 + \mathbf{S}_2)^2$, with eigenvalues $s(s+1)\hbar^2$. This can be shown by using the properties of the spin-$\frac{1}{2}$ operators listed in Sect. 3.1, and the fact that $\mathbf{S}_1$ and $\mathbf{S}_2$ can be applied in either sequence $\mathbf{S}_1\mathbf{S}_2 = \mathbf{S}_2\mathbf{S}_1$, and the same for their individual components. These operators commute because they operate in different Hilbert subspaces.

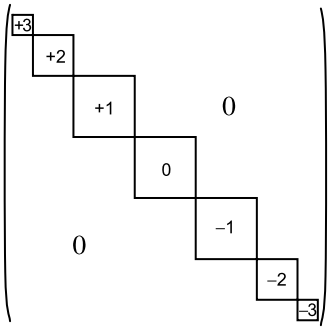


Fig. 8.1 Block-diagonal structure of the table $C(2 \otimes 1)$ of Clebsch–Gordan coefficients $\langle 2m_1 1m_2|jm\rangle$. The subscripts $m = m_1 + m_2$ on the submatrices C_m descend from $m = j_1 + j_2$ down to $m = -j_1 - j_2$. Within each C_m, total angular momentum quantum number j descends from a maximum $j = j_1 + j_2 = 3$ in unit steps down to $j = m$ or $j = |j_1 - j_2| = 1$, whichever is larger

If we collect the $|sm\rangle$ and $|m_1 m_2\rangle$ in one column vector each, then both are related as

$$\begin{pmatrix} |s,m\rangle \\ |1,+1\rangle \\ |1,0\rangle \\ |0,0\rangle \\ |1,-1\rangle \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \sqrt{\frac{1}{2}} & \sqrt{\frac{1}{2}} & 0 \\ 0 & \sqrt{\frac{1}{2}} & -\sqrt{\frac{1}{2}} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} |m_1,m_2\rangle \\ \left|+\frac{1}{2},+\frac{1}{2}\right\rangle \\ \left|+\frac{1}{2},-\frac{1}{2}\right\rangle \\ \left|-\frac{1}{2},+\frac{1}{2}\right\rangle \\ \left|-\frac{1}{2},-\frac{1}{2}\right\rangle \end{pmatrix}. \tag{8.10}$$

The elements of the 4×4 unitary matrix in this equation are the Clebsch–Gordan coefficients for the coupling of two spin-$\frac{1}{2}$ particles, written as $\langle \frac{1}{2} m_1 \frac{1}{2} m_2 | 1m\rangle$. The matrix, which we call $C(\frac{1}{2} \otimes \frac{1}{2})$, has block-diagonal form.

For general angular momenta $\mathbf{J}_1 + \mathbf{J}_2 = \mathbf{J}$, (8.10) reads

$$|j,m\rangle = \sum_{m_1,m_2} |j_1 m_1 j_2 m_2\rangle \langle j_1 m_1 j_2 m_2 | jm\rangle, \tag{8.11}$$

with the Clebsch–Gordan coefficients $\langle j_1 m_1 j_2 m_2 | jm\rangle$.

Due to $m = m_1 + m_2$, the sum in this equation is only over one of the two quantum numbers m_1 or m_2.

Tables $C(j_1 \otimes j_2)$ of the Clebsch–Gordan coefficients have block-diagonal form when grouped with respect to total magnetic quantum number m. The submatrices C_m are arranged along the diagonal as shown for $C(2 \otimes 1)$ in Fig. 8.1. For more details, see Sect. 17.3.1.

8.3 Example: Hyperfine Structure

To give an instructive example of the use of the Clebsch–Gordan coefficients, we derive the Zeeman effect in the *hyperfine interaction* of the ground state of the hydrogen atom. In an atom, nuclear and electron angular momenta $\mathbf{I}$ and $\mathbf{J}$ couple to $\mathbf{F} = \mathbf{I} + \mathbf{J}$. The Hamiltonian that describes this spin coupling must be a scalar derived from the two spin vectors $\mathbf{I}$ and $\mathbf{J}$, and must depend on both vectors in the same way. The simplest scalar formed from two vectors is their scalar product, and this is indeed the Hamiltonian for the spin–spin hyperfine coupling,

$$\mathrm{H_{IJ}} = V(\boldsymbol{r})\mathbf{I} \cdot \mathbf{J}. \tag{8.12}$$

For a free atom, V cannot depend on the relative orientation of the electron and the nucleus, but only on their distance r, with $V(r) \propto 1/r^3$ for the interaction of two permanent dipoles. Averaging over all distances r gives the *hyperfine interaction constant* $A = \langle V \rangle \propto \langle r^{-3} \rangle = \int \psi^\dagger(r) r^{-3} \psi(r)\,\mathrm{d}r$, which can be calculated analytically for the ground state of the hydrogen atom, with $\psi(r) \propto \exp(-r/a_0)$ and *Bohr radius* $a_0 \approx 0.05$ nm.

The hyperfine interaction leads to a splitting of the energy levels of the atomic state. To calculate this splitting, we are free in our choice of basis states and take the eigenstates of total angular momentum $|Fm_F\rangle$ from (8.9a)–(8.9d). These are the eigenstates of the operator $\mathbf{I} \cdot \mathbf{J}$, as we see when we expand $\mathbf{F}^2 = (\mathbf{I} + \mathbf{J})^2$ and resolve for $\mathbf{I} \cdot \mathbf{J} = \frac{1}{2}(\mathbf{F}^2 - \mathbf{I}^2 - \mathbf{J}^2)$. All operators on the right are diagonal in the $|Fm_F\rangle$ basis, which is true by definition for $\mathbf{F}^2$ with eigenvalues $\langle Fm_F|\mathbf{F}^2|Fm_F\rangle = F(F+1)\hbar^2$. The operators $\mathbf{J}^2$ and J_z act only on the $|Jm_J\rangle$ part, and $\mathbf{I}^2$ and I_z act only on the $|Im_I\rangle$ part of the product states $|Jm_J\rangle|Im_I\rangle$, and all these states are orthonormal. Their energy eigenvalues therefore are $\langle Fm_F|\mathbf{J}^2|Fm_F\rangle = J(J+1)\hbar^2$ and $\langle Fm_F|\mathbf{I}^2|Fm_F\rangle = I(I+1)\hbar^2$, respectively. (In atomic physics, the quantum numbers I, J, F are usually written in capital letters.)

The energy eigenvalues of the hyperfine Hamiltonian $\mathrm{H_{IJ}} = A\mathbf{J} \cdot \mathbf{I}$ then are

$$E = \frac{1}{2} A\hbar^2 \big[F(F+1) - J(J+1) - I(I+1)\big]. \tag{8.13}$$

For the case $I = J = \frac{1}{2}$ of the $^2\mathrm{S}_{\frac{1}{2}}$ hydrogen ground state, these eigenvalues are $E_+ = \frac{1}{4}A\hbar^2$ for $F = 1$ and $E_- = -\frac{3}{4}A\hbar^2$ for $F = 0$. This splitting is shown at the origin $B = 0$ of Fig. 8.2. The weighted mean of E_+ and E_- is not changed by the interaction because energy is conserved in the closed system: $\langle E \rangle = [3 \times \frac{1}{4} + 1 \times (-\frac{3}{4})]A\hbar^2 = 0$, with the number of magnetic substates $2F + 1 = 3$ and 1 as statistical weights.

With $A\hbar^2 = 5.9 \times 10^{-6}$ eV, the transition between these states lies at a frequency of 1.42 GHz corresponding to a wavelength of 21.1 cm. In spite of an extremely small probability of 3×10^{-15} s^{-1} for spontaneous transitions between these levels, this wavelength is dominant in radio astronomy, due to the large number of hydrogen atoms in the Universe.

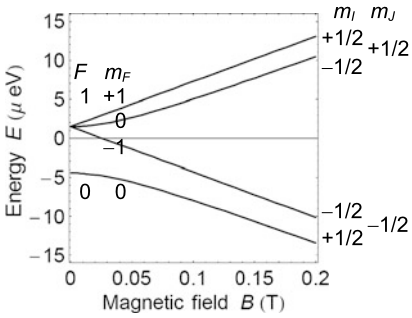


Fig. 8.2 Zeeman effect of the hyperfine interaction of the hydrogen atom in its ground state. At low field, electron and proton spins are coupled to total angular momentum $\mathbf{F} = \mathbf{I} + \mathbf{J}$, with quantum numbers $F = 1$ and $F = 0$. At high field, the hyperfine interaction is progressively decoupled

An external magnetic field $\boldsymbol{B}$ leaves the electron's spatial wave function $\psi(r)$ essentially unaffected, therefore A is independent of $\boldsymbol{B}$, and we can simply add the Zeeman Hamiltonians for the electron and the nucleus to $\mathrm{H_{IJ}}$ and write

$$\mathrm{H} = A\mathbf{J}\cdot\mathbf{I} - (g_J\mu_\mathrm{B}/\hbar)\mathbf{J}\cdot\boldsymbol{B} - (g_I\mu_\mathrm{N}/\hbar)\mathbf{I}\cdot\boldsymbol{B}. \tag{8.14}$$

Therein $\mu_\mathrm{N} = \mu_\mathrm{B} m_e/m_p = \mu_\mathrm{B}/1836 \ll \mu_\mathrm{B}$ from (2.6) is the nuclear magneton. The g-factor is $g_J = -2.002$ for the electron, and $g_I = 5.6$ for the proton.

Having established the Hamiltonian, we next must find its matrix elements. Here the direct product states $|m_J\rangle|m_I\rangle \equiv |m_J m_I\rangle$ from (8.8) are preferably chosen as basis. For the first part $\mathrm{H_{IJ}} = A\mathbf{J}\cdot\mathbf{I} = A(\mathrm{J}_x\mathrm{I}_x + \mathrm{J}_y\mathrm{I}_y + \mathrm{J}_z\mathrm{I}_z)$, we again make use of (3.9a)–(3.9c) for $I = J = \frac{1}{2}$ and obtain

$$\mathrm{J}_x\mathrm{I}_x\left|+\tfrac{1}{2},+\tfrac{1}{2}\right\rangle = \mathrm{J}_x\left|+\tfrac{1}{2}\right\rangle\mathrm{I}_x\left|+\tfrac{1}{2}\right\rangle = \frac{1}{4}\hbar^2\left|-\tfrac{1}{2},-\tfrac{1}{2}\right\rangle, \tag{8.15a}$$

$$\mathrm{J}_y\mathrm{I}_y\left|+\tfrac{1}{2},+\tfrac{1}{2}\right\rangle = \mathrm{i}^2\frac{1}{4}\hbar^2\left|-\tfrac{1}{2},-\tfrac{1}{2}\right\rangle = -\frac{1}{4}\hbar^2\left|-\tfrac{1}{2},-\tfrac{1}{2}\right\rangle, \tag{8.15b}$$

$$\mathrm{J}_z\mathrm{I}_z\left|+\tfrac{1}{2},+\tfrac{1}{2}\right\rangle = \frac{1}{4}\hbar^2\left|+\tfrac{1}{2},+\tfrac{1}{2}\right\rangle. \tag{8.15c}$$

The contributions (8.15a) and (8.15b) cancel each other, and $|+\frac{1}{2},+\frac{1}{2}\rangle$ from (8.15c) is an eigenstate of $\mathrm{H_{IJ}}$ with eigenvalue $\frac{1}{4}A\hbar^2$.

With the matrix elements of $\mathrm{H_{IJ}}$ arranged in the same way as for the Clebsch–Gordan coefficients in (8.10), this eigenvalue $\frac{1}{4}A\hbar^2$ is the first element in the upper left corner with $m_F = m_J + m_I = 1$, and all other elements of the first row of matrix $\mathrm{H_{IJ}}$ vanish. Because $\mathrm{H_{IJ}}$ is hermitian, the same is also true for the first column.

Due to inversion symmetry, the Hamiltonian should be the same when we invert all signs of the magnetic quantum numbers m, therefore the matrix $\mathrm{H_{IJ}}$ is symmetric with respect to its antidiagonal, and the matrix element in the lower right corner is also $\frac{1}{4}A\hbar^2$. Symmetry considerations can save us much work. For the elements of the central 2×2 matrix with $m_F = 0$, a similar procedure gives $\langle+\frac{1}{2},-\frac{1}{2}|\mathrm{H_{IJ}}|+\frac{1}{2},-\frac{1}{2}\rangle = -\frac{1}{4}A\hbar^2$ for the diagonal elements, and $\langle+\frac{1}{2},-\frac{1}{2}|\mathrm{H_{IJ}}|-\frac{1}{2},+\frac{1}{2}\rangle = \frac{1}{4}A\hbar^2$ for the off-diagonal elements.

The matrix elements of the Zeeman terms in (8.14) are the easier part of our exercise. We choose the quantization axis z along $\boldsymbol{B}$, such that $\mathbf{J}\cdot\boldsymbol{B} = \mathrm{J}_z B_z$ and $\mathbf{I}\cdot\boldsymbol{B} = \mathrm{I}_z B_z$, which adds matrix elements $\langle m_J m_I|\mathrm{J}_z B_z|m_J m_I\rangle = m_J\hbar B_z$ and $\langle m_J m_I|\mathrm{I}_z B_z|m_J m_I\rangle = m_I\hbar B_z$ to the diagonal of H. Collecting the pieces, we arrive at the matrix of the hyperfine Hamiltonian (8.14) in the m_J, m_I representation

$$\mathrm{H} = \begin{pmatrix} \frac{1}{4}A\hbar^2 + \mu_+ B & 0 & 0 & 0 \\ 0 & -\frac{1}{4}A\hbar^2 + \mu_- B & \frac{1}{2}A\hbar^2 & 0 \\ 0 & \frac{1}{2}A\hbar^2 & -\frac{1}{4}A\hbar^2 - \mu_- B & 0 \\ 0 & 0 & 0 & \frac{1}{4}A\hbar^2 - \mu_+ B \end{pmatrix}. \tag{8.16}$$

Here $\mu_+ = \frac{1}{2}(g_J\mu_\mathrm{B} + g_I\mu_\mathrm{N}) = 1.0027\mu_\mathrm{B}$ and $\mu_- = \frac{1}{2}(g_J\mu_\mathrm{B} - g_I\mu_\mathrm{N}) = 0.9997\mu_\mathrm{B}$.

Diagonalization of this matrix gives the eigenenergies of the hydrogen atom, with the first subscript on E for the sign of m_J, the second for m_I,

$$F = 1: \quad E_{++} = \frac{1}{4}A\hbar^2 + \mu_+ B \qquad (m_F = +1), \tag{8.17a}$$

$$E_{+-} = -\frac{1}{4}A\hbar^2 + \sqrt{\frac{1}{4}A^2\hbar^4 + \mu_-^2 B^2} \qquad (m_F = 0), \tag{8.17b}$$

$$E_{--} = \frac{1}{4}A\hbar^2 - \mu_+ B \qquad (m_F = -1), \tag{8.17c}$$

$$F = 0: \quad E_{-+} = -\frac{1}{4}A\hbar^2 - \sqrt{\frac{1}{4}A^2\hbar^4 + \mu_-^2 B^2} \qquad (m_F = 0), \tag{8.17d}$$

This is the *Breit–Rabi formula* for $I = J = \frac{1}{2}$.

Figure 8.2 shows the Zeeman effect in the ground state hyperfine structure of the hydrogen atom. At low field the $\mathrm{H_{IJ}}$ term dominates, with $|Fm_F\rangle$ as eigenstates, and the usual Zeeman splitting for $F = 1$ and $F = 0$ is seen. One says that F and m_F are *good quantum numbers* at low field. The E_{++} and E_{--} curves with $m_F = \pm 1$ are linear for all values of B, while the E_{+-} and E_{-+} curves with $m_F = 0$ repel each other at intermediate B. This is not surprising, as E_{+-} and E_{-+} are obtained by diagonalization of a 2×2 matrix with nonzero off-diagonal elements, and therefore show an avoided crossing. At $B = 0$ the $m_F = 0$ states are the symmetric and antisymmetric superpositions $(|+\frac{1}{2}, -\frac{1}{2}\rangle \pm |-\frac{1}{2}, +\frac{1}{2}\rangle)/\sqrt{2}$ of (8.9b), (8.9d). At high magnetic field, the Zeeman term in (8.14) dominates, and the $|Jm_J\rangle|Im_I\rangle$ become eigenstates of the system, with the good quantum numbers m_J and m_I. The levels $E_{++}(B)$, $E_{+-}(B)$ and the levels $E_{-+}(B)$, $E_{--}(B)$ then run almost parallel to each other.

Chapter 9
"Spooky Action at a Distance"

Abstract Since the early days of quantum theory, some disturbing features of quantum physics have lured in the background. Take two particles that have interacted in the past, but meanwhile are so far separated from each other that an interaction is no longer possible. Then the results of measurements on one particle are affected instantaneously by measurements on the other far distant particle. Hidden parameters that might explain this strange entanglement have been ruled out by experiments whose outcomes violate the Bell inequalities, and we conclude that quantum mechanics is a nonlocal theory.

9.1 Quantum Entanglement

The rules of quantum physics are mathematically well defined and make clear predictions, if only statistical, on the possible results of measurements. These rules have been verified experimentally for many systems with high precision. The measured and calculated g-factors of elementary particles, for example, agree to 10 significant decimal digits (see Sect. 4.1). Many physicists take a pragmatic point of view and say that this is all they require from a good theory. Quantum physics, however, has some truly disturbing features.

> As shown in many experiments, particles that have interacted with each other in the past remain *entangled*, in the sense that the results of measurements on one particle depend on what measurements are carried out at the same time on the other particle.

This holds even if both particles meanwhile are so remote from each other that they are no longer causally connected to each other. Strange, is it not? The results of measurements on the particles remain strongly *correlated* even at large distance. God not only plays dice, contrary to the famous statement of Einstein, but two dice thrown simultaneously at places far distant from each other can be strongly correlated.

The "unspeakables" of quantum physics come out most clearly, if one regards angular momentum and its components. When unpolarized atoms have passed a

D. Dubbers, H.-J. Stöckmann, *Quantum Physics: The Bottom-Up Approach*, Graduate Texts in Physics,
DOI 10.1007/978-3-642-31060-7_9,

Stern–Gerlach filter with magnetic field along axis z, half of the atoms (in one partial beam, Fig. 2.3) are polarized along $+z$, the other half along $-z$. Does this mean that the incoming unpolarized beam has its atoms to equal parts aligned along $\pm z$? Not at all: If we rotate the entire system by a certain angle about the beam axis, we rightly expect from isotropy of space that the beam pattern observed behind the Stern–Gerlach magnet rotates by the same angle (just rotate Fig. 2.3 to the left or the right). The beam then is split in direction of the new field axis z', with half of the atoms having their spins pointing along $+z'$, the other half along $-z'$. The two partial beams again are polarized to 100 %, but now with respect to z'. From this we conclude that the notion of spin up and spin down with respect to some axis z cannot be a property of the incoming unpolarized beam. How should the atoms in the beam know along which axis the Stern–Gerlach magnet far downstream is oriented? The notion of spin up and spin down must be imposed onto the atoms during their interaction with the Stern–Gerlach apparatus.

In another setting (see Sect. 6.3), we started with a beam polarized to 100 % along $+z$ and found that, after passage of a transversely oriented Stern–Gerlach filter, spin components along $-z$ showed up that were not at all present in the incoming beam. Particles with spin certainly are not like marbles to which the rules of set theory apply.

On these questions Einstein, Podolsky, and Rosen (1935) (later EPR) wrote a famous paper titled "Can quantum mechanical descriptions of physical reality be complete?" EPR's text is difficult reading, and we first quote the more compact arguments of Bohm and Aharonov (1957) that build on EPR's paper, but refer to spin components where EPR referred to position and momentum.

Let an isolated molecule A_2 with zero angular momentum spontaneously decay into two equal atoms each with spin $s = \frac{1}{2}$,

$$A_2 \to A + A, \tag{9.1}$$

which then fly apart. Total angular momentum is conserved and hence zero also after the decay of the molecule. The Schrödinger equation supplies the wave function of the total system, which at all times is the antisymmetric singlet state

$$\psi_{\text{total}} = \sqrt{\frac{1}{2}}\left(\left|+\tfrac{1}{2}, -\tfrac{1}{2}\right\rangle - \left|-\tfrac{1}{2}, +\tfrac{1}{2}\right\rangle\right) \tag{9.2}$$

from (8.9d). However, with regard to the spin directions of the two separate atoms, quantum physics makes only statistical predictions: 50 % spin up, 50 % spin down. To know more, we must do spin measurements on the individual atoms, using, for instance, Stern–Gerlach filters on both ends of the setup shown in Fig. 9.1. Each filter has two output channels, one for spin up, one for spin down, with respect to the individual Stern–Gerlach axis. These channels are equipped with suitable atom detectors.

Let the two atoms be meanwhile so remote from each other that measurements done on one atom can have no influence on the other atom, and let us measure the

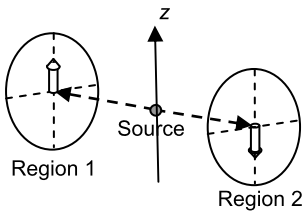


Fig. 9.1 Experimental setup for the measurement of EPR correlations of spin components. The source is a molecule with zero angular momentum, which disintegrates (*dashed arrows*) into two equal atoms with opposite spins (*block arrows*), which are registered in time coincidence after they have passed spin analyzers installed in regions 1 and 2, oriented parallel and antiparallel to z

spin component S_z of the atoms flying to region 1. If we find atom 1 in the spin up state, then quantum theory says that the total wave function of the combined system collapses to the state

$$\psi_{\text{total}+} = \left|+\tfrac{1}{2}, -\tfrac{1}{2}\right\rangle. \tag{9.3}$$

From this we know instantaneously that atom 2 is spin down state. If, instead, in our measurement we find atom 1 in spin down state, the system collapses to

$$\psi_{\text{total}-} = -\left|-\tfrac{1}{2}, +\tfrac{1}{2}\right\rangle \tag{9.4}$$

and atom 2 is instantaneously known to be in spin up state. Hence, quantum theory predicts that results of measurements in region 1 are perfectly anticorrelated with the results of measurements in region 2, independent of the distance the atoms have been traveling in opposite directions. This prediction is confirmed by coincidence measurements in a setup of the type shown in Fig. 9.1. Every time one atom is found spin up, the other inevitably is found spin down, and vice versa. Hence we know with certainty the component S_z also of atom 2, even if we do not measure it.

Let us next measure not S_z but the component S_x of an atom flying to region 1. Based on the same arguments, quantum theory then allows us to know with certainty the component S_x of atom 2, again without measuring it. As we assumed that measurements on atom 1 have no influence on the state of atom 2, we conclude that both S_z and S_x simultaneously are genuine properties of atom 2, measureable with arbitrary precision, given sufficient counting statistics in the experiment. This, however, is at variance with the uncertainty principle, which says (see Sect. 10.2) that different components of angular momentum are *incompatible* quantities that cannot simultaneously be known with certainty. This is the basic conundrum of quantum entanglement.

We saw that random quantum events occurring in different, causally not connected regions of space evidently can be perfectly correlated, due to some unknown *nonlocal* effects. If quantum theory is right, which seems to be the case in view of its high predictive power, then there are two alternatives. Either quantum theory is a conventional *local* theory, but necessarily incomplete because it cannot explain these seemingly nonlocal effects. Or quantum theory is a complete theory, and we

must accept its weird nonlocal character. This is all that needs to be said before proceeding to experimental tests of these two alternatives. Still, let us expand a bit more on the theme of entanglement.

In everyday life, it is not unusual that independent events are correlated. Traffic accidents in Paris and Berlin are independent events, but show significant correlations. Accident rates are high in both cities at night, on weekends, in winter time, etc. Correlations between events do not necessarily imply causality in the sense of cause and effect. Instead, correlations often are due to a common influence. The accident rates not only depend on the spatial variable $\boldsymbol{x}$ for Paris or $\boldsymbol{x}'$ for Berlin, but also on additional variables $\lambda = \{\text{visibility, traffic density, temperature, etc.}\}$. Therefore, the number of accidents $N(\boldsymbol{x}, \lambda)$ in Paris and $N(\boldsymbol{x}', \lambda)$ in Berlin are correlated via the set of *hidden parameters* λ.

If quantum theory is *not* a complete theory, it must be completed by such hidden parameters, which in some unknown way are shared by both atoms of our setup and mimic the nonlocal effects of quantum physics. In practice, it is generally extremely difficult to specify how these hidden parameters should act such as to produce the collapse of the wave function. For successive measurements, even on a small quantum system, the number of hidden parameters needed would be enormous. Still, for a long time it was thought that the existence of hidden variables in quantum physics, unlikely as they might be, could not be falsified.

In everyday life, there also is nothing mysterious about an instantaneous knowledge on distant objects, inferred from looking at nearby objects, even for *space-like separations* of the objects, in the language of special relativity. When, after a long journey, a scientist opens the trunk of his car and sees it empty, he knows immediately that his suitcase is still in the garage. No superluminous signal transmission from the garage to the scientist is needed for this conclusion. However, things are different in quantum physics. There events like clicks in a spin up or spin down counter necessarily are unpredictable statistical processes, the result of which remains uncertain until the moment of measurement.

EPR in their original paper did not use spin components as incompatible quantities, but the observables position and momentum. As these have continuous spectra, EPR's derivation is slightly longer, and we do not repeat it here. Instead, we reproduce their arguments why they thought that quantum theory should be considered as incomplete. For a physical theory to be complete, EPR require a one-to-one correspondence between the elements of the theory and the elements of the world that the theory describes. In their own words: "every element of the physical reality must have a counterpart in the physical theory". How can such an *element of reality* be defined? The authors propose "*If, without in any way disturbing the system, we can predict with certainty* (*i.e., with probability equal to unity*) *the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity*".

EPR base their argument on a proof by contradiction. The uncertainty principle of quantum mechanics states that position and momentum of a particle cannot have simultaneous reality in the above sense, these quantities are incompatible. EPR then assume that quantum theory *is* a complete theory. Based on this, they present as

a counter example the thought experiment of Fig. 9.1, in which positions and momenta of atoms (instead of their spin components) can be measured with arbitrary precision. Then, in the same way as for the spin components, by making measurements solely on atom 1, one can know with certainty both position and momentum of the remote atom 2, without in any way disturbing the latter. However, that which really exists in the region of atom 2 should not depend on what measurement is carried out in the region of atom 1, no reasonable definition of reality could be expected to permit this, so EPR. Therefore, position and momentum of atom 2 must have had definite values all along the atom's trajectory. Position and momentum therefore are elements of reality simultaneously and at all times, contrary to the tenets of the uncertainty principle. The assumption that quantum mechanics is a complete theory leads to a contradiction and therefore is wrong.

EPR did not question the correctness of quantum theory, but its completeness. They concluded that quantum theory could not be the full story, unless one believes in what Einstein later dismissed as *spooky action at a distance*, today called a nonlocal interaction. EPR summarize: "While we have thus shown that the wave function does not provide a complete description of the physical reality, we left open the question of whether or not such a description exists. We believe, however, that such a theory is possible." History has defied this belief, as we shall see in the following, and EPR's conclusion cannot be upheld.

Where does EPR's argument fail? As seen from today, EPR's definition of an "element of physical reality" quoted above (in italics as in the original) looks much like a killer criterion directed against the basic tenets of quantum physics. This, in essence, is also the main complaint in Bohr's (1935) response to EPR. Bohr insisted that the uncertainties of quantum theory should be included as a new fact in the definition of an "element of reality", if only to provide room for the new and undisputed laws of quantum physics.

9.2 Bell's Inequalities

Bell (1964) showed that the question of existence or nonexistence of hidden variables in quantum physics can be decided by way of experiment. To this end, one must measure correlations not only along one axis z, but must do a series of measurements with the two Stern–Gerlach magnets oriented along different axes.

Let the spin of atom 1 be measured in the direction of unit vector $\boldsymbol{a}$, and the spin of atom 2 in the direction of unit vector $\boldsymbol{b}$, see Fig. 9.2(a), by using an appropriately oriented Stern–Gerlach filter at each end of the setup of Fig. 9.1. To search for correlations, one registers the coincidence rates in two detectors positioned behind the analyzers 1 and 2, by means of an electronic AND module. Correlations are measured with the detectors installed on the same output channel of the Stern–Gerlach filters, and anticorrelations are measured with the detectors on opposite output channels.

We recall from (5.15a) and (5.15b) that, if an atom is spin up with respect to one axis, then the spin up and spin down probabilities $p(\alpha)$ and $q(\alpha) = 1 - p(\alpha)$ with

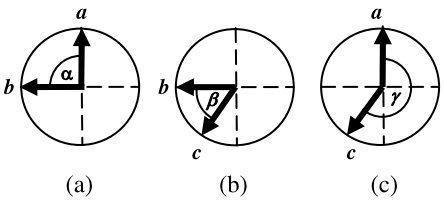


Fig. 9.2 Experimental configurations for tests of Bell's inequality, with views along the decay axis of Fig. 9.1. Coincidence rates are measured for three configurations of the spin analyzer, whose directions are chosen from a set of unit vectors $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$. The coincidence counting rates for these configurations violate Bell's inequality

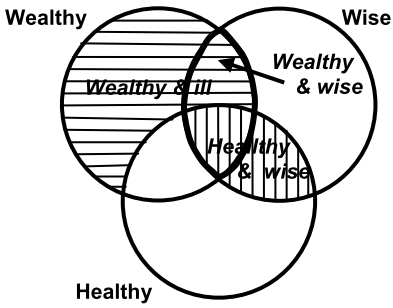


Fig. 9.3 Bell's inequality from set theory: "The number of the wealthy wise (*thick framing*) is smaller than the number of the wealthy ill (*horizontal pattern*) plus the number of the healthy wise (*vertical pattern*)"

respect to another axis, under angle α to the first, is

$$p(\alpha) = \cos^2 \frac{1}{2}\alpha, \qquad q(\alpha) = \sin^2 \frac{1}{2}\alpha. \tag{9.5}$$

To prove the nonexistence of hidden parameters in such systems, Bell also used the method of contradiction. He assumed that hidden variables *do* account for the peculiar predictions of quantum theory. Under this assumption, he then showed that an inequality must be respected regarding the coincidence rates p or q, as registered for different configurations of the two Stern–Gerlach magnets The assumption of hidden variables then leads to a contradiction, because counter examples exist for which quantum mechanics violates this inequality. If such a violation of Bell's inequality is verified in experiment, then this proves that quantum mechanics cannot have hidden parameters.

We give a simple proof that there are cases in quantum theory where Bell's inequality is violated, see also Wigner (1970). To this end, we first make a digression. Many of us want to be wealthy, healthy, and wise. A law from set theory says:

> The number of the wealthy wise is smaller than the number of the wealthy ill plus the number of the healthy wise.

This is evident from Fig. 9.3 because the first group of people is a subset of the other two groups combined. Then we apply this result to three successive atomic-spin correlation experiments, with the fields of the Stern–Gerlach filters 1 and 2 pointing into directions chosen from the three unit vectors $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$, with angles α, β, and γ between them, as shown in Fig. 9.2(a)–(c). We then substitute the above numbers of people N by the coincidence rates p or q from (9.5) for these three configurations:

$$N(\text{wealthy} \wedge \text{wise}) \to p(\alpha) \text{ for correlation } (\boldsymbol{a}, \boldsymbol{b}), \tag{9.6a}$$

$$N(\text{healthy} \wedge \text{wise}) \to p(\beta) \text{ for correlation } (\boldsymbol{c}, \boldsymbol{b}), \tag{9.6b}$$

$$N(\text{wealthy} \wedge \text{not healthy}) \to q(\gamma) \text{ for anticorrelation } (\boldsymbol{a}, \boldsymbol{c}). \tag{9.6c}$$

The symbol $\wedge$ stands for logical AND. Correlation $(\boldsymbol{a}, \boldsymbol{b})$ measures the coincidence rate $p(\alpha)$ with both spins up with respect to $\boldsymbol{a}$ and $\boldsymbol{b}$, and anticorrelation $(\boldsymbol{a}, \boldsymbol{c})$ measures the coincidence rate $q(\gamma)$ with opposite spins with respect to $\boldsymbol{a}$ and $\boldsymbol{c}$. For this set of measurements, the healthy–wealthy–wise inequality then gives

$$p(\alpha) \le p(\beta) + q(\gamma). \tag{9.7}$$

Taking p and q from (9.5), we obtain for the three correlations:

Bell's inequality

$$\sin^2 \frac{1}{2}\alpha \le \sin^2 \frac{1}{2}\beta + \cos^2 \frac{1}{2}\gamma. \tag{9.8}$$

In quantum physics, this inequality is violated for certain combinations of angles α, β, γ. For instance, if field directions $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ of the Stern–Gerlach filters lie in a plane, with $\alpha = 90°$, $\beta = 45°$, $\gamma = 225°$, as shown in Fig. 9.2(a)–(c), we find

$$\sin^2 45° \le \sin^2 22.5° + \cos^2 112.5°, \tag{9.9a}$$

$$0.5 \le 0.146 + 0.146 = 0.293. \tag{9.9b}$$

This is not what we have learned in school.

The above derivation, however, does not refer to hidden variables. The main achievement of Bell was to make a unique link from these correlations to the nonexistence of hidden parameters. We do not reproduce this proof, which is given in modern textbooks on quantum physics; see, for instance, Griffiths (2004), Sakurai (1985), or Liboff (2002).

The violation of Bell's inequality has been established in numerous experiments. Most experiments, starting with Freedman and Clauser (1972), used a two-photon cascade from an excited atomic $J = 0$ state: $A^* \to A + 2$ photons, with subsequent analysis of the polarization of the emitted photons. Dimer decay $Na_2 \to Na + Na$, our example from (9.1), and positronium decay $e^+e^- \to 2\gamma$, have also been used.

The strange correlations show up even if the orientation of the polarizers is chosen at random while the photons are still on their way to their respective detectors, which excludes any hypothetical information exchange at subluminal speed between region 1 and region 2, see Aspect et al. (1982).

The observed nonlocal effects cannot be used to transmit information between regions 1 and 2 at superluminal speed because correlations in Bell-type experiments are established only after the counts from regions 1 and 2 have been brought together and tested for temporal coincidence. Spin measurements in only one of the two regions always produce random results.

As an outlook: In the years since Bell’s discovery, the field of quantum correlations has considerably progressed, and various new forms of Bell’s theorem have emerged. A most interesting correlation is based on what is called a GHZ state, short for Greenberger, Horne, and Zeilinger. In a GHZ state, three spin-$\frac{1}{2}$ particles from a common source are entangled with each other. For certain measurements on three spin components of the three particles, quantum physics predicts a perfect anticorrelation where hidden variable theories would predict a perfect correlation, and one single count measured in the anticorrelation channel therefore decides in favor of quantum mechanics, with no counting statistics involved as in the earlier two-particle correlation experiments, see Greenberger et al. (1990).

Evidently, for a system of two separated atoms that have met in the past, one cannot speak of the state of one atom separately because the two-particle state cannot be expressed as the product of two independent one-particle states, and we must accept that quantum reality requires a nonlocal theory.

References

Aspect, A., Dalibard, J., Roger, G.: Experimental test of Bell’s inequalities using time-varying analyzers. Phys. Rev. Lett. **49**, 1804–1807 (1982)

Bell, J.S.: On the Einstein Podolsky Rosen paradox. Physics **1**, 195–200 (1964)

Bohm, D., Aharonov, Y.: Discussion of experimental proof for the paradox of Einstein, Rosen, and Podolsky. Phys. Rev. **108**, 1070–1076 (1957)

Bohr, N.: Can quantum-mechanical description of physical reality be considered complete? Phys. Rev. **48**, 696–702 (1935)

Einstein, A., Podolsky, B., Rosen, N.: Can quantum-mechanical description of physical reality be considered complete? Phys. Rev. **47**, 777–780 (1935)

Freedman, S.J., Clauser, J.F.: Experimental test of local hidden-variable theories. Phys. Rev. Lett. **28**, 938–941 (1972)

Greenberger, D.M., Horne, M.A., Shimony, A., Zeilinger, A.: Bell’s theorem without inequalities. Am. J. Phys. **58**, 1131–1143 (1990)

Griffiths, D.J.: Introduction to Quantum Mechanics, 2nd edn. Prentice Hall, Upper Saddle River (2004)

Liboff, R.: Introductory Quantum Mechanics, 4th edn. Addison-Wesley, Boston (2002)

Sakurai, J.J.: Modern Quantum Mechanics. Addison-Wesley, Boston (1985)

Wigner, E.P.: On hidden variables and quantum mechanical probabilities. Am. J. Phys. **38**, 1005–1009 (1970)

Chapter 10
The Heisenberg Equation of Motion

Abstract We turn to an alternative description of quantum theory, equivalent to Schrödinger's. Rather simple reasoning led Heisenberg to an equation of motion based on the commutation relations of the operators involved. These commutation relations are uniquely linked to the corresponding uncertainty relations. From Heisenberg's equation, we derive a compact and pictorial equation for magnetic spin precession, valid for any spin quantum number. When applied to the expectation value of the spin operator, this equation turns out to be identical to the classical precession equation of the spinning top. Additional terms that account for spin relaxation lead to the empirical Bloch equations widely used in spin resonance and in quantum optics.

10.1 Matrix Mechanics

Heisenberg discovered the matrix description of quantum matrix mechanics in 1925, shortly before Schrödinger found his wave equation. Heisenberg's aim was to develop a description of atomic processes based exclusively on relations between observable quantities, which at the time was a revolutionary approach. The relevant atomic observables in such a description are, first of all, the measured atomic optical spectral lines at angular frequencies Ω_{nm}. They obey the *Ritz combination principle*, Fig. 10.1,

$$\Omega_{nm} = \Omega_{nk} + \Omega_{km}. \tag{10.1}$$

In the spectrum of the hydrogen atom, for instance, the frequencies Ω_{13} and Ω_{12} of the first two lines of the Lyman series ($n = 1$) are related to the frequency Ω_{23} of the first Balmer line ($n = 2$) as $\Omega_{13} = \Omega_{12} + \Omega_{23}$. This follows directly from the empirical *Rydberg formula* $\Omega_{nm} \propto 1/n^2 - 1/m^2$, $m < n$. Based on his work on stellar spectra, Ritz had conjectured that (10.1) be valid for atomic spectra in general.

From the *photoelectric effect*, Einstein had deduced that light of angular frequency Ω is made up of photons each of energy

$$E_{\text{photon}} = \hbar\Omega. \tag{10.2}$$

D. Dubbers, H.-J. Stöckmann, *Quantum Physics: The Bottom-Up Approach*, Graduate Texts in Physics,
DOI 10.1007/978-3-642-31060-7_10, © Springer-Verlag Berlin Heidelberg 2013

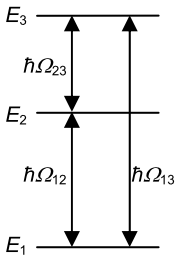


Fig. 10.1 The Ritz combination principle $\Omega_{13} = \Omega_{12} + \Omega_{23}$ for three atomic energy levels

This lead Bohr to postulate the existence of atomic quantized energy levels E_n with

$$\hbar\Omega_{nm} = E_n - E_m. \tag{10.3}$$

Heisenberg associated these angular frequencies with time dependent oscillatory atomic amplitudes

$$X_{nm} = x_{nm}\mathrm{e}^{-\mathrm{i}\Omega_{nm}t}. \tag{10.4}$$

These X_{nm} can be arranged into a matrix $\mathrm{X} = \{X_{nm}\}$. Its time derivative $\dot{X}_{nm} = -\mathrm{i}\Omega_{nm}x_{nm}\exp(-\mathrm{i}\Omega_{nm}t) = -\mathrm{i}\Omega_{nm}X_{nm}$ has the elements, from (10.3),

$$\dot{X}_{nm} = -\frac{\mathrm{i}}{\hbar}(E_n - E_m)X_{nm}. \tag{10.5}$$

The process of atomic radiative transitions must be governed by some Hamiltonian H, which in some way must be connected with this equation of motion of the amplitude matrix X. To find this connection, we write the Hamiltonian as a diagonal matrix $\mathrm{H_D}$ with the energies E_n arranged on the diagonal. We first multiply X from the right with the energy matrix $\mathrm{H_D}$. The product matrix $\mathrm{XH_D}$ and its elements differ from the product matrix $\mathrm{H_DX}$ and its elements by

$$\mathrm{XH_D} - \mathrm{H_DX} = \begin{pmatrix} E_1X_{11} & E_2X_{12} & \dots \\ E_1X_{21} & E_2X_{22} & \dots \\ \dots & \dots & \dots \end{pmatrix} - \begin{pmatrix} E_1X_{11} & E_1X_{12} & \dots \\ E_2X_{21} & E_2X_{22} & \dots \\ \dots & \dots & \dots \end{pmatrix}, \tag{10.6a}$$

$$(\mathrm{XH_D} - \mathrm{H_DX})_{nm} = (E_n - E_m)X_{nm}. \tag{10.6b}$$

In order that the elements X_{nm} of the amplitude matrix X have the oscillatory behavior postulated in (10.5), X must obey the matrix equation

$$\dot{\mathrm{X}} = -\frac{\mathrm{i}}{\hbar}(\mathrm{XH_D} - \mathrm{H_DX}), \tag{10.7}$$

where the dot means the time derivative of the elements of the matrix.

Next we write the Hamiltonian matrix with respect to some other basis, $\mathrm{H} = \mathrm{TH_DT^\dagger}$, as in Sect. 5.1, with a time independent unitary matrix T. Transforming (10.7) with $\mathrm{A} = \mathrm{TXT^\dagger}$ and $\dot{\mathrm{A}} = \mathrm{T\dot{X}T^\dagger}$ and inserting $\mathrm{T^\dagger T} = \mathrm{I}$ wherever needed gives $\dot{\mathrm{A}} = -(\mathrm{i}/\hbar)[(\mathrm{TXT^\dagger})(\mathrm{TH_DT^\dagger}) - (\mathrm{TH_DT^\dagger})(\mathrm{TXT^\dagger})]$, or $\dot{\mathrm{A}} = (\mathrm{i}/\hbar)(\mathrm{HA} - \mathrm{AH})$:

The *Heisenberg equation* of motion of an operator A is

$$\dot{\mathrm{A}} = \frac{\mathrm{i}}{\hbar}[\mathrm{H}, \mathrm{A}]. \tag{10.8}$$

The expression $[\mathrm{H}, \mathrm{A}] = \mathrm{HA} - \mathrm{AH}$ is the *commutator* of H and A. Here $\dot{\mathrm{A}}$ is the total derivative of A, where we have assumed that H is conservative and does not depend explicitly on time, so that $\partial \mathrm{H}/\partial t = 0$. For operators A that depend explicitly on time, we must add the partial time derivative $\partial \mathrm{A}/\partial t$ to the right-hand side of (10.8).

Heisenberg derived this equation at an age when he did not even know what a matrix is. His method then was generalized by postulating that every measurement of a physical quantity A is an operation on the system under study that can be described by a mathematical operator A. As stated before, the possible outcomes of measurements on A are its eigenvalues a_n.

In everyday life, many operations are noncommutative. Getting up in the morning in state ψ, it makes a difference for the further evolution of ψ whether the first operation A of putting on socks is followed by operation B of putting on shoes, to wind up in state $\psi' = \mathrm{BA}\psi$, or whether the reverse order is followed, hence $\mathrm{AB}\psi \neq \mathrm{BA}\psi$. In the evening, the inverse operations A^{-1} and B^{-1} in reverse order lead back to $\psi = \mathrm{A}^{-1}\mathrm{B}^{-1}\psi'$.

Heisenberg's equation can be applied to the operator of any observable. In Sect. 3.2, we had seen that any two-state quantum operator can be written in terms of the Pauli matrices. For later use, it is useful to know the commutator of the Pauli matrices. From (3.11) we know that $\sigma_x\sigma_y = \mathrm{i}\sigma_z$, and $\sigma_y\sigma_x = -\mathrm{i}\sigma_z$, etc. Thus $\sigma_x\sigma_y - \sigma_y\sigma_x = 2\mathrm{i}\sigma_z$, plus cyclic permutations of x, y, z, or

$$[\sigma_x, \sigma_y] = 2\mathrm{i}\sigma_z, \qquad [\sigma_y, \sigma_z] = 2\mathrm{i}\sigma_x, \qquad [\sigma_z, \sigma_x] = 2\mathrm{i}\sigma_y \tag{10.9}$$

and

$$\sigma_x\sigma_y + \sigma_y\sigma_x = 0, \quad \text{plus cyclic permutations.} \tag{10.10}$$

We see that the Pauli matrices *anticommute*. Similar spin commutation relations hold for a general angular momentum operator $\mathbf{J}$,

$$[\mathrm{J}_x, \mathrm{J}_y] = \mathrm{i}\hbar\mathrm{J}_z, \quad \text{plus cyclic permutations.} \tag{10.11}$$

We recall the cross product between two ordinary vectors $(\boldsymbol{A} \times \boldsymbol{B})_z = A_x B_y - A_y B_x$, etc., and write the commutation relations $\mathrm{J}_x\mathrm{J}_y - \mathrm{J}_y\mathrm{J}_x$, etc., for the components of $\mathbf{J}$ as

$$\mathbf{J} \times \mathbf{J} = \mathrm{i}\hbar\mathbf{J}. \tag{10.12}$$

If we write the Schrödinger equation in matrix form from the very beginning, as we did in Sect. 5.1, then it is only a unitary transformation away from the Heisenberg equation, as we shall show in the following. We derive Schrödinger's equation

from Heisenberg's equation (10.8). Let a time dependent observable $A(t)$ be represented by an operator $\mathrm{A_H}$ that obeys Heisenberg's equation. Let $\mathrm{A_S}$ be the corresponding Schrödinger operator, with matrix representation as defined at the beginning of Sect. 3.2. Expectation values are independent of representation, therefore $\langle \mathrm{A_S} \rangle(t) = \langle \mathrm{A_H}(t) \rangle$.

The operator $\mathrm{A_H}$ differs from $\mathrm{A_S}$ at most by a unitary transformation $\mathrm{U}(t)$,

$$\mathrm{A_H}(t) = \mathrm{U}^\dagger(t)\mathrm{A_S}\mathrm{U}(t). \tag{10.13}$$

In a conservative system, the Schrödinger operator $\mathrm{A_S}$ itself is time independent, and the time dependence of its expectation value $\langle \mathrm{A_S} \rangle(t) = \psi^\dagger(t)\mathrm{A_S}\psi(t)$ resides exclusively in the state vector $\psi(t)$. The time derivative of (10.13) then is

$$\dot{\mathrm{A}}_\mathrm{H} = \dot{\mathrm{U}}^\dagger \mathrm{A_S}\mathrm{U} + \mathrm{U}^\dagger \mathrm{A_S}\dot{\mathrm{U}}. \tag{10.14}$$

On the other hand, the Heisenberg equation for A, with (10.13) inserted, gives

$$\dot{\mathrm{A}}_\mathrm{H} = \frac{\mathrm{i}}{\hbar}\big[\mathrm{H}\big(\mathrm{U}^\dagger \mathrm{A_S}\mathrm{U}\big) - \big(\mathrm{U}^\dagger \mathrm{A_S}\mathrm{U}\big)\mathrm{H}\big]. \tag{10.15}$$

The last two equations agree if $\mathrm{U}(t)$ obeys the Schrödinger equation

$$\dot{\mathrm{U}} = -\frac{\mathrm{i}}{\hbar}\mathrm{H}\mathrm{U} \tag{10.16}$$

in its normal and its conjugate transpose form $\dot{\mathrm{U}}^\dagger = (\mathrm{i}/\hbar)\mathrm{H}\mathrm{U}^\dagger$. If $\mathrm{U}(t)$ obeys the Schrödinger's equation, so must $\psi(t) = \mathrm{U}(t)\psi(0)$, with constant $\psi(0)$. We see that $\mathrm{U}(t)$ is nothing but the time evolution operator (6.1) for the solution $\psi(t)$ of Schrödinger's equation. All steps in the above derivation are reversible; therefore,

The Schrödinger equation and the Heisenberg equation are equivalent.

While being equivalent mathematically, from a physics point of view the Heisenberg equation appears to be more fundamental for the following reasons:

- For many observables there is a direct way from Heisenberg's equation to the corresponding equations of motion of classical mechanics. As an example (another example will be given in Sect. 10.3), we apply Heisenberg's equation to the momentum operator $\mathbf{p}$,

$$\dot{\mathbf{p}} = \frac{\mathrm{i}}{\hbar}[\mathrm{H}, \mathbf{p}]. \tag{10.17}$$

With $\mathrm{H} = \mathbf{p}^2/2M + V(\boldsymbol{r})$ and $\mathbf{p} = (\hbar/\mathrm{i})\boldsymbol{\nabla}$ we write the same equation for the expectation values as

$$\langle \dot{\mathbf{p}} \rangle = \frac{\mathrm{i}}{\hbar}\psi^\dagger \left[\frac{\hbar^2 \boldsymbol{\nabla}^2}{2M} + V(\mathbf{r}), \frac{\hbar}{\mathrm{i}}\boldsymbol{\nabla} \right] \psi = -\psi^\dagger \boldsymbol{\nabla} V(\mathbf{r})\psi = \langle -\boldsymbol{\nabla} V \rangle \tag{10.18}$$

where we used $[\nabla^2, \nabla] = 0$ and $[V, \nabla]\psi = -[\nabla, V]\psi$, and conclude:

From Heisenberg's equation we arrive at *Newton's law* $\langle \dot{\mathbf{p}} \rangle = -\langle \nabla V \rangle$.

- The Heisenberg equation directly displays the connection between symmetries and conservation laws. $[\mathrm{H}, \mathrm{A}] = 0$ is equivalent to $\mathrm{AHA}^\dagger = \mathrm{H}$, which means that H is invariant under the transformation A, and from the Heisenberg equation then follows $\dot{\mathrm{A}} = 0$, and A must be time independent, i.e., conserved:

If an operator A commutes with the Hamiltonian H, then the observable A is a *constant of the motion*.

- The commutator on the right-hand side of (10.8) is gives birth to the uncertainty relation, as we will see in the following section.

Though Heisenberg's equation appears to be more fundamental, Schrödinger's equation often is simpler to handle.

10.2 Commutation Relations and Uncertainty Principle

Heisenberg's uncertainty principle states that there are certain incompatible pairs of observables A and B that cannot both be measured simultaneously with arbitrary precision, a well known example being a one-dimensional system which obeys the uncertainty relation for position and momentum $\Delta \mathrm{x} \Delta \mathrm{p}_x \geq \frac{1}{2}\hbar$, with variances as defined in (3.40). If the width x of a beam of light is limited by a slit, then the transverse momentum p_x of the photons increases, and with it the divergence of the beam, a phenomenon called *diffraction*.

Uncertainty relations are known also in classical physics where they hold for any two functions that are Fourier transforms of each other. For the variables time and frequency, the duration Δt of a signal and the bandwidth $\Delta\omega$ of its frequency spectrum obey $\Delta\omega\Delta t \geq 1$. The same holds for a function oscillatory in space, where we have $\Delta k_x \cdot \Delta x \geq 1$ for the position uncertainty Δx and spectral width Δk_x of the component k_x of the wave vector $\boldsymbol{k}$ of magnitude $k = 2\pi/\lambda$, with wavelength λ.

The often cited mysteries of quantum physics lie not so much in the uncertainty relations, relations that are well known to any audio or video technician, but rather in the fact that sometimes particles behave like waves, as in diffraction, and sometimes waves behave like particles, as in the Compton effect. This *wave–particle duality* is based on those strange relations $E = \hbar\omega$ and $\boldsymbol{p} = \hbar\boldsymbol{k}$.

Above we remarked that the commutator $[\mathrm{A}, \mathrm{B}]$ provides a measure of uncertainty, and we want to prove the following statement:

The uncertainty product of two operators A and B is related to the expectation value of the commutator as

$$\Delta\mathrm{A}\Delta\mathrm{B} \geq \frac{1}{2}\left|\langle[\mathrm{A},\mathrm{B}]\rangle\right|. \tag{10.19}$$

We study this inequality in terms of the deviations of A and B from their expectation values

$$\delta\mathrm{A} = \mathrm{A} - \langle\mathrm{A}\rangle, \qquad \delta\mathrm{B} = \mathrm{B} - \langle\mathrm{B}\rangle. \tag{10.20}$$

The right-hand side of (10.19) is not affected by this shift because $\langle\mathrm{A}\rangle$ and $\langle\mathrm{B}\rangle$ are simple numbers such that

$$[\delta\mathrm{A},\delta\mathrm{B}] = [\mathrm{A},\mathrm{B}]. \tag{10.21}$$

For the left-hand side, we go back to the definition (3.39) of variance, namely,

$$(\Delta\mathrm{A})^2 = \left\langle(\delta\mathrm{A})^2\right\rangle, \qquad (\Delta\mathrm{B})^2 = \left\langle(\delta\mathrm{B})^2\right\rangle. \tag{10.22}$$

As a next step, we know that ordinary vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ obey the inequality $|\boldsymbol{a}||\boldsymbol{b}| \geq |\boldsymbol{a}\cdot\boldsymbol{b}|$, with the scalar product defined as $\boldsymbol{a}\cdot\boldsymbol{b} = |\boldsymbol{a}||\boldsymbol{b}|\cos\theta$, and angle θ between $\boldsymbol{a}$ and $\boldsymbol{b}$. The same holds for the expectation values of operators and is known as the *Cauchy–Schwarz inequality* (the proof is found in all major textbooks). Hence the square of the left-hand side of (10.19) obeys

$$\left\langle(\delta\mathrm{A})^2\right\rangle\left\langle(\delta\mathrm{B})^2\right\rangle \geq \langle\delta\mathrm{A}\delta\mathrm{B}\rangle^2. \tag{10.23}$$

We notice that the product of any two quantities x and y can be expanded as

$$\mathrm{xy} = \frac{1}{2}(\mathrm{xy}+\mathrm{yx}) + \frac{1}{2}(\mathrm{xy}-\mathrm{yx}). \tag{10.24}$$

If x and y are hermitian operators (like δA and δB), i.e., $\mathrm{x} = \mathrm{x}^\dagger$ and $\mathrm{y} = \mathrm{y}^\dagger$, then, from $(\mathrm{xy})^\dagger = \mathrm{y}^\dagger\mathrm{x}^\dagger$, the first term in this expansion is hermitian, $(\mathrm{xy}+\mathrm{yx})^\dagger = \mathrm{y}^\dagger\mathrm{x}^\dagger + \mathrm{x}^\dagger\mathrm{y}^\dagger = \mathrm{yx}+\mathrm{xy}$, and the second term is *antihermitian*, namely, $\mathrm{xy}-\mathrm{yx} = -(\mathrm{xy}-\mathrm{yx})^\dagger$. We expand the product $\delta\mathrm{A}\delta\mathrm{B}$ in (10.23) in this way,

$$\delta\mathrm{A}\delta\mathrm{B} = \frac{1}{2}(\delta\mathrm{A}\delta\mathrm{B}+\delta\mathrm{B}\delta\mathrm{A}) + \frac{1}{2}(\delta\mathrm{A}\delta\mathrm{B}-\delta\mathrm{B}\delta\mathrm{A}). \tag{10.25}$$

While the eigenvalues of hermitian operators are all real, those of antihermitian operators are all purely imaginary. The same is true for their expectation values, i.e., the weighted averages of the eigenvalues. Hence we call the first term on the right of (10.25) the real number $s = \frac{1}{2}\langle\delta\mathrm{A}\delta\mathrm{B}+\delta\mathrm{B}\delta\mathrm{A}\rangle$, and the second term the imaginary number $\mathrm{i}d = \frac{1}{2}\langle\mathrm{AB}-\mathrm{BA}\rangle$ from (10.21). Using $|s+\mathrm{i}d|^2 = s^2+d^2$, we write (10.23) as

$$(\Delta\mathrm{A})^2(\Delta\mathrm{B})^2 \geq s^2+d^2 = s^2 + \frac{1}{4}\langle\mathrm{AB}-\mathrm{BA}\rangle^2. \tag{10.26}$$

To arrive at our aim, we simply drop the positive definite number s^2 and are left with

$$(\Delta\mathrm{A})^2(\Delta\mathrm{B})^2 \geq \frac{1}{4}\langle\mathrm{AB}-\mathrm{BA}\rangle^2. \tag{10.27}$$

Taking the square root on both sides completes the proof of (10.19). As an outlook for further reading we point out that the positive definite quantity $s^2 = \frac{1}{4}\langle\delta\mathrm{A}\delta\mathrm{B}+\delta\mathrm{B}\delta\mathrm{A}\rangle^2$ that we had simply dropped on the right of (10.26) has a certain importance that was already known to Schrödinger (1930), but has been largely ignored since.

We give two examples for relation (10.19). Take the two observables position x and momentum component p_x. In the Schrödinger picture, in *position representation* momentum is described by the operator $\mathrm{p}_x = (\hbar/\mathrm{i})\partial/\partial x$ acting on ψ, and position operation x is simply multiplication of ψ with x. For normalized spatial wave functions $\psi(x)$, with $\psi^\dagger\psi = 1$, the expectation value $\langle[\mathrm{p}_x, \mathrm{x}]\rangle = \psi^\dagger(\mathrm{p}_x\mathrm{x} - \mathrm{x}\mathrm{p}_x)\psi$ becomes

$$\psi^\dagger\left(\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x}(x\psi) - x\frac{\hbar}{\mathrm{i}}\frac{\partial\psi}{\partial x}\right) = \frac{\hbar}{\mathrm{i}}\psi^\dagger\left(\psi + x\frac{\partial\psi}{\partial x} - x\frac{\partial\psi}{\partial x}\right) = \frac{\hbar}{\mathrm{i}}\psi^\dagger\psi, \tag{10.28}$$

or

$$\langle \mathrm{p}_x\mathrm{x} - \mathrm{x}\mathrm{p}_x\rangle = \frac{\hbar}{\mathrm{i}}. \tag{10.29}$$

Inserted into (10.19), this gives $\Delta\mathrm{x}\Delta\mathrm{p}_x \geq \frac{1}{2}\hbar$.

The second example is the uncertainty of two spin components, for instance, of $\Delta\sigma_x$ and $\Delta\sigma_y$ for $s = \frac{1}{2}$. We insert the commutator of the Pauli matrices $[\sigma_x, \sigma_y] = 2\mathrm{i}\sigma_z$ into (10.19), and we want to prove that indeed

$$\Delta\sigma_x\Delta\sigma_y \geq \big|\langle\sigma_z\rangle\big|. \tag{10.30}$$

On the right, for pure states α or β, we have $\langle\sigma_z\rangle = \pm 1$ and $|\langle\sigma_z\rangle| = 1$. On the left, if both uncertainties are taken as equal, $\Delta\sigma_x \approx \Delta\sigma_y \approx \Delta\sigma_\perp/\sqrt{2}$, with $\Delta\sigma_\perp = \Delta\sigma = \sqrt{2}$ from Sect. 3.4.1, we have the same result $\Delta\sigma_x\Delta\sigma_y \approx \frac{1}{2}(\Delta\sigma_\perp)^2 = 1$ as on the right, and the equality holds in (10.30). In general, $\Delta\sigma_x$ can experimentally be diminished at the expense of $\Delta\sigma_y$, and vice versa, to obtain what are called *squeezed states*.

10.3 The Bloch Equations

We had applied the Heisenberg equation to the expectation value of the momentum operator $\boldsymbol{p}$ and had recovered Newton's classical equation of motion. In the following, we apply Heisenberg's equation to the expectation value of the angular momentum operator $\mathbf{J}$ and will recover the classical precession equation of the spinning top, valid for the magnetic dipole interaction for any value of quantum number j. Spin precession equations for general multipole interactions will be derived in Sect. 19.4.

The Heisenberg equation for the magnetic interaction with angular momentum $\mathbf{J}$ reads

$$\dot{\mathbf{J}} = \frac{\mathrm{i}}{\hbar}[\mathrm{H}_\mathrm{M}, \mathbf{J}], \tag{10.31}$$

with the Hamiltonian $\mathrm{H}_\mathrm{M} = \gamma\mathbf{J}\cdot\boldsymbol{B} = \gamma(\mathrm{J}_x B_x + \mathrm{J}_y B_y + \mathrm{J}_z B_z)$ and gyromagnetic ratio γ. The first component of $\dot{\mathbf{J}}$ then is

$$\dot{\mathrm{J}}_x = \frac{\mathrm{i}}{\hbar}[\mathrm{H}_\mathrm{M}, \mathrm{J}_x] = \frac{1}{2}\frac{\mathrm{i}}{\hbar}\gamma\big([\mathrm{J}_x, \mathrm{J}_x]B_x + [\mathrm{J}_y, \mathrm{J}_x]B_y + [\mathrm{J}_z, \mathrm{J}_x]B_z\big). \tag{10.32}$$

With $[\mathrm{J}_x, \mathrm{J}_x] = 0$, and with $[\mathrm{J}_y, \mathrm{J}_x]$ and $[\mathrm{J}_z, \mathrm{J}_x]$ from (10.11), we arrive at the equation of motion for the components and for the vector of the angular momentum operator

$$\dot{\mathrm{J}}_x = \gamma(\mathrm{J}_y B_z - \mathrm{J}_z B_y) = \gamma(\mathbf{J} \times \boldsymbol{B})_x, \quad \text{etc.} \tag{10.33a}$$

Hence

$$\dot{\mathbf{J}} = \gamma \mathbf{J} \times \boldsymbol{B}, \tag{10.33b}$$

$$\dot{\mathbf{J}} = \mathbf{J} \times \boldsymbol{\omega}_\mathrm{L}, \tag{10.33c}$$

with the Larmor angular frequency vector $\boldsymbol{\omega}_\mathrm{L} = \gamma \boldsymbol{B}$.

In the end, we are interested only in observable quantities, that is, in expectation values like $\boldsymbol{P} = \langle \mathbf{J} \rangle / \hbar j$, for which the spin precession equation can be alternatively written as

$$\langle \dot{\mathbf{J}} \rangle = \langle \mathbf{J} \rangle \times \boldsymbol{\omega}_\mathrm{L}, \tag{10.34a}$$

$$\dot{\boldsymbol{P}} = \boldsymbol{P} \times \boldsymbol{\omega}_\mathrm{L}, \tag{10.34b}$$

$$\dot{\boldsymbol{M}} = \gamma \boldsymbol{M} \times \boldsymbol{B}, \tag{10.34c}$$

see also Sect. 19.1.2. The last equation is given for the macroscopic magnetization $\boldsymbol{M} = n\mu \boldsymbol{P}$, in general with a time dependent magnetic field $\boldsymbol{B}(t)$.

The equation of motion for the operator $\mathbf{J}$ is formally identical to the *classical precession equation* of a spinning top.

Indeed, with classical angular momentum $\boldsymbol{J}$, subject to an external torque, and with the precession frequency $\boldsymbol{\omega}_\mathrm{P}$ of amplitude $\omega_\mathrm{P} = T/J$, the spinning top obeys

$$\boldsymbol{T} = \dot{\boldsymbol{J}} = \boldsymbol{J} \times \boldsymbol{\omega}_\mathrm{P}. \tag{10.35}$$

Equation (10.34b) for the quantum expectation value $\boldsymbol{P} = \langle \mathbf{J} \rangle / \hbar j$ therefore has the same solution as the classical equation for the spinning top. In the classical case, both $\boldsymbol{J}^2$ and J_z are constants of the motion, like in the quantum case, where the operators $\boldsymbol{J}^2$ and J_z commute with H_M and therefore are constant. The classical transverse component $\boldsymbol{J}_\perp = (J_x, J_y)$ of $\boldsymbol{J}$ performs a rotary motion with precession frequency $\boldsymbol{\omega}_\mathrm{P}$, which corresponds to the Larmor frequency $\boldsymbol{\omega}_\mathrm{L}$ in the quantum case, see Fig. 3.3.

In spite of their close similarity, there is a profound difference in the interpretation of the classical and the quantum mechanical precession equations. The classical

equation describes the motion of a quantity $\boldsymbol{J}$ whose future behavior can be calculated from $\dot{\boldsymbol{J}} = \boldsymbol{J} \times \boldsymbol{\omega}_{\mathrm{P}}$, where the initial conditions can be given with arbitrary precision. The quantum mechanical precession equation, when written not for the operator $\mathbf{J}$ but for the expectation value $\langle \mathbf{J} \rangle$, describes, as its name says, the expected mean result of measurements on the system with variance $(\Delta \mathbf{J})^2 = \langle \mathbf{J}^2 \rangle - \langle \mathbf{J} \rangle^2$. If the measurement is done on an ensemble of many particles, this variance can be very small, see (3.61).

Let a paramagnetic sample be magnetized by an external magnetic field along z. Its magnetization M_{z0} then is given by Curies' law, (6.27). If the magnetization $M_z(t = 0)$ of the system deviates from this equilibrium value, then it will, in the course of time, relax back to this value. This *relaxation* process usually is difficult to calculate, more on this in Chap. 22. The magnetization approaches its equilibrium value asymptotically, in simple cases exponentially, in which case the z-component of magnetization evolves as

$$M_z(t) = M_{z0} + \big[M_z(0) - M_{z0}\big] \mathrm{e}^{-t/T_1}, \tag{10.36}$$

where T_1 is the *longitudinal* or *spin–lattice relaxation time*.

If the nonequilibrium magnetization $\boldsymbol{M}(0)$ has components transverse to the external field, then these often decay to zero as

$$M_\perp(t) = M_\perp(0) \mathrm{e}^{-t/T_2}, \tag{10.37}$$

with a *transverse* or *spin–spin relaxation time* T_2. Transverse relaxation generally is faster than longitudinal relaxation, $T_2 \leq T_1$, because single spin flips ($\uparrow$ to $\downarrow$) induced by thermally fluctuating fields reduce both T_1 and T_2, while *spin diffusion* ("flip-flop") processes between neighboring spins ($\uparrow\downarrow$ to $\downarrow\uparrow$) reduce T_2 but not T_1 because they preserve M_z.

The *Bloch equations* Bloch (1946) have extra terms introduced into the spin precession equation (10.34c) to take relaxation processes phenomenologically into account:

$$\dot{M}_x = \gamma(\boldsymbol{M} \times \boldsymbol{B})_x - M_x/T_2, \tag{10.38a}$$

$$\dot{M}_y = \gamma(\boldsymbol{M} \times \boldsymbol{B})_y - M_y/T_2, \tag{10.38b}$$

$$\dot{M}_z = \gamma(\boldsymbol{M} \times \boldsymbol{B})_z - (M_{z0} - M_z)/T_1. \tag{10.38c}$$

In the absence of a magnetic field, the cross product terms $(\boldsymbol{M} \times \boldsymbol{B})_{x,y,z}$ vanish and only the differential equations for the exponential relaxation that lead to (10.36) and (10.37) remain. If the magnetic field is turned on, $\boldsymbol{B} \neq 0$, these cross product terms lead to an overall rotation of the spin ensemble. The Bloch equations then describe the usual spin precession of $\boldsymbol{M}$ about $\boldsymbol{B}$, but with longitudinal magnetization M_z

exponentially approaching its equilibrium value M_{z0} with time constant T_1, and with a transverse magnetization $\boldsymbol{M}_\perp$ fading away exponentially with time constant T_2. We shall make use of the Bloch equations in the following chapters on spin resonance and on atomic resonance transitions.

References

Bloch, F.: Nuclear induction. Phys. Rev. **70**, 460–474 (1946)

Schrödinger, E.: Zum Heisenbergschen Unschärfeprinzip. Sitzungsber. Preuss. Akad. Wiss. (Phys.-Math. Kl.) **19**, 296–303 (1930), for an English translation see arXiv:quant-ph/9903100v3

Part III
Quantum Physics at Work

Quantum mechanics has many applications in science and industry. We give a number of examples that deepen our understanding of the working of quantum physics. Our main topics are spin resonance, lasers and masers, glasses, superconductors, the systematics of nuclei and particles, and quantum computing.

This page intentionally left blank

Chapter 11
Spin Resonance

Abstract Resonance transitions between energy eigenstates of a quantum system, induced by time dependent perturbations, are a wide-spread phenomenon. In spin resonance, basic methods include Rabi's molecular beam method, continuous wave spectroscopy, free induction decay, spin echo, fast passage, and motional narrowing, phenomena of relevance also in other contexts of quantum physics. We discuss applications of nuclear magnetic resonance in chemical analysis, structural biology, medical imaging, and atomic clocks.

Resonance phenomena are ubiquitous in classical and quantum physics and in many other fields of science. In this chapter, we focus on *spin resonance*, and, in particular, on nuclear magnetic resonance.

The first spin resonance experiments served to measure the unknown magnetic dipole or electric quadrupole moments of the probe particles. Today most of these moments are known to high precision, and spin resonance is mainly used to investigate internal fields in condensed matter, with important applications in structural chemistry and in medical imaging. Many of the subtle effects discovered in spin resonance were later used in other effective spin systems, as shown in the following sections. Later on, Chap. 21 gives a more rigorous treatment of the resonance process.

Spin resonance experiments use various different probe particles. *Electron spin resonance* (ESR) uses paramagnetic electrons on free atoms or in condensed matter, with a gyromagnetic ratio

$$\gamma_e/2\pi = 28.0 \times \left|g_j\right| j \text{ GHz/T}, \tag{11.1}$$

with angular momentum j, and g-factor g_j. Microwave frequencies are required for external fields of the order of 1 T. Free electrons have $|g_j|j \approx 1$.

In *muon spin resonance* (μSR) experiments, muons are stopped in condensed matter and used to measure internal magnetic fields. The muon's gyromagnetic ratio is $\gamma_\mu/2\pi = 135$ MHz/T.

The most frequent application of spin resonance, however, is in *nuclear magnetic resonance* (NMR), with

$$\gamma_I/2\pi = 15.2 \times |g_I| I \text{ MHz/T}, \tag{11.2}$$

D. Dubbers, H.-J. Stöckmann, *Quantum Physics: The Bottom-Up Approach*, Graduate Texts in Physics,
DOI 10.1007/978-3-642-31060-7_11, © Springer-Verlag Berlin Heidelberg 2013

for nuclear spin I and nuclear g-factor g_I. Protons with $g_I I = 2.8$ have a gyromagnetic ratio $\gamma_p/2\pi = 42.6$ MHz/T.

11.1 Basics of Spin Resonance

Let a resonance transition be driven between two energy levels E_1 and E_2 of a quantum system with effective 2×2 Hamiltonian H. The energy difference between the two levels is $E_2 - E_1 = H_{22} - H_{11}$. The transition between the two levels is induced by a time dependent field, with a frequency component ω in its spectrum that meets the resonance condition $\hbar\omega = E_2 - E_1$. The interaction with this field is given by the transition matrix elements $H_{12}(t) = H_{21}^*(t)$.

We follow conventions in spin resonance and write B_0 for the static field and B_1 for the time dependent field amplitude. While the static field is applied along the z-axis, we let this high-frequency field $B_x(t) = 2B_1 \cos\omega t$ oscillate linearly along the x-axis, where the factor 2 is introduced for later convenience. The Hamiltonian then is

$$\mathrm{H} = \frac{1}{2}\hbar\omega_0\sigma_z + \hbar\omega_1\sigma_x \cos\omega t, \tag{11.3a}$$

$$\mathrm{H} = \frac{1}{2}\hbar \begin{pmatrix} \omega_0 & 2\omega_1 \cos\omega t \\ 2\omega_1 \cos\omega t & -\omega_0 \end{pmatrix}, \tag{11.3b}$$

where $\omega_1 = \gamma B_1$ is a measure of the amplitude of the oscillating field.

In spin resonance experiments, the static field B_0 is of the order of 1 T and usually is produced with a superconducting coil, while the field $B_1 \ll B_0$ is of the order of mT and, in NMR, is generated in a small radio frequency (RF) coil, with both coils installed at right angles to each other. In thermal equilibrium, the paramagnetic probe is magnetized to $P_z \approx \mu B_0/k_\mathrm{B}T$, see (6.27). Even at fields as strong as $B_0 = 10$ T, this polarization usually is small at room temperature, $P_z = 2.2$ % for electrons and $P_z = 3.4 \times 10^{-5}$ for protons, see (6.28) and (6.29). Since the sample has up to $n \sim 10^{22}$ spins per cm^3, the magnetization $M_z = n\mu P_z$ is large enough to produce a resonance signal.

The solution of the Schrödinger equation with a linearly oscillating high-frequency field $\boldsymbol{B}_1(t) = 2B_1(\cos\omega t, 0, 0)$ is difficult and is postponed to Sect. 21.3. Instead, we take a circularly rotating field $\boldsymbol{B}_+(t) = B_1(\cos\omega t, \sin\omega t, 0)$, assumed to rotate in the same direction as the precessing polarization $\boldsymbol{P}(t) = (P_x(0)\cos\omega t, P_y(0)\sin\omega t, P_z(0))$. We then go into a new coordinate system x', y', and $z' = z$ that rotates about the z-axis with the same angular frequency ω as the $\boldsymbol{B}_+$ field.

In this rotating frame, the field vector $\boldsymbol{B}_+ = \boldsymbol{B}_1$ is always at rest and points towards the x'-direction, while the polarization $\boldsymbol{P}(t)$ precesses about z with a reduced angular frequency $\omega_0 - \omega$. At resonance $\omega = \omega_0$, this difference frequency disappears, and the polarization is at rest, too. With no precession about z, there is no longer a field along z, and in the rotating frame the longitudinal field is simply transformed away, $B_0' = 0$. Off resonance $\omega \neq \omega_0$, the polarization precesses as if

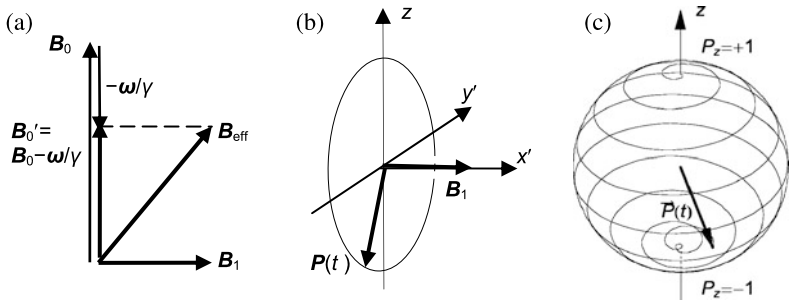


Fig. 11.1 To the rotating frame and back again: (**a**) In the rotating frame, the RF field of amplitude B_1 stands still, and the magnetic field component along z reduces to $B_0' = B_0 - \omega/\gamma$. (**b**) At resonance $\omega = \gamma B_0 = \omega_0$, the static field B_0 is transformed away, the effective field vector becomes $\boldsymbol{B}_{\rm eff} = \boldsymbol{B}_1$, and the polarization vector $\boldsymbol{P}(t)$, originally along z, slowly precesses about the RF field vector $\boldsymbol{B}_1$. (**c**) Back in the laboratory frame, $\boldsymbol{P}(t)$ performs a spiraling motion, up and down on the surface of the unit sphere. In (**b**) and (**c**), $\boldsymbol{P}$ is shown is at polar angle $\theta = \omega_1 t = 200°$

the field along z were reduced to $B_0' = (\omega_0 - \omega)/\gamma = B_0 - \omega/\gamma$, see Fig. 11.1(a), hence in the rotating frame

$$\mathrm{H_R} = \frac{1}{2}\hbar(\omega_0 - \omega)\sigma_z + \frac{1}{2}\hbar\omega_1\sigma_x, \tag{11.4a}$$

$$\mathrm{H_R} = \frac{1}{2}\hbar\begin{pmatrix} \omega_0 - \omega & \omega_1 \\ \omega_1 & -(\omega_0 - \omega) \end{pmatrix}. \tag{11.4b}$$

The magnetic field seen in the rotating frame $\boldsymbol{B}_{\rm eff} = (B_1, 0, B_0 - \omega/\gamma)$ and with it the Hamiltonian no longer depend on time.

Hence the problem is reduced to the time independent case treated in Chap. 5.

Thus at resonance $\omega = \omega_0$, the total effective field $\boldsymbol{B}_{\rm eff}(\omega_0)$ is reduced to the radio-frequency field $\boldsymbol{B}_1$, stationary in the rotating frame. The polarization then precesses about $\boldsymbol{B}_1$ at angular frequency $\omega_1 = \gamma B_1$, see Fig. 11.1(b). Back to the laboratory frame, the polarization $\boldsymbol{P}(t)$, in addition to its slow precession about $\boldsymbol{B}_1$, performs a fast precession with angular frequency $\omega_0 = \gamma B_0$ about z. The tip of the polarization vector then follows a spiral-like trajectory on the surface of the unit sphere, periodically moving from the "north pole" $(+z)$ of the sphere to the "south pole" $(-z)$, and back, see Fig. 11.1(c).

How good is this *rotating wave approximation*? It is perfect if we really apply a rotating RF-field, which can be generated by two orthogonal linear RF coils, one along x, the other along y, with a phase difference of $\pi/2$ in their alternating currents. But even if we do apply a linearly oscillating field $\boldsymbol{B}_1(t) = 2B_1(\cos\omega t, 0, 0)$, this approximation is excellent in most cases. The reason is that $\boldsymbol{B}_1(t)$ can be thought of as being the vector sum $\boldsymbol{B}_1 = \boldsymbol{B}_+ + \boldsymbol{B}_-$ of two counter-rotating circular fields $\boldsymbol{B}_\pm(t) = B_1(\cos\omega t, \pm\sin\omega t, 0)$. At resonance in the rotating frame, the

polarization $\boldsymbol{P}$ sees the rotating field $\boldsymbol{B}_+ = \boldsymbol{B}_1$ at rest, and sees the counter rotating field $\boldsymbol{B}_-$ far off resonance at doubled angular frequency $-2\omega_0$. The adiabatic theorem of quantum mechanics, Sect. 7.2, tells us that, in the nonadiabatic approximation, a rapidly changing external perturbation will not affect a quantum system. If one goes beyond this approximation, one finds:

> The counter-rotating field component leads to a *Bloch–Siegert shift* of the resonance signal, typically of the order of $\Delta\omega_0/\omega_0 \approx (\omega_1/\omega_0)^2 \sim 10^{-6}$.

11.2 Methods of Spin Resonance

There exists an innumerable variety of spin resonance schemes, with probes either in-beam or in condensed matter. We discuss the following methods: (i) The Rabi method; (ii) Continuous wave spectroscopy; (iii) Free induction decay; (iv) Spin echo; (v) Adiabatic fast passage.

(i) The earliest magnetic resonance experiment was done with a molecular beam in which nuclear spin flips are induced while the molecules fly through a resonance region. The *Rabi method* (1939) proceeds in three steps (see Fig. 11.2(a)).

Step 1: In the strongly nonuniform field of a magnet labeled A, one nuclear spin component in a diamagnetic molecule is selected by the Stern–Gerlach effect.
Step 2: The polarized beam enters the resonance region, with the uniform field of a second magnet C and with a radio-frequency field superimposed.
Step 3: A third magnet labeled B, of the same type as magnet A, serves as a spin analyzer and focuses the beam onto a detector of molecules. The fields A, B and C have the same sign, while the gradients of fields A and B differ in sign.

The *Rabi formula* gives the shape of the in-flight resonance signal $P_z = 1 - 2p_{\text{down}}$, see (5.24) and (3.34b), which is written in terms of the field components as

$$P_z(B_0) = 1 - 2\frac{B_1^2}{B_1^2 + (B_0 - \omega/\gamma)^2} \sin^2\left(\gamma\sqrt{B_1^2 + (B_0 - \omega/\gamma)^2}\,T\right). \tag{11.5}$$

Figure 11.2(b) shows a measured resonance signal. To find a previously unknown resonance after many months of work is an unforgettable moment in the life of a scientist.

(ii) In *continuous wave spectroscopy*, one measures the time average response of a polarized atomic, nuclear, or particle spin system as a function of the frequency ω of a continuous radiation field $B_1 \cos\omega t$. We first neglect relaxation, and do not specify the method of signal detection. Figure 11.3(a) shows the energy eigenvalues of the spin resonance Hamiltonian (11.4b) in the rotating frame, which are the same as those in Fig. 7.1(a), but with B_x replaced by B_1 and B_z/B_{x0} replaced by

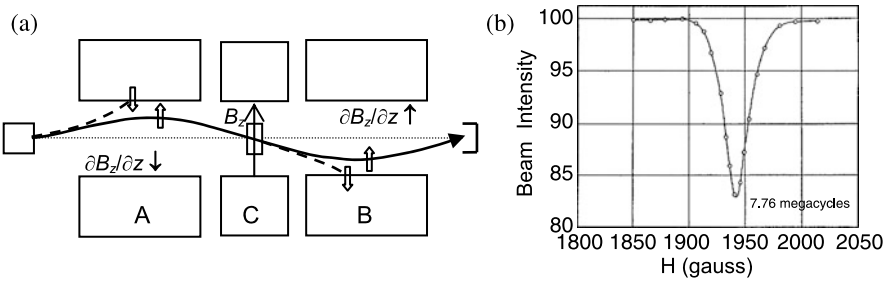


Fig. 11.2 (**a**) Rabi apparatus with A, B, and C magnets, and molecular trajectories without (*full line*) and with π-flip (*dashed line*) in the resonance region (*central rectangle*). In modern versions of the apparatus, magnets A and B are quadrupole or hexapole magnetic lenses that focus one spin component while defocussing the other. The detector is a hot tungsten filament that ionizes the incoming molecules and translates beam intensity into an electric current. (**b**) Rabi resonance curve of ^{19}F nuclei in a beam of NaF molecules, from Rabi et al. (1939). The *line* is drawn to guide the eye. The wiggles expected from (11.5) in the outer wings of the signal are washed out for a broad molecular velocity spectrum

$(\omega - \omega_0)/\omega_1$. In the spin precession picture, the time average polarization is obtained from Fig. 11.1(a)–(b) by first projecting the initial $\boldsymbol{P} = (P_{z0}, 0, 0)$ onto the field $\boldsymbol{B}_{\text{eff}}$ to obtain the time average $\bar{\boldsymbol{P}}(\omega)$ of the precessing polarization, and then projecting this a second time onto the z-component to obtain $\bar{P}_z(\omega) = \cos^2\theta = B^2_{\text{eff}z}/B^2_{\text{eff}}$, which stays the same when going back to the laboratory frame. The same result can alternatively be derived from (5.27), with the cosine term averaged away, and with the same replacements as before, and one obtains

$$\bar{P}_z(\omega) = P_{z0}\frac{(\omega - \omega_0)^2}{\omega_1^2 + (\omega - \omega_0)^2}, \tag{11.6}$$

for an initial nuclear polarization P_{z0}, as shown in Fig. 11.3(b). The wings of the Lorentz-shaped resonance signal fall off quadratically with increasing detuning $|\omega - \omega_0|$. The signal reaches half its minimum at $(\omega - \omega_0)^2 = \omega_1^2$, hence the signal's full width $\Delta\omega = 2\omega_1$ grows linearly with the RF amplitude B_1. This *power broadening* conforms with the uncertainty principle, as $\Delta E = \hbar\Delta\omega = 2\hbar\omega_1 > \hbar/T$, for a mean time $T = 1/\omega_1$ of residence in a given eigenstate of energy. Resonance detection methods that are sensitive to the longitudinal $\bar{P}_z(\omega)$ will be discussed in Chap. 21.

Continuous wave NMR, used with solid or liquid probes of condensed matter, usually is not sensitive to $\bar{P}_z(\omega)$, but uses a *nuclear induction* method that is sensitive to the transverse magnetization $M_x(\omega) = n\mu\bar{P}_x(\omega)$ or $M_y(\omega) = n\mu\bar{P}_y(\omega)$. The equilibrium NMR line shapes are calculated by setting the time derivatives on the left-hand side of the Bloch equations (10.38a)–(10.38c) to zero, again in the rotating frame, and by solving the resulting system of linear equations by inversion of the corresponding 3×3 matrix. The derivation is given in the original paper (Bloch 1946). The response of the probe's magnetization is $M_x(\omega) = \chi'(\omega)M_{z0}$ with the susceptibility $\chi'(\omega)$, in phase with the field component $B_x(t)$,

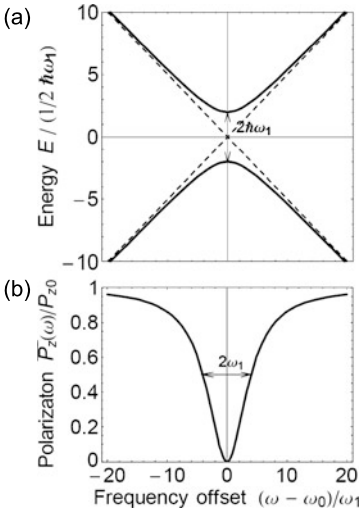


Fig. 11.3 (**a**) Energy diagram of a two-level system under high-frequency irradiation as seen in the rotating frame. (**b**) At resonance $\omega = \omega_0$, the RF field completely mixes the two spin states "up" and "down", similar as in Fig. 7.1(b), and therefore makes the longitudinal polarization vanish. The resonance signal shown is from (11.6)

and $M_y(\omega) = \chi''(\omega) M_{z0}$ with the out-of-phase susceptibility $\chi''(\omega)$,

$$\chi'(\omega) = \frac{(\omega - \omega_0)\omega_1 T_2^2}{1 + (\omega - \omega_0)^2 T_2^2 + \omega_1^2 T_1 T_2}, \tag{11.7a}$$

$$\chi''(\omega) = \frac{\omega_1 T_2}{1 + (\omega - \omega_0)^2 T_2^2 + \omega_1^2 T_1 T_2}, \tag{11.7b}$$

see Fig. 11.4(b) where the RF field amplitude is given by $\omega_1 = B_1/\gamma$. Be aware that the slashes on χ do not mean a derivative. The in-phase signal $\chi''(\omega)$ has Lorentz-shape, and has full width at half maximum $\Delta\omega = 2(1 + \omega_1^2 T_1 T_2)^{\frac{1}{2}}/T_2$. The $\chi'(\omega)$ signal is zero at exact resonance and has peak values at the frequencies $\omega_0 \pm \frac{1}{2}\Delta\omega$ where the absorption signal is at half maximum.

In continuous wave NMR spectroscopy, one measures the time average voltage induced by the precessing magnetization $\boldsymbol{M}(t) = n\mu \boldsymbol{P}(t)$ in a pick-up coil. For in-phase measurements, this coil is wound on top of the B_1-field-generating RF coil. For out-of-phase measurements, the pick-up coil is wound with its axis perpendicular to the RF coil, see Fig. 11.4(a). For practical reasons, when searching the resonance signal one does not vary the angular frequency ω of the tuned L–C resonance circuit, but instead slowly sweeps the magnetic field B_0. In the limit of *slow passage*, this sweep is much slower than all other rates ω, ω_0, ω_1, T_1^{-1}, and T_2^{-1}.

(iii) *Free induction decay* is another frequently used NMR detection method. First, a $\pi/2$-*flip* rotates the initial magnetization M_{z0} into the x–y plane. This is accomplished with a resonant RF pulse of amplitude B_1 and duration T such that the rotation angle about $\boldsymbol{B}_1$ is $\theta = \gamma B_1 T = \pi/2$. After the end of the pulse, the transverse magnetization precesses freely about B_0 and produces a time dependent magnetic field. The RF coil then acts as a pickup coil in which a time dependent voltage signal is induced, which can be Fourier analyzed to give an NMR signal in

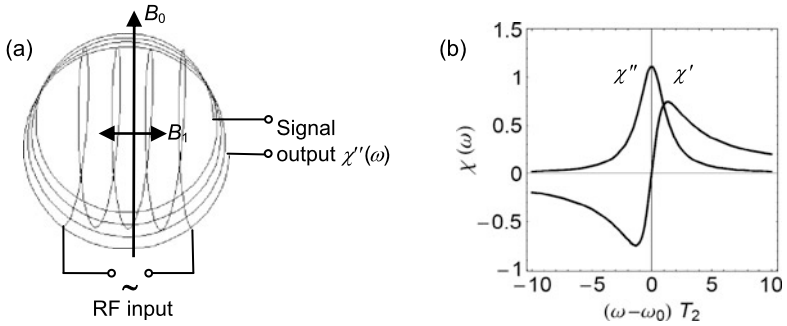


Fig. 11.4 (**a**) Configuration of the NMR coil for RF input and the pick-up coil for $\chi''(\omega)$ signal output. The RF coils are part of an LC resonance circuit. (**b**) NMR $\chi'(\omega)$ in-phase and $\chi''(\omega)$ out-of-phase signals, from (11.7a), (11.7b) with $T_2 = 0.2T_1$ and $\omega_1 = 5T_2^{-1}$

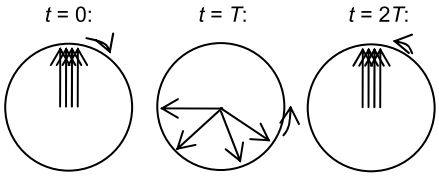


Fig. 11.5 In a spin echo experiment, after a $\pi/2$ flip at time zero, the spins precess in the x–y plane (the paper plane) at different frequencies and become increasingly dispersed. If at time T the direction of precession is reversed, then at time $2T$ all spins are lined up again

the frequency domain. Today, with very low-noise equipment, even the precession of a single proton confined in a trap can be detected (Ulmer et al. 2011).

(iv) The *spin echo* technique gives important information on the spin dynamics inside the NMR probe. In free induction decay, the transverse magnetization relaxes exponentially as $\boldsymbol{M}_\perp \propto \exp(-t/T_2)$. In addition, there is a further decay of $\boldsymbol{M}_\perp$ due to instrumental imperfections, like inhomogeneities of the magnetic field, which make spins at different positions precess at different frequencies $\omega_0(\boldsymbol{x}) = \gamma B_0(\boldsymbol{x})$. These unwanted depolarization channels usually mask the physically more interesting but slower true T_2 relaxation processes. This instrumental $\boldsymbol{M}_\perp$ decay is usually taken into account by adding a $1/T_2^*$ rate to $1/T_2$ in the exponent, although in general this additional decay is not exponential. However, this detail in practise plays no role because, as we shall see, these T_2^* effects can be eliminated experimentally.

In a spin echo experiment, after the $\pi/2$ flip, the individual spins in the probe start their precession in the x–y plane all with the same initial phase. In the course of time, due to dephasing with time constant T_2^*, their relative phases are dispersed over larger and larger angles, see Fig. 11.5, and transverse magnetization is rapidly lost. But while the true T_2 processes are irreversible, the unwanted T_2^* processes are reversible in time. If after a time T the direction of spin rotation is reversed by some means, after twice this time $2T$ the reversible magnetization loss with time constant

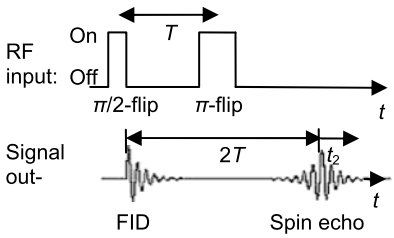


Fig. 11.6 Spin echo sequence of $\pi/2$ and π RF pulses and the probe's response, first by a rapidly decaying free induction signal (FID), and after time $2T$ by a spin echo signal. For t_2 see end of Sect. 11.3(i)

T_2^* is "undone", and the free induction signal reappears as a spin echo, diminished only by the true and irreversible T_2 relaxation processes.

This reminds us of a field of long-distance runners on a circular race-track. After their start, initiated by the usual gun shot, they run at different individual speeds such that the field will become increasingly dispersed. If at a time T, many rounds after the start, and upon a second gun shot, all runners turn back on their track and run at the same individual speed in reverse direction, then after a total time $2T$ all runners will simultaneously cross the starting line again.

Spin relaxation times T_2 can be measured via the spin echo signal of magnitude

$$M_\perp(T) = M_\perp(0)\,\mathrm{e}^{-2T/T_2}. \tag{11.8}$$

From the dependence of T_2 on temperature, magnetic field, or other parameters, one can gain important insight into the internal dynamics of condensed matter, as will be described in Chap. 22. We shall meet various other spin echo applications, both with real spins and with effective spins, in the chapters to follow.

In a spin echo experiment, usually the sign of the magnetic field B_0 cannot be suddenly inverted, unless spin echo is done in-beam with atoms flying through a succession of suitably arranged fields. Therefore, another trick is needed. A time T after the first $\pi/2$-flip one applies a second π-flip to all the spins rotating in the x–y plane. This is accomplished by applying a radio-frequency pulse, oscillating along x like the initial $\pi/2$-pulse, but of twice its duration. This π-flip moves a spin vector with phase angle φ in the x–y plane into a new direction in the same plane, with a new phase angle $2\pi - \varphi$. After this π-flip, the most advanced spins will be the most retarded, and vice versa. After a total time $2T$, all spins are again in phase with each other, and the magnetization signal reappears as a spin echo. A spin echo pulse sequence and the ensuing spin echo signals are shown in Fig. 11.6.

(v) Another NMR method, *adiabatic fast passage*, allows us to invert a given spin polarization in an elegant and efficient way. To this end, one slowly sweeps the field B_0 through resonance. One starts at a field far above resonance, $B_0 \gg \omega/\gamma$, then moves the field through resonance at $B_0 \approx \omega/\gamma$, and ends far below resonance at $B_0 \ll \omega/\gamma$. The effective field in the rotating frame, Fig. 11.1(a), then changes from $B_{\mathrm{eff}} \approx B_0$ along $+z$, through $B_{\mathrm{eff}} \approx B_1$ along x, to $B_{\mathrm{eff}} \approx -B_0$ along $-z$. The polarization follows the effective field adiabatically. When $\boldsymbol{B}_{\mathrm{eff}}$ inverts its sign from $+\boldsymbol{B}_0$ to $-\boldsymbol{B}_0$, the polarization component P_z changes sign from $+P_{z0}$ to $-P_{z0}$.

To be adiabatic, the rate of change $(1/B_{\text{eff}})\partial B_{\text{eff}}/\partial t$ of the external field, i.e., of $\dot{\omega}_{\text{eff}}/\omega_{\text{eff}}$, must be slow compared to the system's precession frequency ω_{eff}, that is, $\dot{\omega}_{\text{eff}} \ll \omega_{\text{eff}}^2$. At the same time, the adiabatic passage must be *fast* enough such that the polarization does not fade away during passage, i.e., $\dot{\omega}_{\text{eff}}/\omega_{\text{eff}} \gg T_1^{-1}, T_2^{-1}$. With the method of adiabatic fast passage, simply turning the knob of the magnet's power supply makes all the $\sim 10^{22}$ spins in the sample flip from spin up to spin down, back and forth if you like, under favorable conditions with essentially no losses.

11.3 Applications of Spin Resonance

We discuss three widespread applications of spin resonance in (i) NMR spectroscopy, (ii) medical imaging, and (iii) atomic clocks.

(i) NMR spectroscopy is big business. Almost every chemistry laboratory has a high resolution NMR spectrometer in its basement. We shall discuss two main applications of NMR spectroscopy: first, chemical analysis where the NMR spectrum serves as a "fingerprint" that signals the presence of a molecule, and second, molecular structure analysis which gives us the spatial position of every atom even for very large molecules.

Let us first discuss the shapes and frequencies of NMR lines. In the solid state, NMR lines are *inhomogeneously broadened* because every probe nucleus sees a slightly different internal magnetic field $B_{\text{int}} \sim (\mu_{\text{N}}/4\pi)r^{-3}$ from the almost randomly oriented magnetic moments of neighboring nuclei. These small internal fields of amplitude $B_{\text{int}} \sim 10^{-4}$ T for internuclear distances $r \sim 1$ nm are superimposed on the strong external holding field B_0 and are approximately Gaussian distributed. In modern NMR spectrometers, with $B_0 \approx 10$ T and $\nu = \omega/2\pi \approx 400$ MHz, inhomogeneous line widths have $\Delta\nu/\nu_0 = B_{\text{int}}/B_0 \sim 10^{-5}$, or $\Delta\nu \sim$ several kHz.

In liquid NMR probes, this inhomogeneous broadening is suppressed by the effect of *motional narrowing*: due to the rapid movements of the molecules in a liquid, amplitudes and directions of internal fields change so fast that their effect on the NMR line averages out; see also Sect. 22.3. The residual *homogeneous* line width $\Delta\nu = 1/(2\pi T_2) \sim$ several Hz is due to transverse relaxation and is a natural width due to energy–time uncertainty, see Sect. 3.4.2. These extremely narrow NMR lines are sensitive to the smallest changes in the environment of the nuclei in a molecule:

> *Chemical shifts* arise because every nucleus in a molecule sees a different site-specific electronic environment, with a slightly different NMR frequency.

In this way, each molecule in a liquid leaves its specific signature in the NMR spectrum. This makes NMR spectroscopy an indispensable tool of chemical analysis. For an example, see Fig. 11.7.

NMR spectroscopy is also highly valuable for molecular structure determination. Most effective to this end is two- or *multidimensional NMR spectroscopy*. With this

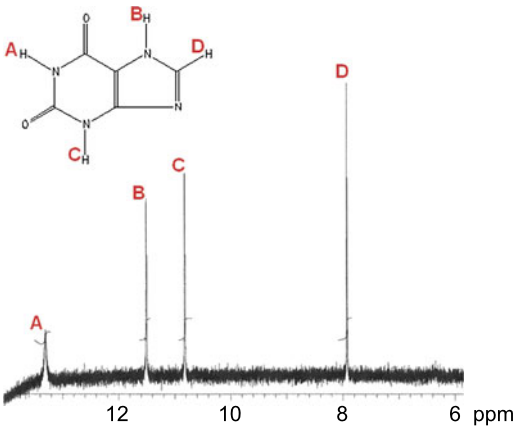


Fig. 11.7 Proton NMR "fingerprint" of the organic molecule xanthine. The NMR lines labeled A, B, C, and D are due to protons at the indicated sites of the molecule, shown in the upper left. The chemical shifts of the NMR lines are given in parts per million (ppm). From Wikimedia public domain

method, the intensities of spin echo signals are recorded not only as a function of the time interval $t_1 = T$ between the two spin echo RF pulses, but also as a function of the time t_2 passed after the spin echo signal, see Fig. 11.6. The entire set of spin echo responses $f(t_1, t_2)$ embodies information on the distances and angular positions of neighboring nuclear spins. From these data one can build up the entire three-dimensional atomic structure even of large molecules. NMR studies of protein structures can be done with the molecules in their natural liquid environment, whereas structural studies with classical X-ray diffraction require crystalline probes.

The structure of thousands of proteins, each containing up to 100 000 atoms, has been fully established with multidimensional NMR spectroscopy, which competes with and complements the methods of X-ray diffraction.

(ii) Another well known application of NMR is in medical diagnostics, known as *magnetic resonance tomography*. MRT records spatially resolved NMR signals $S(\boldsymbol{x})$ in dependence of the position $\boldsymbol{x} = (x, y, z)$ of the probe nuclei within the volume of a patient's body, which then are transformed into three-dimensional MRT images.

In magnetic resonance tomography, the magnitude of the position-sensitive signal $S(x, y, z)$ depends on the local density of the probe nuclei and on their local relaxation times, which vary with the medical condition of the tissue under study.

Three successive steps are required for MRT image formation, one for each dimension. *Step 1*, z-position by *slice selection*: The patient is oriented along the z axis of the holding field $B_0(z)$. The field has a gradient along its z axis, such that the resonance condition

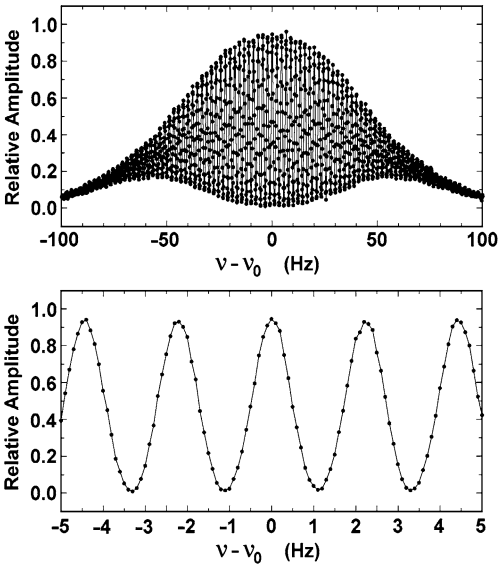


Fig. 11.8 Ramsey signal observed in the hyperfine splitting of the ^{133}Cs ($I = 7/2$) ground state, obtained with a "fountain" of cesium atoms at a temperature <2 μK, after double passage of the rising and falling atoms through the same microwave $\pi/2$-flip region. The atoms' mean time-of-flight of 0.4 s determines the ∼1 Hz width of the Ramsey fringes, while the Fourier transform of the atoms' time-of-flight spectrum determines the ∼100 Hz width of the signal envelope. The lower panel shows an expanded view of the central fringes. From Sullivan et al. (2001)

$B_0(z) = \omega/\gamma$ is fulfilled only for a thin slice of the patient near $z = z_0$. When a $\pi/2$ RF pulse is applied, only nuclei within this slice participate in the subsequent free induction decay.
Step 2, x-position by *phase encoding*: During a short time, small field gradients are applied along x, which produce a gentle push or pull on the precessing transverse magnetization $\boldsymbol{M}_\perp(t)$, which acquires a different phase $\varphi(x)$ for each position x.
Step 3, y-position by *frequency encoding*: During the whole measurement, a small field gradient is applied along y, such that the transverse magnetization $\boldsymbol{M}_\perp(t)$ precesses at a slightly different angular frequency $\omega(y)$ for each position y. The superposition of these free induction signals then shows strong interference beats in time, such that each precession frequency ω in its Fourier spectrum corresponds to a different position y.

(iii) Finally, *atomic clocks* are based on in-beam NMR and are in wide use as time and frequency standards. They nowadays reach a relative precision of 1.5×10^{-15}, or 1 second in about 10^7 years, with rapid progress every year. The GPS global positioning system relies on atomic clocks, as do many high precision tests of special and general relativity. Atomic clocks use *Ramsey's method of separate oscillatory fields*, which we discuss briefly. The same as in spin echo, spins initially polarized along $+z$ are irradiated with two successive radio-frequency pulses, separated by a time interval T. In contrast to spin echo, the oscillations of the two RF pulses are of equal length and stay in phase with each other. To guarantee phase coherence, both pulses come from the same oscillator, which is kept running between the two

RF pulses and is merely gated off at the end of each pulse. In the Ramsey method, the first $\pi/2$ pulse turns the nuclear polarization from $+z$ into the x–y plane, and thereafter the phase of the spins, precessing in the x–y plane, evolves as $\varphi_{\text{spin}} = \omega_0 t$, while the phase of the oscillator evolves as $\varphi_{\text{osc}} = \omega t$. If both frequencies coincide, $\omega = \omega_0$, both phases evolve in parallel, and the second $\pi/2$ pulse has the same effect as the first, i.e., it turns the spins by another 90° from the x–y plane into the $-z$ direction. If, however, ω is detuned from ω_0, then the second pulse flips the spins elsewhere, depending on the phase difference $\Delta\varphi = (\omega - \omega_0)T$. Altogether, one obtains an interference signal in dependence of the detuning $\omega - \omega_0$. This is in analogy to double-slit interference, which relies on the phase difference $\Delta\varphi = (x - x_0)k$, with $k = 2\pi/\lambda$ and with the optical-path difference $x - x_0$. Therefore,

> The Ramsey method is the time domain analogue of a double-slit experiment, and the signals shown in Figs. 1.1 and 11.8 in favorable cases have identical shapes.

The Ramsey method is the more precise the longer the time interval T between the two pulses. Therefore, the most advanced atomic clock uses a "fountain" of 2 μK cold cesium atoms with rise plus fall time $T = 0.4$ s between the RF pulses. The ^{133}Cs nuclei are polarized by optical pumping (see the end of Sect. 18.3), and their polarization is optically monitored. The ^{133}Cs splitting $\hbar\omega_0$ of the $F = 4$ and $F = 3$ hyperfine levels (cf. Sect. 8.3) has a frequency $2\pi\omega_0 = 9\,192\,631\,770$ Hz, which is used as a time standard that defines the second as the time elapsed after 9 192 631 770 oscillation periods of the ^{133}Cs spins.

References

Bloch, F.: Nuclear induction. Phys. Rev. **70**, 460–474 (1946)

Rabi, I.I., Millman, S., Kusch, P., Zacharias, J.R.: The molecular beam resonance method for measuring nuclear magnetic moments. Phys. Rev. **55**, 526–535 (1939)

Sullivan, D.B., Bergquist, J.C., Bollinger, J.J., Drullinger, R.E., Itano, W.M., Jefferts, S.R., Lee, W.D., Meekhof, D., Parker, T.E., Walls, F.L., Wineland, D.J.: Primary atomic frequency standards at NIST. J. Res. Natl. Inst. Stand. Technol. **106**, 47–63 (2001)

Ulmer, S., Rodegheri, C.C., Blaum, K., Kracke, H., Mooser, A., Quint, W., Walz, J.: Observation of spin flips with a single trapped proton. Phys. Rev. Lett. **106**, 253001(4) (2011)

Chapter 12
Two-State Systems in Atomic and Molecular Physics

Abstract Various effective-spin systems in atomic physics are investigated. The first example is the photon with its two states of circular polarization. The photon's coherent superposition states can be mapped onto the surface of the Bloch sphere. The next example presents two atomic states involved in a resonance transition, driven by laser or microwave radiation. This leads to Rabi oscillations, photon free induction decay, quantum beats, photon echoes, and Ramsey signals. In molecular physics, a well known example of a two-state system is a nitrogen atom in the double-well potential of an ammonia molecule. A modern version of this experiment is a Bose–Einstein condensate in a double-well optical trap.

12.1 Photons as Two-State Systems

The photon is a well known two-state quantum system. Classical electrodynamics knows right-circular σ^+ and left-circular σ^- electromagnetic waves. We call these two basis states R and L, for right and left. Linearly polarized light is a coherent superposition of σ^+ and σ^- light. We could as well take two orthogonal linearly polarized waves and call them H and V, for horizontal and vertical.

The photon as a particle has spin quantum number $s = 1$ with only two substates with magnetic quantum numbers $m = 1$ and $m = -1$. In relativistic quantum field theory, it is shown that massless particles of spin j possess only the two $m = \pm j$ substates of spin. The other m-states do not exist because this would violate Lorentz invariance. In a quantum description, linearly polarized photons are a coherent superposition of their $m = \pm 1$ states. Again, there is an infinite variety of superposition states $\psi = a\mathrm{R} + b\mathrm{L}$, with complex amplitudes a and b, and $|a|^2 + |b|^2 = 1$.

In Hilbert space, all possible superposition states $\psi = a\alpha + b\beta$ with $|\psi|^2 = 1$ can be mapped onto the surface of the *Bloch sphere*, in optics also called the *Poincaré sphere*, shown in Fig. 12.1. The Bloch sphere should not be confused with the sphere described by the vector $\boldsymbol{P}(t) = \langle \boldsymbol{\sigma} \rangle(t)$ of length $|\boldsymbol{P}| = 1$, shown in Fig. 3.3, which is defined in three-dimensional vector space.

D. Dubbers, H.-J. Stöckmann, *Quantum Physics: The Bottom-Up Approach*, Graduate Texts in Physics,
DOI 10.1007/978-3-642-31060-7_12, © Springer-Verlag Berlin Heidelberg 2013

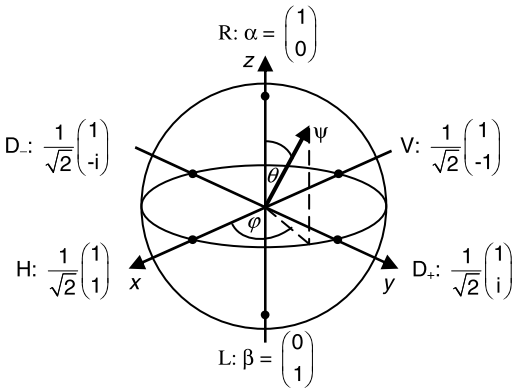


Fig. 12.1 All normalized quantum states $\psi = a\alpha + b\beta$ can be mapped onto the surface of the Bloch sphere in Hilbert space, with $\psi = \alpha$ on the north pole, $\psi = \beta$ on the south pole, and states with equal amplitudes of α and β on the equator. All other states are on the two hemispheres in between

Similarly as in Sect. 3.3.2, we write the general two-state vector ψ in terms of its polar and azimuthal angles θ and φ on the Bloch sphere as

$$\psi(\theta,\varphi) = \mathrm{R}e^{-\mathrm{i}\phi/2}\cos\frac{1}{2}\theta + \mathrm{L}e^{\mathrm{i}\varphi/2}\sin\frac{1}{2}\theta = \begin{pmatrix} e^{-\mathrm{i}\varphi/2}\cos\frac{1}{2}\theta \\ e^{\mathrm{i}\varphi/2}\sin\frac{1}{2}\theta \end{pmatrix}. \tag{12.1}$$

The orthogonal states R and L with $\theta = 0$ and 180° lie on the north and south poles of the sphere, respectively. All states with linear polarization have $\theta = 90°$ and lie on the equator, for instance, $\mathrm{H} = (\mathrm{R}+\mathrm{L})/\sqrt{2}$ with $\varphi = 0$ for horizontal polarization, $\mathrm{V} = (\mathrm{R}-\mathrm{L})/\sqrt{2}$ with $\varphi = 180°$ for vertical polarization, and $\mathrm{D}_\pm = (\mathrm{R} \pm \mathrm{iL})/\sqrt{2}$ with $\varphi = \pm 90°$ for diagonal polarization. The states lying between the equator and the poles correspond to *elliptical polarization*. In the laboratory, the state of a photon can be altered by the use of optical elements.

Loss-free optical elements can be described by 2×2 matrices that move a photon state ψ to some other state on the Bloch sphere.

In the H–V basis, frequently used in optics, the *Jones matrices* for horizontal and vertical linear polarizers are $\mathrm{J_H} = \frac{1}{2}(\mathrm{I} + \sigma_z)$, $\mathrm{J_V} = \frac{1}{2}(\mathrm{I} - \sigma_z)$, and for a phase shifter $\mathrm{J}_\varphi = e^{\mathrm{i}\varphi/2}\sigma_z$. Such devices permit the experimenter to walk around at will in "optical" Hilbert space.

12.2 Optical Resonance Transitions

Two energy levels of an atom, coupled to each other by a laser field, are another example of an $s = \frac{1}{2}$ effective-spin system. When atoms are irradiated with light, their response in most cases is well described in the *semiclassical approximation*,

with light regarded as a classical electromagnetic wave, characterized by its electric field vector $\boldsymbol{E}$. In laser spectroscopy, light typically is a traveling plane wave,

$$\boldsymbol{E}(z,t) = \boldsymbol{E}_0 \mathrm{e}^{\mathrm{i}(kz-\Omega t)}, \tag{12.2}$$

oscillating at angular frequency Ω and propagating along z with wave number $k = 2\pi/\lambda$. Classical field amplitudes are observables and must be real numbers, and the complex presentation is chosen only for mathematical convenience. This means that, without explicitly noting this in the equations, only the real part of all occurring quantities must be taken. In contrast, amplitudes in quantum mechanics are not observables and can be truly complex quantities.

In an optical transition between low-lying states of the valence electrons, the wavelength λ is typically of the order of several 100 nm, while the diameter of the light-emitting atom is of the order of nm or less. Therefore, within the volume of the atom, we set $\mathrm{e}^{\mathrm{i}kz} = \mathrm{e}^{2\mathrm{i}\pi z/\lambda} \approx 1$, and

$$\boldsymbol{E}(z,t) \approx \boldsymbol{E}(t) = \boldsymbol{E}_0 \mathrm{e}^{-\mathrm{i}\Omega t}. \tag{12.3}$$

The time dependent electric field $\boldsymbol{E}$ interacts with the atomic electric dipole moment $\boldsymbol{d}_{\mathrm{el}}$ formed by the negative electron shell of the atom and its positive nucleus, which are assumed to be elastically bound to each other. The classical potential of an electric dipole interaction is $V_{\mathrm{E}} = -\boldsymbol{d}_{\mathrm{el}} \cdot \boldsymbol{E}$.

In the quantum case, we proceed in analogy to the magnetic dipole interaction. There we had the classical magnetic dipole moment $\boldsymbol{\mu} = I\boldsymbol{a}$ of a current I on a circuit of area a, and replaced it by the magnetic moment operator $\boldsymbol{\mu} = \mu\boldsymbol{\sigma}$ to arrive at the Hamiltonian $\mathrm{H_M} = -\boldsymbol{\mu} \cdot \boldsymbol{B}$ of the magnetic dipole interaction. Similarly, we have the classical dipole moment $\boldsymbol{d}_{\mathrm{el}} = e\boldsymbol{r}$ for two charges of magnitude e and of opposite sign, with the displacement vector $\boldsymbol{r}$ pointing from the negative to the positive charge, and replace it by the operator $\mathbf{d}_{\mathrm{el}} = d_{\mathrm{el}}\boldsymbol{\sigma}$ to arrive at the Hamiltonian of the electric dipole interaction

$$\mathrm{H_E} = -\mathbf{d}_{\mathrm{el}} \cdot \boldsymbol{E}. \tag{12.4}$$

This electric dipole Hamiltonian, like any two-state operator, can be expressed in terms of the Pauli spin operator $\boldsymbol{\sigma}$, see (3.17a), which explains why $\mathbf{d}_{\mathrm{el}}$, like any vector operator, is proportional to $\boldsymbol{\sigma}$. In the following, we disregard static electric dipole moments, as well as the effect of the light-wave's magnetic field vector $\boldsymbol{B} = \boldsymbol{E}/c$, which latter is negligible in allowed electric-dipole transitions.

In laser spectroscopy, a time dependent transverse electric field $\boldsymbol{E}_\perp = \boldsymbol{E}_{\perp_0} \cos \Omega t$ with optical frequency Ω can induce electric-dipole transitions between two atomic energy levels $E_1 < E_2$, see also Sect. 18.3. We assume that the other energy levels of the atom are so far away, as compared to the line width of the transition, that they do not affect the resonance transition under study. The two atomic states then are the basis states $\psi_1 = \alpha$ and $\psi_2 = \beta$ of this atomic two-level system, E_1 and E_2 being the diagonal elements of the corresponding Hamiltonian 2×2 matrix. As before, we put the zero of energy halfway between these two diagonal elements, which

then have values $\pm\frac{1}{2}(E_2 - E_1)$ of the order of eV for low-lying atomic states. The laser field provides the off-diagonal matrix elements, which couple both states and induce transitions between them. For linearly polarized light with $\boldsymbol{E}(t)$ along x, the Hamiltonian then is

$$\mathrm{H} = \frac{1}{2}(E_2 - E_1)\sigma_z + 2d_{\mathrm{el}}E_{x0}\sigma_x \cos \Omega t, \tag{12.5a}$$

$$\mathrm{H} = \begin{pmatrix} \frac{1}{2}(E_2 - E_1) & 2d_{\mathrm{el}}E_{x0}\cos \Omega t \\ 2d_{\mathrm{el}}E_{x0}\cos \Omega t & -\frac{1}{2}(E_2 - E_1) \end{pmatrix}. \tag{12.5b}$$

Be aware that E_1, E_2 are energies, while E_{x0} is an electric field component.

The above Hamiltonian equations do not cover spontaneous emission, but describe semiclassically the process of *induced emission* and *induced absorption*. As before, we go to the frame rotating at the photon's angular frequency Ω. When we neglect the counter rotating laser amplitude, the Hamiltonian in the rotating wave approximation becomes

$$\mathrm{H}_{\mathrm{rot}} = \frac{1}{2}\hbar \begin{pmatrix} \Omega_0 - \Omega & \Omega_{\mathrm{R}} \\ \Omega_{\mathrm{R}} & -(\Omega_0 - \Omega) \end{pmatrix}. \tag{12.6}$$

This is the same Hamiltonian as in NMR, see (11.4b), with the energy $\hbar\Omega_0 = E_2 - E_1$ of the atomic optical transition replacing the energy $\hbar\omega_0 = \mu B_0$ of the Zeeman transition. The *Rabi precession frequency*

$$\Omega_{\mathrm{R}} = d_{\mathrm{el}}E_{x0}/\hbar \tag{12.7}$$

about the electric laser field amplitude E_{x0} replaces the precession frequency $\omega_1 = \mu B_1/\hbar$ about the magnetic RF-field amplitude B_1.

> In principle, all the methods developed in spin physics can be applied to free atoms exposed to a laser beam.

12.3 Optical Analogies of Spin Rotation and Spin Resonance

Most techniques developed in both spin rotation and spin resonance were realized, one after the other, also for the atomic two-state system, as there are quantum beats, Rabi oscillations, optical free induction decay (Brewer and Shoemaker 1972), optical adiabatic fast passage (Loy 1974), photon echoes (Brewer and Genack 1976), or Ramsey's method of separate oscillatory fields (Bergquist et al. 1977). We shall have a closer look on the first two methods.

Quantum beats are intensity modulations in the reemitted light from an excited atomic state that has a small energy splitting δE, after this state has been excited

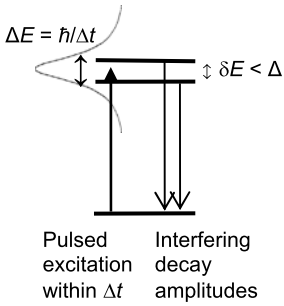


Fig. 12.2 Quantum beats are observed in the reemitted light if the spitting δE of the excited state is smaller than its energy uncertainty $\Delta E = \hbar/\Delta t$ due to pulsed excitation. Alternatively, quantum beats can be described as spin rotation of the excited atoms, detected by their rotating light emission pattern

by a very short light pulse of duration Δt, see Fig. 12.2. Quantum beats are an interference phenomenon: it is uncertain which of the two upper states is excited, if the splitting of the two levels is so small that it is blurred by the energy uncertainty of the incoming laser pulse,

$$\delta E \ll \Delta E = \hbar/\Delta t. \tag{12.8}$$

According to the rules given in Sect. 6.6, the two emitted light amplitudes then must be added coherently, which leads to a beating in the light intensity of frequency $\omega_0 = \delta E/\hbar$.

In nuclear physics, quantum beats are traded under the name *perturbed angular distributions* (PAD) or *perturbed angular correlations* (PAC), and are used to measure internal fields in condensed matter. Nuclear physicists generally use the spin precession picture, while atomic physicists prefer the description in terms of probability amplitudes. In the spin precession picture, the perturbed state is prepared by some nuclear or other reaction within a time interval Δt short enough that the state's transverse polarization does not dephase before spin rotation sets in. With an energy splitting δE of the perturbed state, spins rotate at $\omega_0 = \delta E/\hbar$, which leads to the requirement

$$\delta\varphi = \omega_0 \Delta t \ll 1, \tag{12.9}$$

which is identical to (12.8). The ensuing spin precession then is detected via another polarization sensitive reaction.

Needless to repeat, we prefer the spin precession picture with its three distinct steps, listed at the beginning of Chap. 4, because this picture does not depend on the way the perturbed state is prepared, or on the way it is analyzed, and describes phenomena as different as atomic quantum beats and the precession of polarized muons about an external magnetic field. More on this in Sect. 20.5.

In *Rabi oscillations*, what is it that precesses with angular frequency $\Omega_\mathrm{R} = d_\mathrm{el} E_{x0}/\hbar$, introduced in (12.7)? To find out, let us regard an optical resonance transition driven at angular frequency $\Omega_0 = (E_2 - E_1)/\hbar$. Let us first assume that the tunable laser is switched off and that all atoms are in their lower state ψ_1. For atoms in thermal equilibrium, the initial mixed-state *pseudopolarization* is

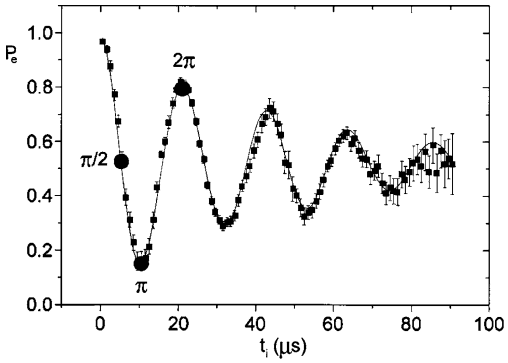


Fig. 12.3 Rabi oscillations of Rb atoms between two highly excited Rydberg states, separated by $\Omega_0/2\pi = (E_{51} - E_{50})/h = 51$ GHz. During the atoms' flight through a superconducting microwave cavity tuned to Ω_0, their pseudopolarization oscillates at the Rabi frequency $\Omega_R/2\pi = d_{el}E_x/h = 47$ kHz, for details see text. The *dots* in the figure indicate the interaction times for Rabi rotations by $\frac{1}{2}\pi$, π, and 2π. The damping is due to instrumental imperfections. From Raimond et al. (2001)

$P_z(0) = p_2 - p_1 = -1$. This atomic pseudopolarization has nothing to do with atomic spin, but is just a way of saying that the atoms are all in their ground state.

If the laser field $\boldsymbol{E}$ is switched on at the atomic resonance frequency $\Omega = \Omega_0$, then in the rotating-frame Hamiltonian equation (12.6), only the off-diagonal elements $H_{12} = d_{el}E_{x0} = \frac{1}{2}\hbar\Omega_R$ are nonzero. From our earlier discussions on spin precession, we know what will happen. The pseudopolarization will precess in the y–z plane about the rotating laser field component E_{x0} as $\boldsymbol{P}(t) = (0, \sin\Omega_R t, -\cos\Omega_R t)$, in the same way as shown in Fig. 11.1(b) for the NMR case, or, more precisely, from (6.20), $\boldsymbol{P}_{\text{mixed}}(t) = (p_2 - p_1)\boldsymbol{P}(t)$. After a time $T = \pi/\Omega_R$, the atoms will no longer be in their ground state, but will all be in their excited state ψ_2, the energy difference $\hbar\Omega_0$ being supplied by the laser field, and after twice that time, they will all be back to their ground state, under induced emission of a photon of energy $\hbar\Omega_0$. At all times in between, the atomic states are coherent superpositions of the atom's upper and lower energy states, moving along the $\varphi = \pi/2$ longitude of the appropriate Bloch sphere. We then go back to the laboratory frame and find that during Rabi oscillations, pseudopolarization $\boldsymbol{P}(t)$ evolves in the same way as the precession of real spins in an NMR experiments, Fig. 11.1(c).

In particular, the longitudinal polarization $P_z(t) = -\cos\Omega_R t = |a(t)|^2 - |b(t)|^2$, and with it the occupation probabilities $|a(t)|^2$ and $|b(t)|^2$, oscillate back and forth at the Rabi frequency Ω_R. Induced photon absorption alternates with induced photon emission, and the intensity of the reemitted light varies at the Rabi frequency. Figure 12.3 shows Rabi oscillations of rubidium atoms between two neighboring highly excited *Rydberg states* with principal quantum numbers $n = 50$ and 51.

When we tune the laser angular frequency Ω through the resonance, then far off resonance, $\Omega \ll \Omega_0$ or $\Omega \gg \Omega_0$, the system will answer with small effective-spin nutations, similar as in Fig. 5.2 for small transverse static fields, and at $\Omega = \Omega_0$ will

show the full Rabi-oscillation signal. The homogeneous width $\Delta\Omega$ of the resonance curve is given by the inverse of the atom's residence time in one of the two states, as required by the uncertainty relation. Hence, for large laser field amplitudes E_x, the width of the resonance signal is proportional to the laser field amplitude, $\Delta\Omega \propto E_x$, as was the signal width $\Delta\omega \propto B_x$ in magnetic resonance.

Thus Rabi oscillations differ considerably from the optical quantum beats discussed above. Quantum beats are spin rotations in the excited state of the type treated in Chap. 4, where the precession of (real or effective) spins is about some internal or external field that we did not specify any further. The laser provides one specific means of initiating and detecting this spin precession, but the amplitude of the laser field does not enter the 2×2 spin rotation Hamiltonian of the excited state. Rabi oscillations, on the other hand, in the present example are a pseudospin resonance effect, in which pseudospin precession is about the electric field vector of the laser; therefore, the laser amplitude *does* enter the Hamiltonian equation (12.6) of the atomic two-state system.

12.4 Particles in a Double Well

12.4.1 The NH_3 Molecule

A prominent cause for the splitting of energy levels is the tunnel effect. As an example we regard the ammonia molecule NH_3 in Fig. 12.4(a). If we squeeze the nitrogen atom along the z-axis through the equilateral triangle formed by the hydrogen atoms, then the N atom sees a *double well potential* $V(z)$ of the type shown in Fig. 12.4(b). The N atom can be on either side of the plane of the H-atoms. Hence the NH_3 molecule forms a two-state system with two basis states

$$\mathrm{L} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \mathrm{R} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{12.10}$$

The system can change from L to R and back because there is a certain probability amplitude A that the nitrogen atom tunnels through the central potential barrier to the other well.

The nitrogen atom has the same energy $H_{11} = H_{22} = E_0$ on both sides of the molecule, L and R being mirror images of each other. The system therefore has the two-state Hamiltonian

$$\mathrm{H} = \begin{pmatrix} E_0 & -A \\ -A & E_0 \end{pmatrix} = E_0 \mathrm{I} - A\sigma_x. \tag{12.11}$$

The tunneling amplitude $A > 0$ depends on the shape $V(z)$ of the potential, main parameters being the width and the height of the barrier. As we do not know $V(z)$ very well, we simply take the measured value for A.

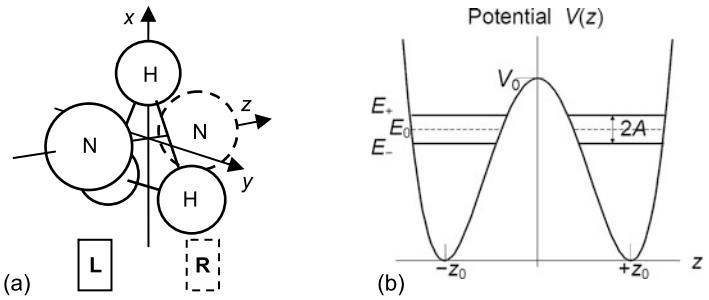


Fig. 12.4 (**a**) Structure of the ammonia molecule NH_3. The nitrogen atom can tunnel through the x–y plane from basis states L to R and back. (**b**) Tunnel splitting of the eigenenergy E_0 of the nitrogen atom in the double well potential of the ammonia molecule, with gap $2A \approx 1 \times 10^{-4}$ eV. The nitrogen atom N tunnels at a frequency $2A/h = 24$ GHz between positions L to R through the barrier of height $V_0 \approx 0.2$ eV

From (12.11) we can at once draw a number of conclusions without exact knowledge of $V(z)$. From the secular equation $(E_0 - E)^2 - A^2 = 0$ we find

$$E_\pm = E_0 \pm A. \tag{12.12}$$

A particle in a double-well potential shows a tunnel splitting of its energy.

The corresponding energy eigenstates of the molecule are symmetric and antisymmetric superpositions $(\mathrm{L} \pm \mathrm{R})/\sqrt{2}$, as in (3.54). This is in analogy to the frequency splitting of two classical coupled pendulums with its symmetric and antisymmetric normal modes. If at time $t = 0$ all molecules are in state L, their wave function evolves as $\psi = a(t)\mathrm{R} + b(t)\mathrm{L}$, and the probability of finding them in states L or R oscillates as

$$p_\mathrm{R} = |a|^2 = \cos^2(At/\hbar), \qquad p_\mathrm{L} = |b|^2 = \sin^2(At/\hbar). \tag{12.13}$$

As in Sect. 3.5.2, we can define a pseudopolarization $\boldsymbol{P}(t)$ with components

$$P_z(t) = |a|^2 - |b|^2 = \cos(2At/\hbar), \tag{12.14a}$$

$$P_y(t) = 2\,\mathrm{Im}\big(a^* b\big) = -\sin(2At/\hbar). \tag{12.14b}$$

The longitudinal component P_z indicates how the system oscillates between the states L and R. The transverse component P_y tells us that, during the tunneling processes, the system possesses phase dependent transverse amplitudes that let the pseudopolarization vector $\boldsymbol{P} = \langle \boldsymbol{\sigma} \rangle$ rotate about a vector $(A, 0, 0)$ in three-dimensional space. The transition matrix element A plays several roles.

- A determines the angular frequency $2A/\hbar$ at which the N-atoms oscillate between the states L and R. This corresponds to the Larmor precession frequency $\omega_\mathrm{L} = \gamma B_x$ in a transverse magnetic field.

- A determines the tunneling-induced energy splitting $E_+ - E_- = 2A$. This corresponds to the Zeeman splitting $\hbar\gamma B_x$ induced by a transverse magnetic field.
- The off-diagonal element A leads to coherence in the tunneling process, necessary for the periodicity of p_R and p_L in (12.13).
- In contrast to the magnetic dipole case, the off-diagonal term A of the ammonia molecule has a fixed value that cannot be varied experimentally.

Neither molecular vibrations nor rotations play a role in the tunneling process: at room temperature, only the lowest vibrational state of the molecule is populated, and the narrowly spaced rotational levels play no role because they all see essentially the same potential $V(z)$, and therefore experience the same tunnel splitting.

12.4.2 The Ammonia Maser

The tunnel splitting of the ammonia NH_3 molecule finds application in the ammonia MASER, an acronym for Microwave Amplification of Stimulated Emission of Radiation, similarly as LASER is an acronym for Light Amplification. While magnetic interactions of the NH_3 molecules are negligible, the molecules' dipole moments are very sensitive to external electric fields. Due to the differences in their electronegativities, each hydrogen atoms carries a mean positive excess charge ε^+, while the N-atoms carries a mean negative excess charge $\varepsilon^- = -3\varepsilon^+$. Hence, with the nitrogen atom in state L, the electric dipole moment of the molecule has amplitude $d_\mathrm{el} = \varepsilon^- z_0$, where z_0 is the mean distance of the N-atom to the x–y plane of the H-atoms. The electric dipole interaction energy then is $-\mathbf{d}_\mathrm{el} \cdot \boldsymbol{E}$. With the z-axis along the electric field, the Hamiltonian is

$$\mathrm{H} = E_0 \mathrm{I} - d_\mathrm{el} E_z \sigma_z - A\sigma_x, \tag{12.15}$$

and the energy eigenvalues $E_\pm$ in dependence of the electric field strength E_z lie on two hyperbolas defined by

$$(E - E_0)^2 - d_\mathrm{el}^2 E_z^2 = A^2, \tag{12.16}$$

similar as in the magnetic case, (7.4) and Fig. 7.1(a). Therefore, at small field E_z, we see a *quadratic Stark effect* which goes over into a *linear Stark effect* for strong E_z. With increasing field strength E_z, the negatively charged N-atoms are pulled more and more into the L state, and the probability to find them in the R state diminishes in the same way as the spin down probability in the magnetic case, (5.15a) and Fig. 7.1(b), namely, with

$$\cos^2 \frac{1}{2}\theta = \frac{1}{2}\left[1 + \frac{d_\mathrm{el} E_z}{(A^2 + d_\mathrm{el}^2 E_z^2)^{1/2}}\right]. \tag{12.17}$$

Just as the magnetization $M_z = n\mu P_z(B_z)$ is defined as the magnetic-field derivative of magnetic dipole energy, see (7.3), the effective electric dipole moment $\langle \mathrm{d}_\mathrm{el} \rangle$

is defined as the derivative of electric-dipole energy with respect to field amplitude E_z, which for the ground state energy E_- reads

$$\langle \mathrm{d_{el}} \rangle = \frac{\partial E_-}{\partial E_z} = \frac{d_{\mathrm{el}}^2 E_z}{(A^2 + d_{\mathrm{el}}^2 E_z^2)^{1/2}}. \tag{12.18}$$

At zero electric field, the slope of this curve is zero, and hence the electric dipole moment vanishes and changes sign with the applied field, as was already shown in Fig. 7.1(a) for the analogous magnetic case. Therefore:

Molecules do not have permanent, but only *induced electric dipole moments*.

The orientation of a particle can be specified only by the orientation of its angular momentum. Therefore, a permanent dipole moment would violate *invariance under time reversal* operation T, which changes the sign of $\langle \boldsymbol{\sigma} \rangle$ but not that of $\langle \mathbf{d}_{\mathrm{el}} \rangle$, so under time reversal the particle would change appearance.

Ammonia molecules in thermal equilibrium at room temperature have both energy levels E_+ and E_- nearly equally populated because of the smallness of the tunnel splitting $2A$. One can spatially separate the E_+ state from the E_- state by passing a beam of NH_3 molecules through an electric quadrupole lens with a strong radial electric field gradient. Like in the Stern–Gerlach case, a force

$$F_{r\pm} = -\frac{\partial E_\pm}{\partial r} = -\frac{\partial E_\pm}{\partial E_r}\frac{\partial E_r}{\partial r} = \mp \langle \mathrm{d_{el}} \rangle \frac{\partial E_r}{\partial r} \tag{12.19}$$

acts on the molecules, focusing the upper energy state E_+ and defocussing the lower state E_-. In this way, a *population inversion* is obtained, a necessary condition for maser and laser action. Indeed, the first maser was built before the invention of the laser. It used a beam of NH_3 molecules in the E_+ state, which, during passage of a tuned microwave resonator, made a transition to the E_- state, under stimulated emission of microwaves at the tunneling transition frequency of $2A/\hbar = 24$ GHz. Hence a maser amplifier can be used as a coherent source of microwaves.

12.4.3 Bose–Einstein Condensate in a Double Trap

Next we present an experiment in which instead of an atom an entire Bose–Einstein condensate of about 1000 atoms tunnels in a double-well potential. As bosons like to congregate (see Sect. 8.1), they can make a phase transition from a classical gas to a *Bose–Einstein condensate* (BEC) if temperature T is sufficiently low and density n is sufficiently high. We state without proof:

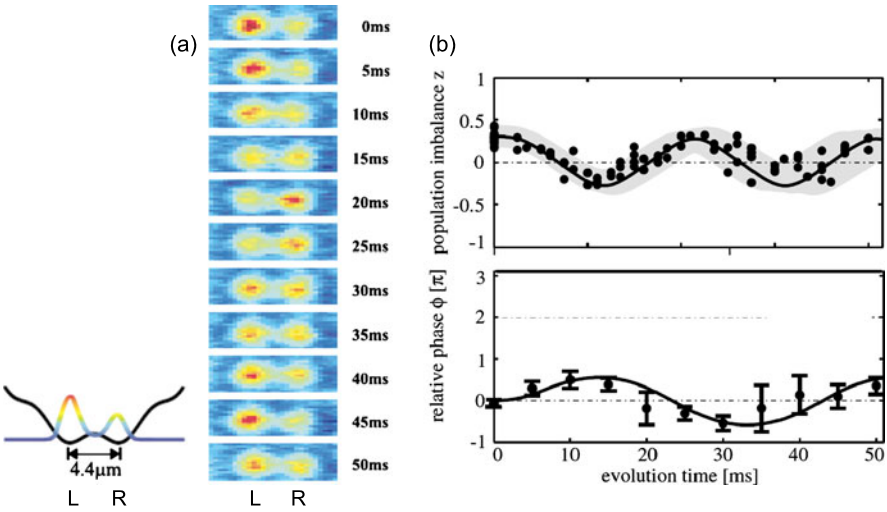


Fig. 12.5 Oscillations of a Bose–Einstein condensate in the potential of a double optical trap: (**a**) Measured time evolution of the populations of weakly linked ^{87}Rb Bose–Einstein condensates in the left and right wells of the double-well potential shown *on the left*. The size of the absorption images is about 10 μm × 20 μm. (**b**) *Upper panel*: Time dependence of the measured population imbalance, which is the longitudinal pseudopolarization $P_z(t)$ as derived from (**a**). (**b**) *Lower panel*: Time dependence of the measured phase factor of the transverse polarization $P_y(t)$ between the two BEC condensates in the double well, see (12.14b). The *solid lines* are theoretical predictions with no free parameter. From Albiez et al. (2005)

A phase transition of a gas to a Bose–Einstein condensate can occur if the phase space density exceeds one atom per phase space cell.

An atom at the origin $\boldsymbol{x} = \boldsymbol{p} = 0$ of phase space, and inside the cell volume $\Delta\boldsymbol{x}\Delta\boldsymbol{p} = \Delta x\Delta y\Delta z\Delta p_x\Delta p_y\Delta p_z$ has $\Delta\boldsymbol{x} \sim d^3$, where d is the mean distance between atoms, and $\Delta\boldsymbol{p} \sim p^3$, with $p = h/\lambda$. If d equals the de Broglie wavelength λ, then the cell volume becomes $\Delta\boldsymbol{x}\Delta\boldsymbol{p} = \lambda^3(h/\lambda)^3 = h^3$, as required by (6.39):

A Bose–Einstein transition can occur if the mean distance between neighboring atoms approaches their de Broglie wavelength.

A gas in thermal equilibrium has $p = (3Mk_BT)^{1/2}$, with atomic mass M and Boltzmann's constant k_B, and $d = 1/n^{1/3}$ for atomic density n. We resolve $\Delta\boldsymbol{x}\Delta\boldsymbol{p} \sim d^3p^3 = (3Mk_BT)^{3/2}/n = h^3$ for T and obtain the *critical temperature* for this phase transition

$$T_c \approx h^2n^{2/3}/3k_BM. \tag{12.20}$$

For ^{87}Rb atoms of density $n = 10^{12}$ cm^{-3}, $T_c \approx 0.1$ μK.

Bose–Einstein condensates have been the object of many investigations in recent years. Like single atoms, a BEC can be manipulated with *light forces*, and therefore can be stored in an *optical trap* formed by two counterpropagating laser beams, which produce an approximately harmonic three-dimensional confining potential. Likewise, in a standing light wave in front of a mirror, the BEC is confined to the crests of the light field, which is periodic in one dimension. If a harmonic three-dimensional trap and a one-dimensional periodic standing light wave are properly superimposed, one produces the optical double-well potential shown on the left of Fig. 12.5(a).

Shown are several consecutive snapshots of a BEC of rubidium atoms in this optical double-well potential, or more precisely, we see the shadow of such a BEC thrown onto a CCD camera when the BEC is illuminated from behind. If for $t = 0$ the BEC is in the left potential well L, then some tens of milliseconds later one finds the BEC in the right potential well R, and so on, back and forth. Figure 12.5(b) shows both the intensity oscillation L $\leftrightarrow$ R of the BEC's longitudinal pseudopolarization $P_z(t)$ and the transverse pseudopolarization $P_y(t)$ from (12.14b). To measure the phase φ of P_y, the double-well potential was switched off, and the partial BEC waves were dropped in the Earth's gravitational field and were allowed to interfere with each other. By tuning the parameters of the double-well potential, important information on the properties of the BEC can be deduced.

References

Albiez, M., Gati, R., Fölling, J., Hunsmann, S., Cristiani, M., Oberthaler, M.K.: Direct observation of tunneling and nonlinear self-trapping in a single bosonic Josephson junction. Phys. Rev. Lett. **95**, 010402(4) (2005)

Bergquist, J.C., Lee, S.A., Hall, J.L.: Saturated absorption with spatially separated laser fields: observation of optical "Ramsey" fringes. Phys. Rev. Lett. **38**, 159–162 (1977)

Brewer, R.G., Genack, A.Z.: Optical coherent transients by laser frequency switching. Phys. Rev. Lett. **16**, 959–962 (1976)

Brewer, R.G., Shoemaker, R.L.: Optical free induction decay. Phys. Rev. A **6**, 2001–2007 (1972)

Loy, M.M.T.: Observation of population inversion by optical adiabatic rapid passage. Phys. Rev. Lett. **32**, 814–817 (1974)

Raimond, J.M., Brune, M., Haroche, S.: Colloquium: manipulating quantum entanglement with atoms and photons in a cavity. Rev. Mod. Phys. **73**, 565–582 (2001)

Chapter 13
Two-State Systems in Condensed Matter

Abstract Condensed matter physics offers various examples of two-state quantum systems. We focus on two such examples. Glasses at very low temperature, with their excessive heat capacity and other peculiar properties, since long are known to behave as if they were some kind of a two-state quantum system. Spin echo experiments on glasses helped to elucidate the mechanism behind this phenomenon. The second example involves superconducting Josephson junctions, macroscopic two-state quantum devices that have applications in metrology, magnetometry, and quantum informatics.

13.1 Glasses

When cooled rapidly, many liquids avoid a phase transition to the crystalline state by turning into an amorphous glass. Glasses are in a metastable disordered state that has not had time to find its state of minimal energy. The macroscopic time scales for global changes in glasses can be of the order of many years, much longer than molecular time scales, which are of the order of picoseconds. Glasses solidify near the empirical *glass transition temperature*, though this transition has no theoretically well defined meaning. In spite of the practical importance of glasses, their slow dynamics is far from being satisfactorily described by current theories, see Berthier and Biroli (2011), Langer (2007), and references therein.

For most glasses, the glass transition temperature is between 200 K and 1000 K. At temperatures far below 1 K, glasses show another type of peculiar and universal behavior (Zeller and Pohl 1971). In this regime, glasses there have heat capacities too high by factors of one thousand and more, with a temperature dependence $C \propto T$, as compared to the usual $C \propto T^3$ from Debye theory, and in addition show a number of other anomalies. This was quite unexpected because phonons of long wavelength, mainly responsible for heat transport in condensed matter, were thought to be insensitive to the microscopic structure of matter.

These anomalies of glasses can be phenomenologically explained by assuming that atoms or groups of atoms in the glass are sitting in some double-well potential of unknown origin (Phillips 1973; Anderson et al. 1972). As amorphous glasses lack periodicity, these potentials, as seen by different atoms in a glass, should have a broad distribution of barrier widths and heights, and with it of tunnel splittings.

D. Dubbers, H.-J. Stöckmann, *Quantum Physics: The Bottom-Up Approach*,
Graduate Texts in Physics,
DOI 10.1007/978-3-642-31060-7_13, © Springer-Verlag Berlin Heidelberg 2013

These two-state systems can be described and investigated with the methods developed in previous chapters, for a survey see Phillips (1987). To this end, one defines a pseudospin-$\frac{1}{2}$ Hamiltonian matrix

$$\mathrm{H} = \frac{1}{2}(\Delta_{\mathrm{asy}}\sigma_z + \Delta_0\sigma_x). \tag{13.1}$$

As in Sect. 12.4, the off-diagonal Δ_0 is the matrix element responsible for tunneling between two wells, while the *asymmetry energy* Δ_{asy} on the diagonal accounts for unequal depths of two adjacent wells, the values of both parameters having a broad distribution. This two-state system then is coupled to time varying electric fields from thermally induced rearrangements of the surrounding atoms. These additional off-diagonal oscillatory elements lead to excitations and energy dissipation by phonon absorption. For a uniform distribution of tunnel splittings, the heat capacity (obtained in the usual way by integration over phonon energies, weighted with the phonon density of states) indeed gives the linear temperature dependence of the heat capacity $C \propto T$ observed in glasses. Other properties like thermal conductivity, thermal expansion, or energy relaxation can equally well be explained in terms of this two-state model.

In the laboratory, one can induce transitions between the putatively tunnel-split energy levels. To this end, one irradiates the glass either with photon pulses from an external microwave source, or with phonon pulses from a piezo oscillator coupled to the probe, both at frequencies in the GHz range. These pulses are applied in spin echo experiments, carried out as described in Sect. 11.2, with a $\pi/2$ flip followed by a π-flip of pseudospin, recall Fig. 11.6. For quite some time, this two-state approach was purely phenomenological, following the postulate:

> In a glass at very low temperatures, groups of atoms tunnel between different molecular configurations, separated by a potential barrier.

In recent years, the mechanism behind these interacting tunnel states was elucidated by such photon- or phonon-induced spin echo experiments (Enss 2002). For the glass former glycerol, for example, of molecular formula $C_3H_8O_3$, it was possible to identify a group of three hydrogen atoms which as a whole tunnel between two configurations, separated by a potential barrier. Figure 13.1 shows the configurations that are at the origin of the two-state behavior in this glass.

The results shown in the figure were obtained on the following way. Below about 50 mK, a quadrupole interaction can be identified, which is present in fully deuterated glycerol $C_3D_8O_3$ with deuteron spin $I = 1$, but not in ordinary $C_3H_8O_3$ with proton spin $I = \frac{1}{2}$ (cf. Sect. 7.1.3). With its broad range of interaction strengths, this quadrupole interaction washes out the spin echo signals and makes them disappear, unless a sufficiently strong magnetic field decouples the quadrupole interaction and makes the spin echo signal reappear. From the magnetic-field dependence of the spin echo amplitudes and from their time evolution for both natural and deuterated

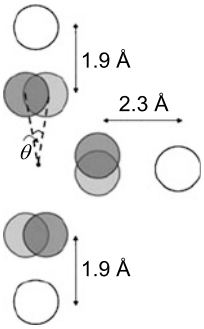


Fig. 13.1 In glass-like glycerol at very low temperatures, three hydrogen atoms tunnel as a whole between two stable triangular configurations, separated from each other by a potential barrier, without changing the shape of the triangle, see the *dark and light gray spheres*. The in-plane rotation angle is measured to $\theta = 16°$. The white spheres denote neighboring stationary H-atoms. From Fickenscher et al. (2009)

glycerol, it was possible to identify the two-state configurations responsible for the low-temperature glass anomalies, shown in Fig. 13.1.

13.2 Josephson Effects

13.2.1 Basics of Superconductivity

We turn to another well known two-state system, the *Josephson junction* of two weakly coupled superconductors. Though let us first recall some facts about superconductivity. A superconducting current meets zero resistance and is carried by a small fraction of conduction electrons that form the *Cooper pairs* in the following way. When moving through the crystal lattice, the (dielectrically screened) electron charge induces a distortion of the surrounding crystal lattice. When two such distortions overlap, this leads to an attractive interaction of the two electrons, essentially an electron–phonon interaction. The electrons in a Cooper pair formed in this way carry their spins and momenta opposite to each other. In a classical analogue, two pins are supported on a water surface by surface tension, which attract each other and at the same time turn parallel.

Below a critical temperature T_c, the spinless Cooper pairs as bosons form a condensate of the type of the Bose–Einstein condensate discussed in Sect. 12.4.3.

The electronic ground state of a superconductor is a boson condensate that is separated from the normally-conducting electron states by an energy gap of several meV.

In a metal, electrical resistance has its origin in the inelastic scattering of the conduction electrons on deviations from lattice periodicity, arising from lattice defects or from lattice vibrations, the latter dominating at high temperatures. In a superconductor, on the other hand, for $T < T_c$, the inelastic scattering of superconducting electrons across the energy gap Δ into their normally-conducting state is suppressed, and electrical resistance disappears. A voltage applied across a superconductor will make the entire superconducting coherent ensemble move through the crystal as a supercurrent.

We can mimic the formation of such a coherent bosonic condensate of Cooper pairs with a highly simplified toy model. We take as basis states the subset of those pairs of electrons that already have antiparallel momenta and spins, and we assume that the matrix elements responsible for the attractive interaction of such pair states all have the same small negative value $\langle n'_\uparrow n'_\downarrow | \mathrm{H} | n_\uparrow n_\downarrow \rangle = -\delta$ for $n' \neq n$, and are zero otherwise. The Hamiltonian and its eigenvalues therefore are, for the example of $N = 5$ Cooper pairs,

$$\mathrm{H} = -\delta \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{pmatrix}, \tag{13.2}$$

$$E_1 = E_2 = E_3 = E_4 = +\delta, \qquad E_5 = -4\delta. \tag{13.3}$$

All eigenstates have their energy $E_1 = \cdots = E_{N-1}$ increased by δ, except for one state whose energy E_N is lowered to $-(N-1)\delta$, such as to keep the energy balance at zero. The corresponding (unnormalized) eigenfunctions are

$$\psi_1 = (-1, +1, 0, 0, 0), \tag{13.4a}$$

$$\psi_2 = (-1, 0, +1, 0, 0), \tag{13.4b}$$

$$\psi_3 = (-1, 0, 0, +1, 0), \tag{13.4c}$$

$$\psi_4 = (-1, 0, 0, 0, +1), \tag{13.4d}$$

$$\psi_5 = (1, 1, 1, 1, 1), \tag{13.4e}$$

as can be checked by numerical calculation. The proof for general N is straightforward and is omitted here. The wave function of the last eigenstate is a coherent superposition of all Cooper pairs involved, and is separated from other states by a gap of $\Delta = E_1 - E_N = N\delta$. This last state forms the superconducting condensate. For a large number N of Cooper pairs, this gap can reach the order of meV even for extremely small δ. The above simplified treatment is meant to give an indication as to how a condensate with an energy gap may come into existence in a many-body quantum system.

13.2.2 Josephson Junctions and Their Applications

We now come to our second example of a two-state system in condensed matter. When two pieces of the same superconducting material are joined together, separated from each other only by an insulating oxide layer a few atomic layers thick, the Cooper pairs can tunnel through this layer from one side of the junction to the other. In the early 1960s, Josephson had theoretically studied such junctions and predicted a number of peculiar effects, which were realized experimentally shortly after, and which carry his name. At the time, Josephon's predictions were met with disbelief because the idea that coherent Cooper pairs could tunnel through an isolator without disruption was refuted by most theorists. We treat three Josephson effects and their applications, (i) the *DC*, (ii) the *AC*, and (iii) the inverse Josephson effect, plus (iv) another useful device called a SQUID.

We first derive the basic properties of a Josephson junction. Let ψ_1 be the wave function of the superconducting Cooper pairs on one side of the junction, and ψ_2 that on the other side, and let A be the probability amplitude for their tunneling across the junction. The coupled system then has the same Hamiltonian $\mathrm{H} = A\sigma_x$ as any other coupled two-state tunnel system, with E_0 again set to zero. In a complete theory, we would need to treat the whole device as a single quantum system, including the interfaces connecting the terminals of the battery. Our simplified treatment will turn out to give the correct result, but we must be aware that it is insufficient.

Let a voltage U be applied across this device by connecting both ends of the junction to a battery, which acts on the charges $q = -2e$ of the Cooper pairs. The voltage drop is only across the insulating layer of the junction, with potentials $\pm\frac{1}{2}qU$ on the two sides, but not inside the two superconductors.

With the Hamiltonian $\mathrm{H} = -eU\sigma_z + A\sigma_x$ the Schrödinger equation $\dot{\psi} = -(\mathrm{i}/\hbar)\mathrm{H}\psi$ for this coupled quantum system has two parts:

$$\dot{\psi}_1 = \frac{\mathrm{i}}{\hbar}(eU\psi_1 - A\psi_2), \tag{13.5a}$$

$$\dot{\psi}_2 = -\frac{\mathrm{i}}{\hbar}(eU\psi_2 + A\psi_1). \tag{13.5b}$$

So far there is nothing new. A main difference, though, to all previous cases is that we have to deal with *macroscopic wave functions* of Cooper pairs that are coherent throughout the bulk of the superconductor. Therefore, we set

$$\psi_1 = \sqrt{n_1}\mathrm{e}^{\mathrm{i}\theta_1}, \qquad \psi_2 = \sqrt{n_2}\mathrm{e}^{\mathrm{i}\theta_2}, \tag{13.6}$$

with the macroscopic Cooper pair densities $n_i = |\psi_i|^2$ and the phases θ_1 and θ_2 of the wave functions on the two sides of the junction, with time derivative

$$\dot{\psi}_i = \left(\frac{\dot{n}_i}{2n_i} + \mathrm{i}\dot{\theta}_i\right)\psi_i, \quad i = 1, 2. \tag{13.7}$$

The ψ_i must obey both the conventional two-state equations (13.5a), (13.5b) and the new macroscopic equations (13.7). This gives us two macroscopic equations for the densities n_i and for the phase difference $\delta = \theta_2 - \theta_1$.

Subtracting the real and imaginary parts of (13.5a), (13.5b) from those of (13.7) gives four expressions that we call R_1, R_2, I_1, I_2. With $c_i = \cos\theta_i$, $s_i = \sin\theta_i$, the expressions $R_i c_i + I_i s_i = 0$ $(i = 1, 2)$ and $R_1 s_1 - I_1 s_1 - R_2 s_2 + I_2 s_2 = 0$ give equations for the Cooper pair densities and for the phase difference.

First, for the densities we obtain

$$\dot{n}_1 = \frac{2A}{\hbar} n \sin\delta = -\dot{n}_2. \tag{13.8}$$

The Cooper pair densities n_i must both be the same because both parts of the junction are made from the same material. The gain in Cooper pair density $\dot{n}_1$ on one side is compensated by the loss $-\dot{n}_2$ on the other side. The role of the battery is to keep the overall electron density at a constant level n. In (13.8), no voltage is involved, still a current $I = 2e\dot{n}_1$ of Cooper pairs of charge $q = 2e$ is expected. Our first result is:

(i) The *DC Josephson effect*: Without a voltage applied, there is a *DC* tunneling current of Cooper pairs across a Josephson junction of size

$$I = I_0 \sin\delta. \tag{13.9}$$

With a device dependent phase δ across the junction, this current can have values $-I_0 \le I \le +I_0$, with $I_0 = (2Ae/\hbar)n$.

Second, for the phase difference $\delta = \theta_2 - \theta_1$ across the junction we obtain

$$\dot{\delta} = \frac{2e}{\hbar} U, \tag{13.10}$$

which depends only on the voltage U across the junction, but not on the properties of the junction that enter the tunneling amplitude A. The current-phase relation (13.9) and the phase evolution equation (13.10) describe the dynamics of a Josephson junction. With a voltage U applied, the phase difference, starting at some device dependent δ_0, grows with time t as

$$\delta(t) = \delta_0 + \int_0^t U(t')\,\mathrm{d}t'. \tag{13.11}$$

With a constant voltage $U = U_0$ applied, this growth is linear, $\delta = \delta_0 + \omega_0 t$. From (13.9) then follows

(ii) The *AC Josephson effect*: If a constant voltage is applied across a Josephson junction, an alternating current flows through the junction

$$I = I_0 \sin(\delta_0 + \omega_0 t). \tag{13.12}$$

Its angular frequency is

$$\omega_0 = \frac{2e}{\hbar} U_0. \tag{13.13}$$

Hence Josephson junctions act as perfect voltage-to-frequency converters, used worldwide as voltage standards, with frequency and voltage connected by the *Josephson constant* $K_\mathrm{J} = 2e/h = 0.4835979$ GHz/μV. As a general rule, high-precision measurements of analogue quantities are best done electronically in terms of a high frequency.

The Josephson constant K_J is the inverse of the *magnetic flux quantum* $\Phi_0 = h/2e$, with the charge $2e$ of a Cooper pair, which is the smallest unit of magnetic flux that can be supported by a ring made of superconducting material. We had invoked the quantity $h/e = 2\Phi_0$, with the charge e of single electrons, in Sect. 7.3.4 on the Aharonov–Bohm effect.

In the time average, the Josephson AC current (13.12) vanishes. If the voltage U_0 that generates this AC current is turned off, without disconnecting the voltage supply, then the constant current of the DC Josephson effect reappears, see (13.9). Hence, we have a nonzero DC current for zero voltage, and zero DC average current for nonzero voltage, not quite what we expect from Ohm's law.

Next we turn to the inverse Josephson effect and apply a high-frequency voltage $u\cos\omega t$ of small amplitude $u \ll U_0$ across the junction, in addition to the constant voltage U_0. Whenever its frequency ω coincides with the DC voltage-induced frequency ω_0, or with an integer fraction thereof, then a DC current $\bar{I} \neq 0$ across the junction reappears, as we shall prove below.

(iii) In the *inverse Josephson effect*, the junction acts as a perfect frequency to voltage converter.

The inverse Josephson effect is applied in gauging offices worldwide as a reference voltage supply.

With the additional high-frequency voltage $u\cos\omega t$ applied, the phase in (13.11) is

$$\begin{aligned}\delta(t) &= \delta_0 + (2e/\hbar)\int_0^t (U_0 + u\cos\omega t')\,\mathrm{d}t' \\ &= \delta_0 + \omega_0 t + (\omega_0/\omega)(u/U_0)\sin\omega t.\end{aligned} \tag{13.14}$$

With $\varepsilon \equiv (\omega_0/\omega)(u/U_0) \ll 1$ we have

$$\sin\delta(t) \approx \sin(\delta_0 + \omega_0 t) + \varepsilon \sin\omega t \cos(\delta_0 + \omega_0 t). \tag{13.15}$$

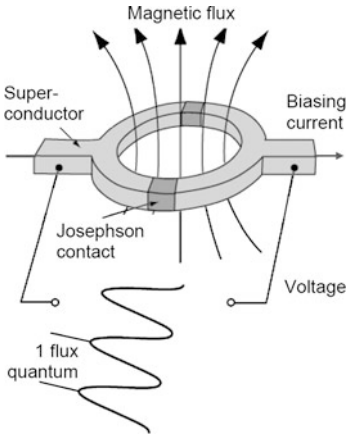


Fig. 13.2 In a SQUID magnetometer, with increasing magnetic field the current through the two parallel Josephson junction oscillates once back and forth for each additional flux quantum Φ_0. Magnetic fields are monitored by counting the changes in the number of flux quanta $\Phi_0 = 2.07 \times 10^{-15}$ Tm^2 going through the loop. From Wikipedia public domain

As the time average of this is zero, the DC current vanishes, $\bar{I} = I_0 \overline{\sin\delta} = 0$, unless the externally applied frequency coincides with the voltage-induced frequency, $\omega = \omega_0$. In this case, the time average of the second term on the right of (13.15) becomes nonzero, being the time average of $I_0 \varepsilon \sin(\omega_0 t)\cos(\delta_0 + \omega_0 t)$, hence $\bar{I} = \frac{1}{2}(u/U_0) I_0 \cos\delta_0$, which in general is nonzero. Similar signals appear for commensurate frequencies $\omega = \omega_0/n$, with integer n.

The SQUID, short for Superconducting QUantum Interferometry Device, is a superconducting loop with two Josephson junctions in parallel, see Fig. 13.2, used as a hypersensitive magnetometer. If a magnetic field B is applied at right angles to this loop, then the Aharanov–Bohm effect on the Cooper pairs shifts the phase around the loop in proportion to B, as in (7.25), but with e replaced by $2e$. This change of phase takes place across the two junctions, and changes the current $I = I_0 \sin\delta$ flowing through them. With rising field B, for every additional flux quantum $\Phi_0 = h/2e$, the current through the device changes from zero to maximum and back.

(iv) In a SQUID, changes in magnetic flux $\Phi = \int \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{a}$ can be counted in units of the magnetic flux quantum $\Phi_0 = 2.07 \times 10^{-15}$ Wb.

Recall that 1 Wb = 1 Tm^2. SQUIDs have important applications in medicine (magnetoencephalography, cardiology), industry (magnetic property measurement systems, oil prospecting, mineral exploration), and space and other research, where magnetic fields down to 10^{-18} T can be monitored with a SQUID.

References

Anderson, P.W., Halperin, B.I., Varma, C.M.: Anomalous low-temperature thermal properties of glasses and spin glasses. Philos. Mag. **25**, 1–9 (1972)

Berthier, L., Biroli, G.: Theoretical perspectives on the glass transition and amorphous materials. Rev. Mod. Phys. **83**, 587–645 (2011)

Enss, C.: Beyond the tunnelling model: quantum phenomena in ultracold glasses. Physica B **316–317**, 12–20 (2002)

Fickenscher, G., Bazrafshan, M., Reinhold, K., Fleischmann, A., Enss, C.: Influence of nuclear dipole-dipole interaction on atomic tunnelling systems in glycerol. J. Phys. Conf. Ser. **150**, 1–4 (2009)

Langer, J.: The Mysterious Glass Transition. Phys. Today, **Feb**, 8–9 (2007)

Phillips, W.A.: Tunneling states and the low-temperature thermal expansion of glasses. J. Low Temp. Phys. **11**, 757–763 (1973)

Phillips, W.A.: Two-level states in glasses. Rep. Prog. Phys. **50**, 1657–1708 (1987)

Zeller, R.C., Pohl, R.O.: Thermal conductivity and specific heat of noncrystalline solids. Phys. Rev. B **4**, 2029–2041 (1971)

This page intentionally left blank

Chapter 14
Two-State Systems in Nuclear and Particle Physics

Abstract The quantum states "neutron" and "proton" are another two-state effective-spin system, called an isospin doublet. We first answer why we need isospin and discuss its role in nuclear physics. We then generalize the concept of isospin, and introduce the quantum numbers flavor and color of particle physics. For neutrinos and neutral mesons, a phenomenon called flavor oscillations is observed. Neutron–antineutron oscillations could also be allowed, but have not been observed, yet.

14.1 Isospin

The proton and the neutron, collectively called the *nucleons*, form a two-state system called an *isospin-*$\frac{1}{2}$ system. The concept of isospin was introduced by Heisenberg shortly after the discovery of the neutron. It is based on the assumption that the strong nuclear force is the same for the proton and the neutron. Indeed, in all nuclear reactions, proton–proton, proton–neutron, and neutron–neutron interactions give very similar results. Only the much weaker electromagnetic Coulomb interaction between the protons requires a small correction.

A strong support for this *isospin symmetry*, also called *charge independence of nuclear forces*, comes from the γ spectra of *mirror nuclei* which turn out to be almost identical. In a pair of mirror nuclei with a given nuclear mass number $A = Z + N$, the number N of neutrons of one mirror nucleus equals the number Z of protons of the other mirror nucleus. An example is the mirror pair ^{12}B with $Z = 5$, $N = 7$ and ^{12}N with $Z = 7$, $N = 5$, also written as ${}^{12}_{7}\mathrm{B}_5$ and ${}^{12}_{5}\mathrm{N}_7$. For the strong-interaction states of these nuclei, it apparently does not make a difference if a proton is replaced by a neutron, or vice versa. Isospin is shorthand for *isobaric* spin. In the *nuclear chart* in which the isotopes are assembled in a scheme Z vs. N, isobaric nuclei have all the same mass number and lie on a minor diagonal $A = Z + N =$ const. from the upper left to the lower right of the chart.

In the isospin description, the neutron and the proton are written as the basis states p and n. With respect to the strong interaction, it is undetermined whether a nucleon is a neutron, a proton, or any linear combination $\psi = a\mathrm{p} + b\mathrm{n}$ of the two. The state ψ can be chosen anywhere on the corresponding Bloch sphere of the nucleon, and any such ψ will give the same expectation value for any strong-interaction operator.

D. Dubbers, H.-J. Stöckmann, *Quantum Physics: The Bottom-Up Approach*, Graduate Texts in Physics,
DOI 10.1007/978-3-642-31060-7_14, © Springer-Verlag Berlin Heidelberg 2013

For the two-state system "electron", we had solely used the fact that it has two internal states called spin up and spin down to define the Pauli spin operator $\boldsymbol{\sigma}$, which was represented by the 2×2 matrices in (3.8b) with respect to the basis states α and β. In the same way, for the two-state system "nucleon", we solely use the fact that it has two internal states that we call proton and neutron to define the isospin-$\frac{1}{2}$ operator $\boldsymbol{\tau}$, represented by the same 2×2 Pauli matrices in (3.8b), but with respect to the basis states p and n. One uses a different symbol $\boldsymbol{\tau}$ for isospin-$\frac{1}{2}$ because the corresponding Pauli matrices operate in a different space, the space of nucleon wave functions. Likewise we define an operator $\mathbf{t} = \frac{1}{2}\hbar\boldsymbol{\tau}$, in analogy to the spin operator $\mathbf{s} = \frac{1}{2}\hbar\boldsymbol{\sigma}$. The proton state p and the neutron state n then are eigenstates both of $\mathbf{t}^2$ with eigenvalues $t(t+1)\hbar^2 = \frac{3}{4}\hbar^2$, and of t_z with eigenvalues $t_z = \pm\frac{1}{2}\hbar$.

What could be a Hamiltonian acting on such single-nucleon states p and n? Because the nuclear force is very short-ranged, in general there is nothing that would resemble a classical external $\boldsymbol{B}$ field in the magnetic spin Hamiltonian. The only 2×2 Hamiltonian in our context is the nucleon *mass matrix* $\mathrm{M} = \mathrm{H}/c^2 = m_{\mathrm{nucleon}}\mathrm{I}$ with unit matrix I and with nucleon masses $m_{\mathrm{nucleon}} = m_p = m_n$ assumed to be equal, which corresponds to our earlier energy base line E_0. Therefore, isospin plays a role mainly in coupled nucleon systems. Isospin symmetry is not perfect because neutron mass exceeds proton mass by about one part per thousand.

All this sounds perfectly reasonable. Still, the fact that the proton and the neutron should behave in all respects like the up and down components of a quantized angular momentum usually is accepted rather by force of habit than by conviction. To strengthen the case, we add the following observations. The nuclear chart of isotopes is constructed by filling up the quantized nuclear energy levels one by one with fermionic nucleons, in a similar way as the periodic table of elements is constructed by filling up the quantized atomic energy levels one by one with fermionic electrons.

In both cases, the Pauli-exclusion principle plays a decisive role. For the Pauli principle, all that counts is whether the constituents of a system are the same or whether they differ in some respect, no matter in what respect. Hence when filling the nuclear chart, the properties "proton" or "neutron" play exactly the same role as the properties "spin up" or "spin down". For the required symmetrization procedures under particle exchange to work (see Sects. 8.1 and 8.2), both spin and isospin must follow the same mathematical rules, the *angular momentum algebra* of quantum mechanics. Rather than saying that isospin behaves like angular momentum, we recall the following:

> The statement "a particle has spin $s = \frac{1}{2}$" has no other meaning than that the particle has two internal degrees of freedom called "spin up" and "spin down" that impose the 2×2 spin algebra. In the same way, the two internal degrees of freedom "proton" and "neutron" of the strong interaction impose the 2×2 isospin algebra on the nucleon.

The equivalence between spin and isospin has remarkable consequences.

- In the same way as the individual nucleon angular momenta add up to nuclear total spin **I** (which may comprise spin and orbital angular momenta), the isospins of the nucleons add up to a total nuclear isospin **T**.
- In nucleon–nucleon interactions, the coupling of two ordinary spins is usually described by a term $V_{\mathrm{S}}(\boldsymbol{r})\mathbf{S}_1 \cdot \mathbf{S}_2$, as in (8.12). Similarly, the isospin dependence of the nuclear force is described by additional terms $V_{\mathrm{T}}(\boldsymbol{r})\mathbf{T}_1 \cdot \mathbf{T}_2$ or $V_{\mathrm{ST}}(\boldsymbol{r})(\mathbf{T}_1 \cdot \mathbf{T}_2)(\mathbf{S}_1 \cdot \mathbf{S}_2)$ in the nuclear Hamiltonian.
- Like ordinary angular momentum **I**, isospin **T** is conserved in a closed, i.e., isolated, system like a free nucleus. By *Noether's theorem*, conservation of angular momentum is linked to isotropy of space, that is, to the invariance of interactions under rotations in ordinary space. In the same way, conservation of isospin is linked to the invariance of the strong interaction under rotations in isospin space. While nuclear forces depend on isospin quantum number T, under isospin symmetry they are independent of its component T_z.
- For ordinary spin I, a nuclear state is a multiplet of $2I+1$ different m_I substates, $m_I = -I, \ldots, I$, where $I_z = \frac{1}{2}(N_+ - N_-)$ for N_+ nucleons with spin up, N_- nucleons with spin down. Similarly, a nuclear *isospin T multiplet* has $2T+1$ different T_z substates $T_z = -T, \ldots, T$, made up of isobaric isotopes $A = Z + N = \text{const.}$, with

$$T_z = \frac{1}{2}(Z - N). \tag{14.1}$$

  $T = T_z = 0$ implies $Z = N$, that is, all $T_z = 0$ nuclei, and only those, lie on the main diagonal of the nuclear chart.
- The total nuclear wave function is a product of spatial, spin, and isospin functions, and, by Pauli's principle, is antisymmetric under the exchange of any two of its constituents, the fermionic nucleons.

To give an example of nuclear states that obey these rules, we look at the boron isotope ${}^{10}_{5}\mathrm{B}_5$ in its ground state with $A = 10$ nucleons. Each nuclear level can hold at most four nucleons, namely, the four isospin–spin direct-product states

$$\mathrm{p}\alpha, \quad \mathrm{p}\beta, \quad \mathrm{n}\alpha, \quad \mathrm{n}\beta, \tag{14.2}$$

which are single-particle states, in contrast to (8.1), which are two-particle states. The first two energy levels of ${}^{10}\mathrm{B}$ are each filled with four nucleons coupling to total spin zero and total isospin zero. The third nuclear level is only half filled with one proton and one neutron, and all higher energy levels are empty. We then add two more nucleons to form a nucleus with mass number $A + 2 = 12$. If the two nucleons are placed in the half filled third level, then Pauli's principle requires them to be a proton and a neutron, with spins opposite to the spins of the proton and neutron already present, which gives the ground state of ${}^{12}_{6}\mathrm{C}_6$ with $I = T = 0$.

If instead one of the two additional nucleons occupies the empty fourth nuclear level, then there is no restriction from Pauli's principle, and the configurations can be

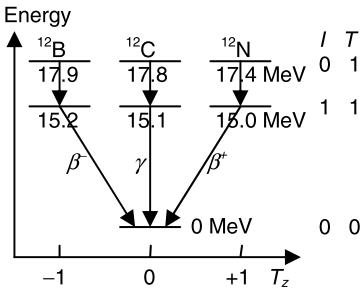


Fig. 14.1 Spin and isospin assignments in the isospin triplet ^{12}B, ^{12}C, and ^{12}N. The nuclear level scheme is shown for three values $T_z = \frac{1}{2}(Z - N) = 0$, and ± 1. The ^{12}C ground state is a spin and isospin singlet. The excited states are the isospin $T = 1$ triplets with spin $I = 1$ and $I = 0$

either singlets or triplets, both for spin and isospin. The lowest isospin $T = 1$ triplet has $I = 1$ and lies about 15 MeV above the stable isosinglet ground state of ^{12}C. This isotriplet comprises the boron isotope $^{12}_{7}\mathrm{B}_5$ with $T_z = -1$, the excited $^{12}_{6}\mathrm{C}^*_6$ isotope with $T_z = 0$, and the nitrogen isotope $^{12}_{5}\mathrm{N}_7$ with $T_z = +1$, see the nuclear level scheme Fig. 14.1. The energies of these $T = 1$ triplets are almost degenerate. The energy of this spin $I = 1$ triplet differs from the energy of the next spin $I = 0$ singlet, both with $T = 1$, by about 15 %, due to the sizable spin dependence of the nuclear force.

Isospin conservation requires that isospin T must be the same on the left and on the right of any reaction formula. This requirement may have drastic consequences. For example, it is possible to excite the $T = 1$ level of ^{12}C by inelastic scattering of protons on the $T = 0$ ground state of ^{12}C. However, if we try to do the same thing with deuterons, which have just one neutron added to the proton, nothing happens. The reason: protons have $T = \frac{1}{2}$ and $T_z = +\frac{1}{2}$. Proton scattering $^{12}\mathrm{C} + p \to {}^{12}\mathrm{C}^* + p$ is allowed because isospin on the left is $T = 0 + \frac{1}{2} = \frac{1}{2}$. On the right, isospins 1 and $\frac{1}{2}$ can couple to $T = \frac{1}{2}$ or to $T = \frac{3}{2}$. By isospin conservation, only the $T = \frac{1}{2}$ channel contributes to the reaction with weight $|\langle 1, 0, \frac{1}{2}, \pm\frac{1}{2} | \frac{1}{2}, \pm\frac{1}{2}\rangle|^2 = \frac{1}{3}$, as determined by the appropriate Clebsch–Gordan coefficients. The $T = \frac{3}{2}$ channel, however, in spite of its twofold weight $|\langle 1, 0, \frac{1}{2}, \pm\frac{1}{2} | \frac{3}{2}, \pm\frac{1}{2}\rangle|^2 = \frac{2}{3}$ does not contribute.

The state of the deuteron, on the other hand, is found to be an antisymmetric isospin singlet with $T = 0$. On the left of the scattering reaction $^{12}\mathrm{C} + d \to {}^{12}\mathrm{C}^* + d$ therefore $T = 0 + 0 = 0$, but on the right $T = 1 + 0 = 1$. Therefore, this reaction is isospin forbidden, and only protons can promote ^{12}C to its first excited state.

In conclusion, in strong-interaction processes, isospin has the same crucial importance as has angular momentum.

14.2 Flavor and Color

Next we generalize the concept of spin and isospin to other multivalued quantum numbers. The isospin quantum numbers $T_z = \pm\frac{1}{2}$, formerly referring to the proton and the neutron, in particle physics are attributed to the up quark u and the down quark d. *Baryons* like the proton *uud* and neutron *udd* are composed of three quarks, written as *qqq*. *Mesons* like the pions π^+, π^0, π^- are bound quark–antiquark states $q\bar{q}$. When more baryons like Σ^+, Σ^0, Σ^- and mesons like K^+, K^0, $\bar{K}^0$, K^- were discovered, a new quantum number *strangeness* S was introduced to accommodate these particles in new larger multiplets.

When these additional particle states are constructed from quarks, the basic building blocks are no longer two-valued as was isospin with quark states u and d as building blocks, but are three-valued, with three quark flavors u, d, and s. In group theory, one says that these quarks have $SU(3)$ flavor symmetry, instead of $SU(2)$ symmetry of spin or isospin.

In ordinary spin physics, to couple two spin-$\frac{1}{2}$ particles, we first had taken the direct-product states αα, αβ, βα, and ββ, with substate quantum numbers added to $m = m_1 + m_2$. These direct-product states then were rotated into an antisymmetric $s = 0$ spin singlet and a symmetric $s = 1$ spin triplet, see Sect. 8.2. In group theory, this procedure is symbolized as $2 \otimes 2 = 1 \oplus 3$, both for spin and isospin.

In quark physics, mesons, for example, are formed by coupling one of the u, d, or s quarks with one of the $\bar{u}, \bar{d}$ or $\bar{s}$ antiquarks. This task must be accomplished with group-theoretical methods. Here we merely demonstrate graphically what kind of particle multiplets can be expected from such couplings. To this end, we represent each quark by a dot in the T_z–S plane isospin vs. strangeness, with quark u at position $(T_z, S) = (+\frac{1}{2}, 0)$, d at $(-\frac{1}{2}, 0)$, and s at $(0, 1)$, and the same for the antiquarks with all signs inverted, see Fig. 14.2(a). We then form the nine direct-product states $q\bar{q}$, add their substate quantum numbers T_z and S, and put them into the T_z–S plane, as shown in Fig. 14.2(b). One can abbreviate this procedure by placing the midpoint of the antiquark triangle on the right of Fig. 14.2(a) onto each corner of the particle triangle on the left of the same figure, redrawn in Fig. 14.2(b) with dashed lines. Finally, these direct-product states are rotated into appropriately (anti-) symmetrized states, giving an octet, composed of pions, kaons, and the η meson shown in Fig. 14.2(c). The ninth meson, the η', must be in a separate singlet, because the $T_z = S = 0$ center position cannot accommodate more than two particles (in two orthogonal isospin states) without violating Pauli's principle. The multiplets shown are for mesons with spin $I = 0$. Similar multiplets exist for mesons with spin $I = 1$.

We summarize flavor $SU(3)$ results:

> The coupling of u, d, or s quarks and of $\bar{u}$, $\bar{d}$ or $\bar{s}$ antiquarks leads to a meson singlet and a meson octet, symbolized as $3 \otimes 3 = 1 \oplus 8$.

In a similar way, the coupling of three quarks leads to one baryon singlet, two baryon octets, and one baryon decouplet, $3 \otimes 3 \otimes 3 = 1 \oplus 8 \oplus 8 \oplus 10$, all of which are found to be realized in nature.

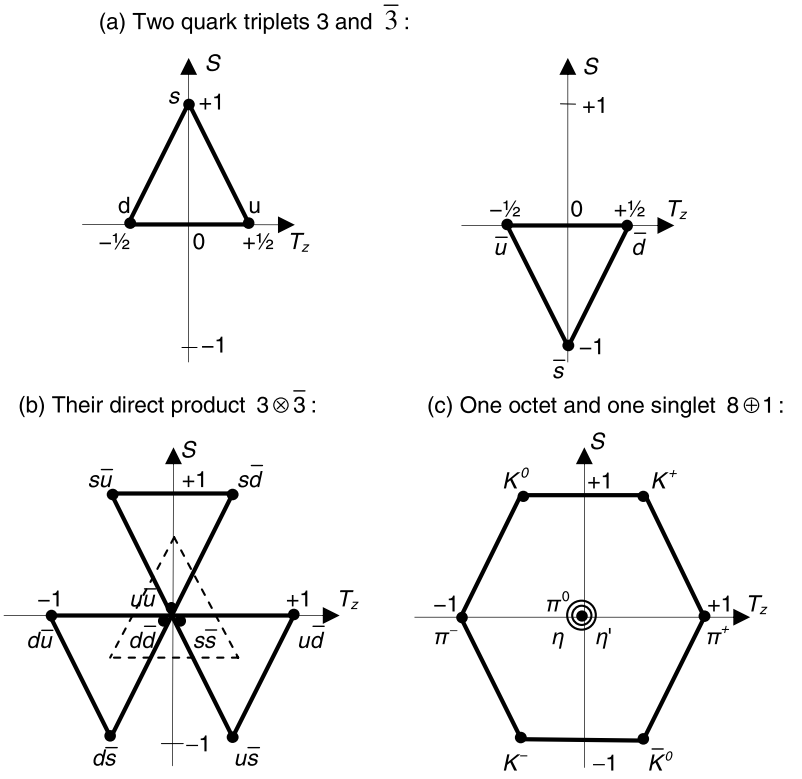


Fig. 14.2 Graphical representation of the coupling of three quarks and three antiquarks to mesons: (**a**) The quarks u, d, and s and their antiquarks are symbolized by three points (T_z, S) in the plane isospin vs. strangeness, which indicate their quantum numbers. (**b**) The nine direct-product quark–antiquark states shown find their place in the T_z–S plane by simply adding their single-quark quantum numbers. (**c**) These direct-product states then are transformed into the properly antisymmetrized singlet and octet meson states, as indicated by the *dots*. *Circles around a dot* signal multiple occupations

Today, the three-family standard model has altogether six different quarks, accommodated in three quark doublets with flavors down d and up u, strange s and charmed c, bottom b and top t. While the neutron–proton mass difference is very small, particles with flavors beyond s show much larger mass differences. Therefore, for quarks beyond the strange quark, flavor symmetry is no longer very helpful.

When these particle multiplets were established, for internal consistency another three-valued quantum number was needed, arbitrarily called *color*, with quantum numbers red, blue, and green. It later turned out that the strong coupling constant of *quantum chromodynamics* or QCD, the theory of the strong nuclear interaction, is based on color:

Color is the three-valued charge that is the source of the strong interaction.

This is the analogue to electric charge $\pm e$, which is the two-valued source of the electromagnetic interaction. The two states of electric charge do not lead to pseudospin behavior because electromagnetic interactions are sensitive to the sign of the charge, whence both charge states are distinguishable. In contrast, the strong interaction, whose source is the color charge, is found to be blind against the individual colors red, blue, or green. Of all the possible multiplets of color-$SU(3)$, only "white" color singlets $3 \otimes 3 \otimes 3 = 1$ are realized in nature, and it remains uncertain which of the three colors is acting in a given process. In contrast to flavor symmetry, color-charge symmetry in QCD is a perfect symmetry.

14.3 Particle Oscillations

Some pairs of particles can periodically transform into each other, a process called *particle oscillations*. Examples are kaon–antikaon oscillations, neutrino oscillations between different neutrino species, and possibly neutron–antineutron oscillations. We give a survey of the quantum description of these processes, but do not have space for going into the experimental details or into their deep theoretical implications.

14.3.1 Kaon Oscillations

Kaon oscillations were discovered about half a century ago. In these K^0–$\bar{K}^0$ oscillations, a neutral kaon with quark content $d\bar{s}$ periodically changes into an antikaon $\bar{d}s$, and back into a kaon, etc., with an oscillation period of 0.2 ns. Particle–antiparticle oscillations are possible only for electrically neutral particles such as to obey charge conservation.

Neutral kaons K^0 have strangeness quantum number $S = 1$, while their antiparticles $\bar{K}^0$ have $S = -1$, see Fig. 14.2(c). Kaon oscillations therefore are strangeness oscillations with $|\Delta S| = 2$. Strangeness is found to be conserved under the strong interaction, therefore strangeness changing kaon oscillations must be attributed to the weak interaction. Both strange K^0 and $\bar{K}^0$ particles decay weakly into the same zero-strangeness two-pion final states, either to $\pi^+ + \pi^-$ or to $2\pi^0$, and via their common final states are virtually coupled to each other.

The coupled K^0–$\bar{K}^0$ system has the 2×2 mass matrix

$$\mathrm{M} = m_K \mathrm{I} + \delta_K \sigma_x , \tag{14.3a}$$

with $m_K = m_{\bar{K}} = 498\ \mathrm{MeV}/c^2$, and a coupling term δ_K. Basis states are the mass eigenstates $|K^0\rangle$ and $|\bar{K}^0\rangle$. Diagonalization of this matrix leads to $\mathrm{M_D} =$

$m_K \mathrm{I} + \delta_K \sigma_z$, with a mass splitting $m_{1,2} = m_K \pm \delta_K$, similar as in (12.12), and an experimentally determined gap of $2\delta_K = 3.5\ \mu\mathrm{eV}/c^2$, or $\delta_K/m_K \sim 10^{-15}$. The corresponding weak-interaction eigenstates $|K_\mathrm{S}\rangle$ and $|K_\mathrm{L}\rangle$ are related to these mass eigenstates as

$$|K_\mathrm{S}\rangle = \sqrt{\frac{1}{2}}\big(|K^0\rangle + |\bar{K}^0\rangle\big),\ |K_\mathrm{L}\rangle = \sqrt{\frac{1}{2}}\big(|K^0\rangle - |\bar{K}^0\rangle\big), \tag{14.4}$$

cf. (3.54). K_S is symmetric under exchange of K^0 and $\bar{K}^0$, while K_L is antisymmetric. The subscripts S and L mean short and long lifetimes, respectively, for reasons explained in the following.

Let the "up state" K^0 be produced at time $t = 0$ by a strong interaction reaction like $\pi^- + p^+ \to K^0 + \Lambda$. The strong-interaction eigenstates have no off-diagonal couplings as have the weak eigenstates, but are identical to the mass eigenstates. We know from (3.58b) that the probability of finding the system in the "down state" $\bar{K}^0$, with $p_{K^0\to\bar{K}^0} \equiv p_\mathrm{down}$, evolves from the initial mass eigenstate as

$$p_{K^0\to\bar{K}^0} = \frac{1}{2}(1 - \cos\omega_K t), \tag{14.5}$$

with the kaon oscillation angular frequency $\omega_K = (2c^2/\hbar)\delta_K$.

In K_S and K_L decay, the final two-pion states $\pi^+ + \pi^-$ or $2\pi^0$ evidently are symmetric under particle exchange. Therefore, for the equally symmetric K_S these decays are allowed with a lifetime $\tau_\mathrm{S} = 90$ ps. For the antisymmetric K_L, these decays are forbidden, only decay into three pions can produce an antisymmetric final state. The decay probability, i.e., the phase space available for three-pion decay is much smaller than for two-pion decay, hence the lifetime $\tau_\mathrm{L} = 51$ ns of the K_L is much longer.

To take the finite lifetimes of K_S and K_L into account, we subtract a diagonal decay matrix Γ from the mass matrix $\mathrm{M_D}$. This matrix Γ has the two decay rates τ_L^{-1} and τ_S^{-1} on its diagonal. The eigenfunctions ψ_D of the diagonal matrix $\mathrm{M} = \mathrm{M_D} - (\mathrm{i}\hbar/c^2)\Gamma$ evolve as $\dot{\psi}_\mathrm{D} = -(\mathrm{i}/\hbar)\mathrm{M}\psi_\mathrm{D}$, whose solution has the usual coefficients

$$c_\mathrm{S}(t) = c_\mathrm{S}(0)\mathrm{e}^{\mathrm{i}(m_K+\delta_K)c^2 t/\hbar}\mathrm{e}^{-t/\tau_\mathrm{S}}, \tag{14.6a}$$

$$c_\mathrm{L}(t) = c_\mathrm{L}(0)\mathrm{e}^{\mathrm{i}(m_K-\delta_K)c^2 t/\hbar}\mathrm{e}^{-t/\tau_\mathrm{L}}. \tag{14.6b}$$

From this we form the downstate $b(t) = [c_\mathrm{S}(t) - c_\mathrm{L}(t)]/\sqrt{2}$ and obtain the oscillation probability $p_\mathrm{down} = |b|^2$, or

$$\begin{aligned} p_{K^0\to\bar{K}^0} &= \sqrt{\frac{1}{2}}\big[c_\mathrm{S}(t) - c_\mathrm{L}(t)\big]\sqrt{\frac{1}{2}}\big[c_\mathrm{S}^*(t) - c_\mathrm{L}^*(t)\big] \\ &= \frac{1}{2}\mathrm{e}^{-2t/\tau_\mathrm{S}} + \frac{1}{2}\mathrm{e}^{-2t/\tau_\mathrm{L}} - \mathrm{e}^{-(1/\tau_\mathrm{S}+1/\tau_\mathrm{L})t}\cos\omega_K t. \end{aligned} \tag{14.7}$$

Over the years, these damped kaon oscillations have been studied in much detail, and, in addition to the amplitudes in (14.6a), (14.6b), a small time reversal violating

amplitude was found which, within the standard model, can be elegantly absorbed in the *CKM quark mixing matrix* that relates the weak-interaction eigenstates of the quarks to their mass eigenstates via a rotation in the flavor space of the 3-family standard model. These topics are discussed in many textbooks on particle physics, and we shall not repeat the details here. For classical analogues of kaon and other meson oscillation effects, see Schubert and Stiewe (2011).

14.3.2 Neutrino Oscillations

A deficit in the number of solar neutrinos arriving on earth had been seen since the 1970s, and was tentatively attributed to neutrino disappearance into other neutrino flavors. Other neutrino ν_e–ν_μ and ν_e–ν_τ oscillations were observed in early 2000 and studied in great detail since, with neutrino flight lengths ranging from tens of meters up to the distance to the sun of 1.5×10^8 km.

In the three-family standard model, besides the three quark doublets mentioned, there are also three *lepton* doublets, consisting of the electron e and the electron neutrino ν_e, the muon μ and the muon neutrino ν_μ, and the tau τ and the tau neutrino ν_τ, together with their respective antiparticles. The results from different neutrino oscillation experiments show that at least two of the three neutrino flavors ν_e, ν_μ, ν_τ must have nonzero masses m_1, m_2. In the standard model, however, all neutrinos must be massless. Therefore, neutrino oscillations must be a clear signal for new physics beyond the standard model. Again, we only derive the general phenomenology of neutrino oscillations, for the simplified case that only the two neutrino flavors ν_e and ν_μ are involved.

We could, as we did in the case of kaon oscillations, write down a mass matrix for the coupled ν_e–ν_μ system, with nonzero off-diagonal elements. Nowadays, however, an equivalent ansatz is preferred, which leads to the same results, but is based on neutrino mixing. It assumes that the weak-interaction eigenstates ψ' of the neutrinos are not identical to their mass or energy eigenstates ψ, but are related to them by a neutrino mixing matrix V, with rotation angle θ in lepton-flavor space, as

$$\psi' = \mathrm{V}\psi, \tag{14.8a}$$

$$\begin{pmatrix} \nu_\mu \\ \nu_e \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix}, \tag{14.8b}$$

with mass eigenstates ν_1 and ν_2. This is in analogy to CKM quark mixing, with the above matrix equations being the analogue of the original 2-family version of quark mixing, here applied to lepton mixing.

A complete derivation of the time evolution of relativistic neutrino wave packets would require inclusion of the production, propagation, and detection of neutrinos. Our simplified textbook version happens to come to the same results. We assume

very small neutrino masses, $m_i c^2 \ll E_i$, and write the neutrino relativistic energies as

$$E_i = \left(p_i^2 c^2 + m_i^2 c^4\right)^{1/2} \approx p_i c + \frac{(m_i c^2)^2}{2 p_i c}, \quad i = 1, 2, \tag{14.9}$$

where we used $(1+\varepsilon_i)^{1/2} \approx 1 + \frac{1}{2}\varepsilon_i$ with $\varepsilon \equiv (m_i c^2/p_i c)^2 \ll 1$.

Let the neutrinos be born as electron neutrinos, i.e., in a flavor eigenstate, in a weak interaction process, for instance, by the nuclear fusion reaction $p + p \to d + e^+ + \nu_e$ in the sun, and let them be detected in another weak interaction reaction, for instance, by the inverse reaction of deuteron breakup $d + \nu_e \to p + p + e^-$ in a neutrino detector on earth. Between creation and detection, the neutrinos propagate as mass eigenstates ν_1 and ν_2. To find the time evolution of the ν_e, we use $p_{\nu_e \to \nu_e} \equiv p_{\text{up}} = 1 - \sin^2 2\theta \sin^2 \frac{1}{2}\omega_{\text{L}} t$ from (5.24), with 2θ instead of θ because we use V as given in the matrix in (14.8b) and not the matrix T from (5.12). We then replace ω_{L} by $(E_2 - E_1)/\hbar$ from (14.9), write $\Delta m^2 \equiv m_2^2 - m_1^2$, set $p_1 \approx p_2 \approx E/c$ with mean neutrino beam energy E, and $L = ct$ for the distance between source and detector. This gives the probability that a ν_e arrives as such at the detector

$$p_{\nu_e \to \nu_e} \approx 1 - \sin^2 2\theta \sin^2 \frac{\Delta m^2 c^4}{4 E \hbar c} L. \tag{14.10}$$

Insertion of the *oscillation length* $\lambda = 4\hbar c E/|\Delta m^2| c^4$ gives the oscillation rates

$$p_{\nu_e \to \nu_e} = 1 - \sin^2 2\theta \sin^2(L/\lambda) \quad (\textit{survival}), \tag{14.11}$$

$$p_{\nu_e \to \nu_\mu} = 1 - p_{\nu_e \to \nu_e} = \sin^2 2\theta \sin^2(L/\lambda) \quad (\textit{disappearance}). \tag{14.12}$$

If E is measured in MeV, Δm in eV/c^2, then λ given in m is $\lambda = 0.79 \times E/|\Delta m^2|$.

From the observed ν_e–ν_μ oscillations, one finds $\Delta m_{21}^2 = 7.6(3) \times 10^{-5}$ eV^2 and $\sin^2\theta_{12} = 0.31(2)$. From ν_e–ν_τ oscillations, one finds $|\Delta m_{32}^2| = 2.4(1) \times 10^{-3}$ eV^2 and $\sin^2\theta_{23} = 0.42(1)$. At present ν_e–ν_τ oscillations one finds $\sin^2\theta_{13} = 0.025(3)$ with $|\Delta m_{31}^2| \approx |\Delta m_{32}^2|$. From the existing data, it follows that all three types of neutrinos are light with masses $m_i < 0.5$ eV and $m_1 \neq m_2 \neq m_3$, and there is no experimental evidence for the existence of more than three light neutrinos, see Nakamura and Petcov (2012).

14.3.3 Neutron Oscillations

Could it be that neutrons, too, oscillate into antineutrons? Before and after a nuclear reaction involving baryons, the number of baryons minus the number of antibaryons is always found to be the same. One says that *baryon number B* is conserved. A similar conservation law was thought to hold for the lepton numbers L_e, L_μ, and L_τ, which were found to be separately conserved before the discovery of neutrino oscillations. An example for the conservation of baryon and lepton number is neutron decay $n \to p^+ + e^- + \bar{\nu}_e$, where $B = 1$ and $L_e = 0$ are the same before and after decay.

There is, except for the evident stability of the universe, no theoretical compelling reason for a conservation law for baryon or lepton number. To the contrary, should there exist a grand unified state of matter beyond the standard model (i.e., at very high energies, realized either in the laboratory or in the very early universe), then baryons and leptons would be members of one big particle multiplet. Within this multiplet, baryons and leptons would be indistinguishable from each other, and therefore should be able to transform into each other. Baryon and lepton number violating processes like proton decay $p^+ \to e^+ + \pi^0$ should therefore be allowed, but so far have not been observed, with a lower proton lifetime limit of 8×10^{31} years at 90 % C.L. This reaction has $B = 1$ and $L = 0$ before decay, and $B = 0$, $L = -1$ after decay, the pion as a quark–antiquark pair having $B = 0$.

In various models beyond the standard model, neutron–antineutron oscillations $n \leftrightarrow \bar{n}$ should be possible as a higher order process, with baryon number oscillating by two units $|\Delta B| = 2$. Hence, in a free neutron beam it may be that some neutrons are converted in-flight into antineutrons. Let us calculate the time evolution of this process. A magnetic field B_0 lifts the degeneracy between neutrons and antineutrons, necessary for their oscillations, because their magnetic moments have opposite signs, $\mu_{\bar{n}} = -\mu_n$; therefore, we include the magnetic interaction from the beginning. The energy matrix for neutron and antineutron basis states n and $\bar{\text{n}}$ of equal mass $m_n = m_{\bar{n}}$ is

$$\mathrm{H} = m_n c^2 \mathrm{I} + \mu_n B_0 \sigma_z + \delta_n \sigma_x, \tag{14.13}$$

with the off-diagonal amplitude $\delta_n = \hbar/\tau_{n\bar{n}}$, and neutron–antineutron oscillation time $\tau_{n\bar{n}}$. The probability of the appearance of an antineutron in a beam of neutrons varies, from (5.24), with the neutron time of flight t as

$$p_{n\to\bar{n}}(t) = \frac{\delta_n^2}{\delta_n^2 + (\mu_n B_0)^2} \sin^2\left(\frac{1}{\hbar}\sqrt{\delta_n^2 + (\mu_n B_0)^2}\, t\right). \tag{14.14}$$

Under all realistic conditions, the interaction $\mu_n B_0$ with residual magnetic fields is much larger than the present limits on δ_n, and therefore the multiplier on the right of (14.14) becomes $\approx (\delta_n/\mu_n B_0)^2$, which is extremely small, and which at first sight seems to strongly suppress $n\bar{n}$ oscillations. However, as long as the time of observation is short enough, we can safely replace the sine function in (14.14) by its argument, which then cancels with the denominator. In this case, the neutrons are *quasifree* because observation time is too short to resolve the magnetic energy splitting within energy–time uncertainty, and $p_{n\to\bar{n}}(t) \approx \delta_n^2 (t/\hbar)^2$, or

$$p_{\bar{n}}(t) \approx t^2/\tau_{n\bar{n}}^2. \tag{14.15}$$

This is in agreement with the observation in Sect. 5.3.2 that for short times t any quantum system evolves proportional to t^2.

In the early 1990s, a search for neutron–antineutron oscillations was made (Baldo-Ceolin et al. 1994), with neutrons flying over a distance of 75 m in almost field free space with a mean time of flight $t \approx 0.1$ s. A lower limit $\tau_{n\bar{n}} > 0.86 \times 10^8$ s ≈ 3 y at 90 % C.L. was obtained for the oscillation time.

References

Baldo-Ceolin, M., et al.: A new experimental limit on neutron–antineutron oscillations. Z. Phys. C **63**, 409–416 (1994)

Nakamura, K., Petcov, S.T.: Neutrino mass, mixing, and oscillations. In: J. Beringer et al. (PDG), Phys. Rev. D **86**, 010001 (2012). http://pdg.lbl.gov

Schubert, K.R., Stiewe, J.: Demonstration of K_0–$\bar{K}_0$, B_0–$\bar{B}_0$, and D_0–$\bar{D}_0$ transitions with a pair of coupled pendula (2011). arXiv:1108.2772 [hep-ph]

Chapter 15
Quantum Informatics

Abstract The field of quantum informatics is in rapid evolution. It has three main objectives: the simulation of physical quantum systems that are inaccessible to conventional computing; quantum computing of hard mathematical problems; long-distance quantum communication. We list basic requirements that quantum hardware has to meet, discuss the difficulties encountered, and give a snapshot on experimental approaches to quantum computing and communication in atomic, solid state, and photonic systems.

15.1 Quantum Information Theory

Over the past few decades, a new field of *quantum informatics* has emerged, which uses entangled quantum states or *qubits* as carriers of information. A qubit is any state vector $\psi = a\alpha + b\beta$ of a two-state quantum system of pseudospin $s = \frac{1}{2}$. In quantum informatics, the basis states α and β are called $|1\rangle$ and $|0\rangle$, in reminiscence of classical information theory with its two discrete bits 1 for "true" and 0 for "false". A number N of such qubits, allowed to interact with each other in some predefined way, is represented by a state vector Ψ in 2^N-dimensional Hilbert space.

Each qubit ψ_i, $i = 1, \ldots, N$, can be prepared to be initially either in its upstate α_i, or in its downstate β_i, or in a superposition of both. In this way, at time $t = 0$ an input state $\Psi(0)$ of N qubits is prepared. This state is then manipulated by a sequence of quantum operations under which it evolves as $\Psi(t) = U(t)\Psi(0)$, with some time evolution operator $U(t)$. At readout time $t = T$, the output state is measured, and the N qubits are found in a certain sequence of up and down states, with a probability distribution $p_i(T) = |a_i(T)|^2 = 1 - |b_i(T)|^2$ for the expectation values of the pseudospins. Hence, like ordinary spin rotation experiments, quantum computing is separable into three successive steps:

Step 1: Initialization $\Psi(0)$ of the system;
Step 2: Evolution of the system as described by a unitary operator $\mathrm{U}(t)$;
Step 3: Final readout of $\Psi(T)$ of the system.

So far, quantum computing does not differ from the ordinary evolution of a quantum system, as described in preceding chapters. How then can one use such a coupled quantum system as a quantum computer? As was pointed out by Feynmann

D. Dubbers, H.-J. Stöckmann, *Quantum Physics: The Bottom-Up Approach*, Graduate Texts in Physics,
DOI 10.1007/978-3-642-31060-7_15,

(1982), classical digital computers are not very efficient in simulating complex quantum systems, and he proposed using quantum computers instead. In particular, Feynmann (1986) showed that sequences of spin raising and lowering operators of the type given in (3.6), also called *creation* and *annihilation operators*, could be used to represent *quantum logic gates*. A NOT gate, for instance, turns an input qubit $|0\rangle$ into an output qubit $|1\rangle$, and vice versa. An AND gate produces output $|1\rangle$ if its two input qubits are alike, and output $|0\rangle$ if not. This is nearly all one needs to build a computer. It was then shown (Deutsch 1985) that quantum entanglement would allow a *universal quantum computer* to perform certain tasks with unprecedented speed.

In a quantum computer, each qubit carries a complex vector space with it that allows it to be not only in states $|0\rangle$ or $|1\rangle$, but in all superposition states.

It was far from evident that one could make an entangled quantum system perform useful mathematical computations. Some of the problems encountered in quantum computing are the following:

Uniqueness: A classical computer delivers definite results, while a quantum computer at readout time T merely produces a probabilistic result, namely, a spectrum of possible outcomes with weights $p_i, i = 1, \ldots, N$.

Stability: Classical computing is an inherently stable digital process working with discrete bits of 0 or 1. Quantum computing is an analog process working with a continuum of entangled and extremely fragile quantum states. For a larger number N of qubits, the information content of the system almost entirely resides in the nonlocal "off-diagonal" correlations between input and output states. These states are very sensitive not only to imperfections in the quantum gate operations, but also to *decoherence* due to an uncontrollable entanglement with the environment, see also Sect. 22.4. In conventional computers, reliability is enhanced by cooling the system via dissipative coupling to a heat bath, thereby suppressing thermal activation of erroneous bits. In a quantum computer, in contrast, coupling to a heat bath enhances decoherence and reduces reliability.

Correctability: In a classical computer, error correction codes are relatively simple and reliable. The states $|0\rangle$ and $|1\rangle$ usually are represented by two voltage levels, and small deviations from their setpoints can be measured and reset to zero essentially after each computational step. In addition, accidental bit flips can be detected and repaired with arbitrarily high reliability if a sufficient number of control bits are added. In a quantum computer, the uncertainty principle threatens that just looking at the system leads to uncontrolled disturbances. How can errors be detected and repaired if one seems not even be allowed to read the file? Under these conditions, even to know when readout time T has come may be a problem.

In view of these problems, for some time the opinion prevailed that quantum computing has a certain intellectual appeal but is of little practical relevance. Still,

it had to be acknowledged that quantum informatics must be studied if only to investigate problems that lie ahead when going from micrometer to nanometer sized logic gates where quantum effects may dominate.

15.2 Quantum Computing and Quantum Communication

A number of recent discoveries tackled many of the problems listed above, and shifted the rating of quantum computing from "impossible" to "difficult, though not hopeless".

Uniqueness: Important possible applications of quantum computing were found for which the spectrum of possible results $p_i(T)$ is sufficiently narrow and can rapidly be checked. One of the hardest problems in mathematics, the factoring of large numbers $n = pq$, with integer p and q, can be done exponentially faster, with computing time $T \propto n^2$ as compared to $T \propto \mathrm{e}^{n/3}$ on a classical computer, see references in Chuang et al. (1995). In another possible application, the time needed for searches in databases with n entries scales as $T \propto n^{1/2}$ on a quantum computer, as compared to $T \propto n$ on a classical computer, and hence needs only half the number of digits to do the job. The list of useful mathematical applications is still rather short. On the other hand, for the simulation of physical quantum systems, uniqueness is not a problem because the spectrum of possible outcomes $p_i(T)$ is the aim of the study.

Stability: Some progress has been achieved by choosing processes that do not strongly depend on the precise trajectory of state vector $\Psi(t)$, such as purely adiabatic or purely nonadiabatic transitions, or adiabatic excursions with stable Berry phases. If this is not enough, one has to rely on error correction codes, our next point.

Correctability: Correction codes were found that make use of a peculiar quantum feature: measurements in one subspace of a high-dimensional Hilbert space can give information on another disjoint subspace without disturbing the latter. However, such error correction schemes strongly inflate the number of qubits needed.

> We give an example of an error-correction scheme. For each qubit, instead of the basis states $|0\rangle$ and $|1\rangle$ we take the three qubits of a GHZ states $|0\rangle|0\rangle|0\rangle \equiv |000\rangle$ and $|1\rangle|1\rangle|1\rangle \equiv |111\rangle$ (see the end of Sect. 9.2), and in place of $\psi = a|0\rangle + b|1\rangle$ we encode a superposition $\psi' = a|000\rangle + b|111\rangle$. A measurement on one individual substate of a GHZ state makes this wave function collapse either to $|000\rangle$ if the substate is found to be $|0\rangle$, or to $|111\rangle$ if it is found to be $|1\rangle$. If one of the three qubits has suffered an accidental single-bit flip, how can we identify the culprit? Let s be the binary one-digit sum of the simultaneously measured first and third GHZ qubits. If both are the same, then $s = 0 + 0 = 1 + 1 = 0$, see the first and third rows of the list below. If both are not the same then $s = 1$, see the second and fourth rows. Let s' be the binary one-digit sum of the second and the third qubits. The pair (s, s'), when read as a two-digit binary number, then tells us the position $i = s + 2s'$ of a single bit flip. We check this by listing below the numbers s, s', and i for the three possible single-bit flip errors in an initially error-free state ψ', where the last column instructs us

which bit has to be corrected ($i = 0$ for "no bit flip").

Possible single-bit flips	Binary pair (s, s')	Error position $i = s + 2s'$
$a\|000\rangle + b\|111\rangle$	(0, 0)	0
$a\|100\rangle + b\|011\rangle$	(1, 0)	1
$a\|010\rangle + b\|101\rangle$	(0, 1)	2
$a\|001\rangle + b\|110\rangle$	(1, 1)	3

We next discuss technical realizations of quantum information processing. Many of the two-state systems and procedures discussed in earlier chapters have been tested for their utility as quantum hardware. To give a snapshot of the situation at the time of writing, we cite a number of recent references. For specific applications of Rabi, Ramsey, spin echo, double quantum, dressed atom, or magnetic resonance imaging techniques, see Grinolds et al. (2011), Neumann et al. (2010), and Timoney et al. (2011). For error-avoiding schemes based on adiabatic or nonadiabatic transitions and for geometric excursions, see the review by Nayak et al. (2008). Below we list four prominent realizations of quantum hardware. The first two realizations use qubits in free space, while the other two use qubits in condensed matter. For a comprehensive overview see Ladd et al. (2010).

Ions in an electrostatic trap: A qubit here is realized by two states of an ion's electronic shell. A system of at present up to $N = 8$ such ions is stored in a trap. It is initialized as a polarized state via optical pumping. Individual ions are addressed by sequences of laser pulses. The qubits are coupled to each other by their collective quantized motion in the trap's harmonic potential. The system is read out via vibrational-state dependent fluorescence. The ions can be stored on a chip, with typical nearest-neighbor distances of 1 μm. Advantages of ionic qubits are near 100 % fidelity of initialization and readout, and long coherence times T_1, $T_2 \geq 1$ s; for a survey, see Häffner et al. (2008).

Polarized photons: A qubit is realized by the two states of photon polarization. Any $N \times N$ unitary operation U can be realized in the laboratory by $\frac{1}{2}N(N-1)$ beam splitters plus a number of phase shifters, and the eigenvalues of any hermitian matrix H can then be measured with a number of N photon detectors (Reck et al. 1994). Photonic quantum devices are fast and have coherence times of typically 0.1 ms. Assembling more than six photonic qubits before seemed to be an insurmountable challenge with present technology. However, entanglement between polarization and spatial degrees of freedom of single photons recently allowed the production of a ten-qubit GHZ state. The main obstacles to further progress are uncontrolled photon losses; for a survey, see Kok et al. (2007).

Dopants in condensed matter: Nitrogen-vacancy centers is a naturally occurring lattice defect in diamond, consisting of a substitutional nitrogen atom and an adjacent carbon vacancy. Such centers, artificially created in ultrapure synthetic diamond, are a favorite system for studies on qubits in the solid state. The lattice, isotopically enriched to 99.99 % of ^{12}C (nuclear spin $I = 0$), is essentially "spin free", with long relaxation times even at room temperature. Both nuclear and atomic spin polarization of the dopant can be used for quantum operations. The position of such a color center can be made visible either by optically detected tomographic electron-magnetic resonance imaging (MRI; Grinolds et al. 2011), or by stimulated

emission microscopy (STED; Neumann et al. 2010). As another condensed-matter example, NMR on liquid ^{13}C-labeled ($I = \frac{1}{2}$) chloroform was used as a quantum simulator, on which the binding energy of the H_2-molecule was calculated to -1.851570929351124 a.u., with 45 bit precision, in agreement with the binding energy calculated on a conventional computer to -1.85157092935119 a.u. (Du et al. 2010).

Superconducting circuits: Here macroscopic qubits are realized by superconducting ring currents, with spin up corresponding to a clockwise, spin down to an anticlockwise circulating current. Neighboring qubits are coupled to each other by programmable superconducting elements such as Josephson junctions (see Sect. 13.2.2), which make the "spin–spin" interaction between neighboring qubits continuously tunable in the range from a parallel ("ferromagnetic") to an antiparallel coupling ("antiferromagnetic"). The state of a qubit can be measured with a SQUID; for a review, see Clarke and Wilhelm (2008). A programmable array of eight superconducting qubits has successfully attacked the classically intractable problem of *quantum annealing*, i.e., of finding the ground state that a system of interacting spins will settle to when temperature is slowly lowered. These commercially available devices signal a remarkable progress in quantum simulation (Johnson et al. 2011).

Finally, we shortly touch on quantum communication. Entangled photons can transmit quantum information over large distances; see, e.g., Ursin et al. (2007). Quantum communication devices are commercially available and offer unbreakable communication channels that are protected by the laws of physics: the uncertainty principle lets the sender and receiver know whether an eavesdropper has read the message because any measurement on the transmitted quantum information will inevitably alter its content. In quantum communication, *teleportation* is the successful reconstruction of a quantum state by the receiver. Teleportation is based on purely classically transmitted information, while sender and receiver both are coupled to a common source of entangled qubits. In principle, the *no-clone theorem* does not allow making a perfect copy of a quantum state. If it were otherwise, one could do measurements on the copy without disturbing the original, and in this way bypass the uncertainty principle. Teleportation, though, does not violate the no-clone theorem because nothing forbids the preparation of a perfect copy at some distant place if the original is destroyed in the process. For a review on quantum communication, see Yuan et al. (2010).

References

Chuang, I.L., Laflamme, R., Shor, P.W., Zurek, W.H.: Quantum computers, factoring and decoherence. Science **270**, 1633–1635 (1995)

Clarke, J., Wilhelm, F.K.: Superconducting quantum bits. Nature **453**, 1031–1042 (2008)

Deutsch, D.: Quantum theory, the Church–Turing principle and the universal quantum computer. Proc. R. Soc. Lond. A **400**, 97–117 (1985)

Du, J., Xu, N., Peng, X., Wang, P., Wu, S., Lu, D.: NMR implementation of a molecular hydrogen quantum simulation with adiabatic state preparation. Phys. Rev. Lett. **104**, 030502(3) (2010)

Feynmann, R.P.: Simulating physics with computers. Int. J. Theor. Phys. **21**, 467–488 (1982)
Feynmann, R.P.: Quantum mechanical computers. Found. Phys. **16**, 507–531 (1986)
Grinolds, M.S., Maletinsky, P., Hong, S., Lukin, M.D., Walsworth, R.L., Yacoby, A.: Quantum control of proximal spins using nanoscale magnetic resonance imaging (2011). arXiv:1103.0546v1
Häffner, H., Roos, C.F., Blatt, R.: Quantum computing with trapped ions. Phys. Rep. **469**, 155–203 (2008)
Johnson, M.W., et al.: Quantum annealing with manufactured spins. Nature **473**, 194–198 (2011)
Kok, P., Munro, W.J., Nemoto, K., Ralph, T.C., Dowling, J.P., Milburn, G.J.: Linear optical quantum computing with photonic qubits. Rev. Mod. Phys. **79**, 135–174 (2007)
Ladd, T.D., Jelezko, F., Laflamme, R., Nakamura, Y., Monroe, C., O'Brien, J.L.: Quantum computers. Nature **464**, 45–53 (2010)
Nayak, C., Simon, S.H., Stern, A., Freedman, M., Sarma, S.D.: Non-Abelian anyons and topological quantum computation. Rev. Mod. Phys. **80**, 1083–1159 (2008)
Neumann, P., et al.: Quantum register based on coupled electron spins in a room-temperature solid. Nat. Phys. **6**, 249–253 (2010)
Reck, M., Zeilinger, A., Bernstein, H.J., Bertani, P.: Experimental realization of any discrete unitary operator. Phys. Rev. Lett. **73**, 58–61 (1994)
Timoney, N., Baumgart, I., Johanning, M., Varón, A.F., Plenio, M.B., Retzker, A., Wunderlich, Ch.: Quantum gates and memory using microwave-dressed states. Nature **476**, 185–189 (2011)
Ursin, R., et al.: Entanglement-based quantum communication over 144 km. Nat. Phys. **3**, 481–486 (2007)
Yuan, Z.-S., Bao, X.-H., Lu, C.-Y., Zhang, J., Peng, C.-Z., Pan, J.-W.: Entangled photons and quantum communication. Phys. Rep. **497**, 1–40 (2010)

Part IV
Multilevel Systems and Tensor Operators

In the preceding chapters, we worked mainly with 2×2 matrices, and treated various two-state quantum systems as $s = \frac{1}{2}$ effective-spin entities, subject to a dipole interaction of arbitrary origin. In the following chapters, we restart our treatise on a higher level and generalize our findings to spins $s > \frac{1}{2}$ and to interactions beyond the magnetic dipole. We give a compact description of quantum processes in terms of what is called the irreducible tensor representation. Written in terms of irreducible tensors, the Liouville equation of motion for the observables allows for a simple interpretation of the underlying processes, providing us with a generalization of the magnetic spin-precession equation to higher multipole interactions.

This page intentionally left blank

Chapter 16
Rotations and Angular Momentum

Abstract In order to generalize the spin-$\frac{1}{2}$ results, we develop the necessary tools in the sections to follow. The rotation of classical objects is found already to possess the ingredients needed for a quantum description of rotation. We continue with the quantum properties of orbital angular momentum and its eigenstates, the spherical harmonics, derive the operators that rotate quantum states in Hilbert space, and establish the corresponding rotation matrices.

16.1 Symmetries

Symmetry is one of the most successful concepts in physics, the importance of which can be hardly overestimated. All known fundamental equations governing physics are invariant with respect to shifts of time, of position, and to rotations. Immediate consequences of these symmetries are the conservation laws for energy, momentum, and angular momentum. In addition to these symmetries that depend on the continuous variables time, position and angle, there are also discrete symmetries, for example, with respect to a reflection at some origin or at a mirror plane, with a corresponding conserved quantity named parity. The method of choice to cope with symmetries is group theory. We shall not make explicit use of it here, but readers familiar with group theory will recognize concepts and terminologies borrowed from it.

In the present context, symmetries with respect to rotations are of particular importance. Let us take for illustrative purposes a system we all know very well, our earth. For the description, we introduce an external Cartesian coordinate system x, y, z fixed in time, and an internal coordinate system x', y', z' fixed to the earth. Let us assume that at an arbitrary time zero, e.g., 12:00 CET, both coordinate systems coincide. Six hours later the earth has rotated by an angle of $\pi/2$. What does this mean for the two coordinate systems? Let us take the rotation axis, aligned to the north, as the z-axis. The relation between the two coordinate systems is then given by

$$x' = x\cos\alpha + y\sin\alpha, \tag{16.1a}$$

D. Dubbers, H.-J. Stöckmann, *Quantum Physics: The Bottom-Up Approach*, Graduate Texts in Physics,
DOI 10.1007/978-3-642-31060-7_16, © Springer-Verlag Berlin Heidelberg 2013

$$y' = -x\sin\alpha + y\cos\alpha, \tag{16.1b}$$

$$z' = z, \tag{16.1c}$$

where α is the rotation angle, assuming counter-clockwise rotation.

Let us now consider an arbitrary classical quantity characterizing the earth, e.g., its mass density $\rho(x, y, z)$. This density, expressed in terms of the external coordinates, is

$$\rho\big(x', y', z'\big) = \rho(x, y, z, \alpha) = \rho(x\cos\alpha + y\sin\alpha, -x\sin\alpha + y\cos\alpha, z). \tag{16.2}$$

For the example given above, α takes the value of $\pi/2$. Taking now an infinitesimally small angle ε and expanding up to terms linear in ε, this reads

$$\rho(x, y, z, \varepsilon) = \rho(x, y, z) + \varepsilon\left(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}\right)\rho(x, y, z). \tag{16.3}$$

In terms of the previously introduced operator for the momentum

$$\mathbf{p} = \frac{\hbar}{\mathrm{i}}\nabla, \tag{16.4}$$

the last equation can be more concisely written as

$$\rho(x, y, z, \varepsilon) = \left(1 - \frac{\mathrm{i}}{\hbar}\varepsilon \mathrm{L}_z\right)\rho(x, y, z). \tag{16.5}$$

Herein L_z is the z-component of the angular momentum operator

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} = \frac{\hbar}{\mathrm{i}}\mathbf{r} \times \nabla. \tag{16.6}$$

Passing to the limit as $\varepsilon \to 0$ in (16.5), we obtain the differential equation

$$\dot{\rho} = -\frac{\mathrm{i}}{\hbar}\mathrm{L}_z\rho. \tag{16.7}$$

The dot means differentiation with respect to rotation angle α. The equation is immediately solved with the result

$$\rho\big(x', y', z'\big) = \rho(x, y, z, \alpha) = \mathrm{e}^{-\mathrm{i}\alpha \mathrm{L}_z/\hbar}\rho(x, y, z). \tag{16.8}$$

Quite unexpectedly we stumbled across concepts of quantum mechanics, $\hbar$ and a momentum operator $\mathbf{p}$ instead of a variable, notwithstanding the fact that the whole derivation was purely classical!

The derivation shows that the operators for momentum and angular momentum appear naturally already in classical mechanics, and have nothing to do with quantization.

Equation (16.8) looks quantum mechanical in origin but this is deceptive since the $1/\hbar$ in the exponent on the right-hand side is canceled by the $\hbar$ introduced in the definition (16.6) of $\mathbf{L}$. Any other constant with the dimension of an action would have done as well. It needs "new physics", in this case the uncertainty relation, to fix this otherwise arbitrary constant to the value of $\hbar$.

Equation (16.8) tells us how to transform a scalar quantity such as the density from one coordinate system to another. Let us look at what happens with operators under rotations? As an example, we take a typical Hamilton operator

$$\mathrm{H} = \frac{\mathbf{p}^2}{2M} + V(x, y, z). \tag{16.9}$$

We denote the same Hamilton operator in the rotated frame by

$$\mathrm{H}' = \mathrm{R}_z(\alpha)\mathrm{H}, \tag{16.10}$$

where we have for abbreviation introduced the operator $\mathrm{R}_z(\alpha)$ for the counter-clockwise rotation by an angle α about the z-axis. A possible misunderstanding should be avoided: In (16.10), the application of $\mathrm{R}(\alpha, \beta, \gamma)$ to H means a rotation, but *not* a matrix multiplication! $\mathrm{R}_z(\pi/2)$, for instance, is nothing but a shorthand notation of the prescription: "Rotate the coordinate system counter clockwise by an angle of $\pi/2$ about the z-axis."

Transferring the results obtained in (16.8) for the rotation of a scalar to the rotation of an operator, we now get

$$\mathrm{H}' = \mathrm{e}^{-\mathrm{i}\alpha \mathrm{L}_z/\hbar} \mathrm{H} \mathrm{e}^{\mathrm{i}\alpha \mathrm{L}_z/\hbar}. \tag{16.11}$$

Note that now, in contrast to (16.8), H is sandwiched between $\exp(-\mathrm{i}\alpha \mathrm{L}_z/\hbar)$ and $\exp(\mathrm{i}\alpha \mathrm{L}_z/\hbar)$. This assures that the rotation of a product of two operators AB or the product of an operator with a scalar equals the product of the rotations of the factors, e.g.,

$$\begin{aligned} (\mathrm{AB})' &= \mathrm{e}^{-\mathrm{i}\alpha \mathrm{L}_z/\hbar} \mathrm{AB}\, \mathrm{e}^{\mathrm{i}\alpha \mathrm{L}_z/\hbar} \\ &= \mathrm{e}^{-\mathrm{i}\alpha \mathrm{L}_z/\hbar} \mathrm{A}\, \mathrm{e}^{\mathrm{i}\alpha \mathrm{L}_z/\hbar}\, \mathrm{e}^{-\mathrm{i}\alpha \mathrm{L}_z/\hbar} \mathrm{B}\, \mathrm{e}^{\mathrm{i}\alpha \mathrm{L}_z/\hbar} = \mathrm{A}'\mathrm{B}'. \end{aligned} \tag{16.12}$$

For systems with rotational symmetry about the z-axis, $\mathrm{H}' = \mathrm{H}$, we see that L_z must commute with H because from $\exp(\mathrm{i}\alpha \mathrm{L}_z/\hbar)\mathrm{H} = \mathrm{H}\exp(\mathrm{i}\alpha \mathrm{L}_z/\hbar)$ follows $\mathrm{L}_z\mathrm{H} = \mathrm{H}\mathrm{L}_z$, cf. Sect. 6.1, or

$$[\mathrm{H}, \mathrm{L}_z] = 0. \tag{16.13}$$

The Heisenberg equation of motion (10.8) then requires that the time derivative of L_z vanishes, $\dot{\mathrm{L}}_z = (\mathrm{i}/\hbar)[\mathrm{H}, \mathrm{L}_z] = 0$.

This is the quantum mechanical version of the fact that *each symmetry generates a conserved quantity*, here the component L_z of the angular momentum.

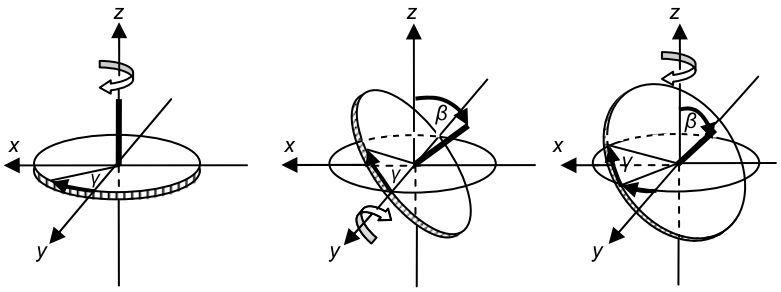


Fig. 16.1 The three Euler angles, γ for a rotation about z, β about y, and α again about z

In quantum mechanics, (16.13) means that an eigenfunction of H at the same time must be an eigenfunction of L_z.

The calculation can be immediately generalized to arbitrary rotations about any axis. It is known from geometry that any rotation can be decomposed into three successive rotations, first about the z-axis, second about the y-axis, and third again about the z-axis, yielding the rotation operator

$$\mathrm{R}(\alpha, \beta, \gamma) = \mathrm{R}_z(\alpha)\mathrm{R}_y(\beta)\mathrm{R}_z(\gamma), \tag{16.14}$$

where α, β, γ are the *Euler angles*, see Fig. 16.1. They are sufficient for an unambiguous characterization of arbitrary rotations.

16.2 Properties of Angular Momentum

There are a number of important relations for the angular-momentum operators that will be needed repeatedly. From definition (16.6) we obtain the following commutation relations for the components of the angular momentum operator,

$$[\mathrm{L}_y, \mathrm{L}_z] = \mathrm{i}\hbar\mathrm{L}_x, \tag{16.15a}$$

$$[\mathrm{L}_z, \mathrm{L}_x] = \mathrm{i}\hbar\mathrm{L}_y, \tag{16.15b}$$

$$[\mathrm{L}_x, \mathrm{L}_y] = \mathrm{i}\hbar\mathrm{L}_z. \tag{16.15c}$$

In terms of the vector $\mathbf{L}$, this can be written more concisely as

$$[\mathbf{L}, \mathbf{L}] = \mathrm{i}\hbar\mathbf{L}. \tag{16.16}$$

Note that exactly the same commutation relations follow from (10.9) for the spin matrices $\mathbf{S} = \frac{1}{2}\hbar\boldsymbol{\sigma}$.

The following steps, too, are more or less a repetition of the corresponding procedure for the spin matrices. The raising and lowering operators

$$\mathrm{L}_+ = \mathrm{L}_x + \mathrm{i}\mathrm{L}_y, \tag{16.17a}$$

$$\mathrm{L}_- = \mathrm{L}_x - \mathrm{i}\mathrm{L}_y, \tag{16.17b}$$

and the operator for the square of the total angular momentum

$$\mathbf{L}^2 = \mathrm{L}_x^2 + \mathrm{L}_y^2 + \mathrm{L}_z^2, \tag{16.18}$$

are of particular importance. A direct application of the commutation relation equations (16.15a)–(16.15c) shows that $\mathbf{L}^2$ commutes with each component of $\mathbf{L}$, in particular,

$$\left[\mathbf{L}^2, \mathrm{L}_z\right] = 0. \tag{16.19}$$

It is thus possible to find functions ψ_{lm}, or, in ket notation, $|lm\rangle$, which at the same time are eigenfunctions of $\mathbf{L}^2$ and L_z,

$$\mathrm{L}_z \psi_{lm} = m\hbar\psi_{lm}, \tag{16.20a}$$

$$\mathbf{L}^2 \psi_{lm} = l(l+1)\hbar^2 \psi_{lm}. \tag{16.20b}$$

For the raising and lowering operator, we obtain from (16.15a)–(16.15c) the commutation relations

$$[\mathrm{L}_z, \mathrm{L}_+] = \hbar\mathrm{L}_+, \tag{16.21a}$$

$$[\mathrm{L}_z, \mathrm{L}_-] = -\hbar\mathrm{L}_-. \tag{16.21b}$$

The first equation may be rewritten as

$$\mathrm{L}_z\mathrm{L}_+ = \mathrm{L}_+(\mathrm{L}_z + \hbar). \tag{16.22}$$

Application to ψ_{lm} yields

$$\mathrm{L}_z\mathrm{L}_+\psi_{lm} = (m+1)\hbar\mathrm{L}_+\psi_{lm}. \tag{16.23}$$

This shows that $\mathrm{L}_+\psi_{lm}$ again is an eigenfunction of L_z but now for the eigenvalue $m+1$, or

$$\mathrm{L}_+\psi_{lm} = \text{const.} \times \psi_{l,m+1}. \tag{16.24}$$

Analogous relations exist for L_-. This motivates the notations "raising" and "lowering" operators for L_+ and L_-, respectively. The constant in (16.24) is fixed by the condition that wave functions have to be normalized to $\langle lm|lm\rangle = 1$. Details of the calculation can be found in textbooks on quantum mechanics and shall not be reproduced here.

Let us just recapitulate the well known results. Using exclusively the commutation relation equations (16.15a)–(16.15c), one finds:

(i) l has to be integer or half-integer;
(ii) m can take $2l+1$ different values from $m = -l$ to $m = l$ in integer steps;

(iii) The raising relation (16.24) and the corresponding lowering relation with the correct constants reads

$$\mathrm{L}_{\pm}\psi_{lm} = \hbar\sqrt{l(l+1) - m(m\pm 1)}\,\psi_{l,m\pm 1}. \tag{16.25}$$

In particular, we have $\mathrm{L}_{+}\psi_{l,+l} = 0$ and $\mathrm{L}_{-}\psi_{l,-l} = 0$. It is thus not possible to raise or to lower the m quantum number beyond the allowed range between $m = -l$ to $m = l$ as it should be. When replacing l by $s = \frac{1}{2}$ the corresponding relations for the spin matrices given in Sect. 3.1 are recovered. This is no surprise since in the derivation only the commutation relations had been used, which are identical for the angular momentum operators and the spin matrices.

16.3 Representations

Taking the ψ_{lm} as basis functions, we may associate with each operator A a $(2l + 1) \times (2l + 1)$ matrix $A^{(l)}$, similar as in (3.14) for the case of a 2×2 matrix. We write this both in bra-ket notation $|lm\rangle$ and for the case of spatial wave functions

$$A^{(l)}_{mm'} = \langle lm|\mathrm{A}\big|lm'\big\rangle = \int \mathrm{d}^3 r\, \psi^*_{lm}(\boldsymbol{r})\mathrm{A}\psi_{lm'}(\boldsymbol{r}). \tag{16.26}$$

Such a mapping of an operator onto a matrix is called representation. The rank $2l+1$ of the representation is called the *dimension*.

Let us illustrate this for the components of angular momentum. Using $\mathrm{L}_x = (\mathrm{L}_+ + \mathrm{L}_-)/2$ and $\mathrm{L}_y = (\mathrm{L}_+ - \mathrm{L}_-)/2\mathrm{i}$, for the spin-$\frac{1}{2}$ case, we recover the two-dimensional representations (3.5) and (3.7a), (3.7b) for S_x, S_y, S_z.

For $l = 1$ we obtain in the same way

$$\mathrm{L}_x = \frac{\hbar}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad \mathrm{L}_y = \frac{\hbar}{\sqrt{2}}\begin{pmatrix} 0 & -\mathrm{i} & 0 \\ \mathrm{i} & 0 & -\mathrm{i} \\ 0 & \mathrm{i} & 0 \end{pmatrix},$$
$$\mathrm{L}_z = \hbar\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \tag{16.27a}$$

$$\mathbf{L}^2 = 2\hbar^2\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{16.27b}$$

The upper left element in the angular momentum matrices is for $m = m' = l$, the lower right element for $m = m' = -l$. We check that the matrix representation for the components of spin and angular momentum, namely, the spin-$\frac{1}{2}$ matrices and the above spin-1 matrices obey the same commutation relations (16.15a)–(16.15c) as the corresponding operators.

Looked upon in this way, the spin matrices lose their peculiar significance. They are just one representation, among many others, of the components of the angular momentum operator. The only thing that matters are the commutation relations.

Whether these relations are realized in terms of 2×2 matrices as for spin-$\frac{1}{2}$ or by differential operators as for the "ordinary" angular momentum is completely irrelevant in this context.

16.4 The Spherical Harmonics

In Sect. 16.2, we obtained eigenvalue equations for L_z and $\mathbf{L}^2$ using exclusively the commutation relations (16.15a)–(16.15c). Alternatively, we may start with the stationary Schrödinger equation

$$\left(\frac{\mathbf{p}^2}{2M} + V(x, y, z)\right)\psi = E\psi. \tag{16.28}$$

The first term representing the kinetic energy may be rewritten using the vector identity

$$[\boldsymbol{r} \times \boldsymbol{\nabla}]^2 = r^2\boldsymbol{\nabla}^2 - (\boldsymbol{r}\boldsymbol{\nabla})^2, \tag{16.29}$$

or

$$\mathbf{p}^2 = \mathrm{p}_r^2 + \frac{1}{r^2}\mathbf{L}^2, \tag{16.30}$$

where $p_r = (\boldsymbol{r} \cdot \mathbf{p})/r$ is the radial component of the momentum. This yields

$$\left(\frac{p_r^2}{2M} + \frac{\mathbf{L}^2}{2Mr^2} + V(x, y, z)\right)\psi = E\psi. \tag{16.31}$$

The two new terms are easily identified: The first term corresponds to the kinetic energy of the radial movement, and the second term is nothing but the rotational energy. Let us now rewrite the Schrödinger equation in polar coordinates, as given in (3.36). The operators for L_z and for $\mathbf{L}^2$ become

$$\mathrm{L}_z = \frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial\varphi}, \tag{16.32a}$$

$$\mathbf{L}^2 = -\hbar^2\left[\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2} + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)\right]. \tag{16.32b}$$

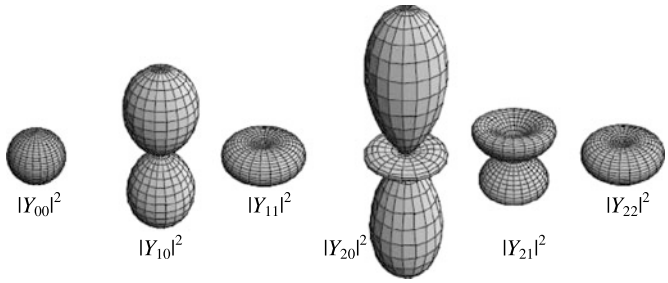


Fig. 16.2 The absolute values squared of the spherical harmonics for $l = 0, 1, 2$

The Schrödinger equation then reads

$$\left(-\frac{\hbar^2}{2Mr^2}\frac{\partial^2}{\partial r^2} + \frac{\mathbf{L}^2}{2Mr^2} + V(r,\theta,\varphi)\right)\psi(r,\theta,\varphi) = E\psi(r,\theta,\varphi). \tag{16.33}$$

The special case of full rotational symmetry, where the potential depends only on the radial coordinate, but not on the angle variables, $V(r,\theta,\varphi) = V(r)$, is of particular importance, This is the situation found for the hydrogen atom, for example. In this case, the dependence of the wave function on the angular variables may be separated,

$$\psi(r,\theta,\varphi) = \psi_{lm}(r)Y_{lm}(\theta,\varphi), \tag{16.34}$$

where $Y_{lm}(\theta,\varphi)$ is an eigenfunction both of L_z and $\mathbf{L}^2$,

$$\mathrm{L}_z Y_{lm}(\theta,\varphi) = m\hbar Y_{lm}(\theta,\varphi), \tag{16.35a}$$

$$\mathbf{L}^2 Y_{lm}(\theta,\varphi) = l(l+1)\hbar^2 Y_{lm}(\theta,\varphi). \tag{16.35b}$$

The $Y_{lm}(\theta,\varphi)$ are called *spherical harmonics*. They are orthogonal and normalized on the surface of the unit sphere,

$$\frac{1}{4\pi}\int_0^{2\pi} \mathrm{d}\varphi \int_0^{\pi} \sin\theta \,\mathrm{d}\theta\, Y^*_{lm}(\theta,\varphi)Y_{l'm'}(\theta,\varphi) = \delta_{ll'}\delta_{mm'}. \tag{16.36}$$

The spherical harmonics $Y_{lm}(\theta,\varphi)$ for l up to 2 (see Fig. 16.2) are given in Appendix A.4.

Spherical harmonics form a *complete set of orthonormal functions*. Therefore, the angle dependencies of physical quantities are often given in terms of spherical harmonics. Examples are the probability distribution of the electrons in the atomic shell, or the charge density distribution in a deformed nucleus, as sketched in Fig. 7.3, which is written as

$$\rho(r,\theta,\varphi) = \rho_0(r) + \rho_2(r)Y_{20}(\theta,\varphi). \tag{16.37}$$

A multipole expansion will be explicitly calculated in Sect. 18.2.

We know from textbook quantum mechanics that the eigenvalue equations (16.35a)–(16.35b) can only be realized if (i) l is an integer, and (ii) m is integer, too, and can take the $2l+1$ values $m=-l,\,-l+1,\ldots,l$. In Sect. 16.2, we came to the same conclusion from an algebraic analysis of the commutation relation, but there is one important difference:

From the commutation relations in (16.15a)–(16.15c), we got the result that l can take both integer and half-integer values, whereas an analysis of the Schrödinger equation allows only integer values for l. To get half-integer values in the Schrödinger equation as well, one has to introduce the Pauli spin matrices as an extra ingredient, which leads to a hybrid representation of the wave function in terms of a spinor part multiplied with a part depending on the spatial variables. This somewhat dissatisfactory situation is avoided if one uses matrix representations from the very beginning.

With the separation ansatz (16.34), the Schrödinger equation (16.33), a partial differential equation for the three coordinates r, φ, and θ, is reduced to an ordinary differential equation for the radial part of the wave function,

$$\left(-\frac{\hbar^2}{2Mr^2}\frac{\partial^2}{\partial r^2}+\frac{l(l+1)\hbar^2}{2Mr^2}+V(r)\right)\psi_{lm}(r)=E\psi_{lm}(r). \tag{16.38}$$

16.5 The Rotation Matrices

In view of the successful concept of replacing operators by their respective matrix representations, we now follow the same strategy for the rotation operators introduced in Sect. 16.1. Applying the rotation operator $\mathrm{R}(\alpha,\beta,\gamma)$ to the eigenfunctions ψ_{lm} for L_z and $\mathbf{L}^2$, from now on written as $|lm\rangle$, we get

$$\mathrm{R}(\alpha,\beta,\gamma)|lm\rangle=\mathrm{e}^{-\mathrm{i}\alpha\mathrm{L}_z/\hbar}\mathrm{e}^{-\mathrm{i}\beta\mathrm{L}_y/\hbar}\mathrm{e}^{-\mathrm{i}\gamma\mathrm{L}_z/\hbar}|lm\rangle. \tag{16.39}$$

The application of L_z does not change m, and the application of $\mathrm{L}_y=(\mathrm{L}_+-\mathrm{L}_-)/2\mathrm{i}$ raises and lowers m by one, whereas in both cases l is not changed at all. The application of $\mathrm{R}(\alpha,\beta,\gamma)$ to $|lm\rangle$, after expanding the exponentials in (16.39), is nothing but a repeated application of L_z, L_+, and L_-. It therefore results in a linear combination of the $|lm\rangle$, with m ranging from $-l$ to l, or

$$\mathrm{R}(\alpha,\beta,\gamma)|lm\rangle=\sum_{m'}D^{(l)}_{m'm}(\alpha,\beta,\gamma)\big|lm'\big\rangle. \tag{16.40}$$

Making use of the orthogonality relation for the $|lm\rangle$, one obtains the matrix elements of the rotation matrix

$$D^{(l)}_{m'm}(\alpha,\beta,\gamma)=\big\langle lm'\big|\mathrm{R}(\alpha,\beta,\gamma)|lm\rangle. \tag{16.41a}$$

Hence

$$D^{(l)}_{m'm}(\alpha,\beta,\gamma) = \langle lm'|\mathrm{e}^{-\mathrm{i}\alpha \mathrm{L}_z/\hbar}\mathrm{e}^{-\mathrm{i}\beta \mathrm{L}_y/\hbar}\mathrm{e}^{-\mathrm{i}\gamma \mathrm{L}_z/\hbar}|lm\rangle \tag{16.41b}$$

is a matrix representation of the rotation operator $\mathrm{R}(\alpha,\beta,\gamma)$ of rank l.

The equations above are written in standard bra-ket notation, as used in most modern textbooks. To smooth the transition from the matrix notation used in earlier chapters, we rewrite (16.40) as

$$\psi_l(\alpha,\beta,\gamma) = \mathrm{D}^{(l)}(\alpha,\beta,\gamma)\psi_l(0), \tag{16.42a}$$

where (0) stands for (0, 0, 0), or even shorter as

$$\psi'_l = \mathrm{D}^{(l)}\psi_l, \tag{16.42b}$$

with the original states assembled in the column $\psi_l = (|l,l\rangle,\ldots,|l,-l\rangle)^{\mathrm{T}}$, the rotated states in the column ψ'_l, and with the rotation matrix elements (16.41b) assembled into the matrix $\mathrm{D}^{(l)}$ of dimension $2l+1$.

The rotation matrices obey the orthogonality relation

$$\sum_m D^{(l)*}_{mm'}(\alpha,\beta,\gamma)D^{(l)}_{mm''}(\alpha,\beta,\gamma) = \delta_{m'm''}, \tag{16.43a}$$

$$\mathrm{D}^{(l)\dagger}\mathrm{D}^{(l)} = \mathrm{I}, \tag{16.43b}$$

which means that a back rotation $\mathrm{D}^{(l)\dagger}(\alpha,\beta,\gamma) = \mathrm{D}^{(l)}(-\gamma,-\beta,-\alpha)$ applied to ψ'_l restores the original state ψ_l.

The dependence of $D^{(l)}_{m'm}(\alpha,\beta,\gamma)$ on α and γ is trivial. Since the $|lm\rangle$ are eigenfunctions of L_z, that is, $\mathrm{L}_z|lm\rangle = m\hbar|lm\rangle$, we have

$$\mathrm{e}^{-\mathrm{i}\gamma \mathrm{L}_z/\hbar}|lm\rangle = \mathrm{e}^{-\mathrm{i}\gamma m}|lm\rangle, \tag{16.44a}$$

$$\langle lm'|\mathrm{e}^{-\mathrm{i}\gamma \mathrm{L}_z/\hbar} = \langle lm'|\mathrm{e}^{-\mathrm{i}\alpha m'}. \tag{16.44b}$$

This yields for the elements of the rotation matrices

$$D^{(l)}_{m'm}(\alpha,\beta,\gamma) = \mathrm{e}^{-\mathrm{i}(\alpha m'+\gamma m)}d^{(l)}_{m'm}(\beta), \tag{16.45}$$

where

$$d^{(l)}_{m'm}(\beta) = \langle lm'|\mathrm{e}^{-\mathrm{i}\beta \mathrm{L}_y/\hbar}|lm\rangle. \tag{16.46}$$

The rotation of spinors with spin-$\frac{1}{2}$ was derived in Sect. 6.2, based on the properties of Pauli matrices $\sigma_y^2 = 1$ and $\sigma_y^3 = \sigma_y$, with the result, see (6.8),

$$\mathrm{d}^{(1/2)}(\beta) = \begin{pmatrix} \cos\frac{1}{2}\beta & -\sin\frac{1}{2}\beta \\ \sin\frac{1}{2}\beta & \cos\frac{1}{2}\beta \end{pmatrix}. \tag{16.47}$$

For $l = 1$ one may proceed in an analogous way, using the relation $\mathrm{L}_y^3 = \mathrm{L}_y$, and finds

$$\mathrm{d}^{(1)}(\beta) = \begin{pmatrix} \frac{1}{2}(1+\cos\beta) & -\sqrt{\frac{1}{2}}\sin\beta & \frac{1}{2}(1-\cos\beta) \\ \sqrt{\frac{1}{2}}\sin\beta & \cos\beta & -\sqrt{\frac{1}{2}}\sin\beta \\ \frac{1}{2}(1-\cos\beta) & \sqrt{\frac{1}{2}}\sin\beta & \frac{1}{2}(1+\cos\beta) \end{pmatrix}. \tag{16.48}$$

The construction of rotation matrices $\mathrm{d}^{(l)}(\beta)$ for arbitrary values of l is somewhat more evolved and can be found in textbooks on angular momenta in quantum mechanics. The rotation matrices up to $l = 2$ are given in Appendix A.5. The full rotation matrix $\mathrm{D}^{(l)}(\alpha, \beta, \gamma)$ then is obtained from (16.45). The central element of a rotation matrix is given by the *Legendre polynomial* $D_{00}^{(l)} = P_l(\cos\theta)$, and the elements on the column and on the row through the center of the matrix are given by the spherical harmonics $D_{m0}^{(l)} = D_{0m}^{(l)\dagger} \propto Y_{lm}(\beta, -\alpha)$. As an example, for $l = 1$ and $m = 0$ (16.40) reads

$$\begin{aligned} \mathrm{R}|1,0\rangle &= D_{10}^{(1)}|1,+1\rangle + D_{00}^{(1)}|1,0\rangle + D_{-10}^{(1)}|1,-1\rangle \\ &= -\mathrm{e}^{-\mathrm{i}\alpha}\sqrt{\frac{1}{2}}\sin\beta|1,-1\rangle \\ &\quad + \cos\beta|1,0\rangle + \mathrm{e}^{\mathrm{i}\alpha}\sqrt{\frac{1}{2}}\sin\beta|1,+1\rangle. \end{aligned} \tag{16.49}$$

This page intentionally left blank

Chapter 17
Irreducible Tensors

Abstract Irreducible tensors are a powerful tool of quantum physics. These tensors are subsets of wave functions or operators that obey particularly simple transformation rules when experiencing a symmetry transformation. From the rotational symmetry of space follows that irreducible tensors of a given rank transform under rotation into linear combinations of each other. They all have the same universal structure. We then define the coupling of two such tensors to a product tensor. We explore the properties of irreducible tensors, with main emphasis on the Wigner–Eckart theorem.

17.1 Scalars, Vectors, and Tensors

In classical physics, a system of particles is characterized by quantities which can have different directional properties in space. A quantity whose magnitude V does not depend on the choice of the coordinate system, with $V' = V$ after a change of coordinates, is a *scalar*. Examples of scalar quantities are mass, charge, or energy. Other quantities do depend on the choice of coordinate system. A quantity composed of several components V_i may, upon a rotation of coordinates, change its components to a linear combination $V_i' = \sum_k a_{ik} V_k$, with a set of coefficients a_{ik}, for example with $i, k = x, y, z$ as in (6.6). This transformation property is used to define a *vector*. Vectors not only have a magnitude but also an orientation in space. Examples of vectors are position, linear or angular momentum, and electric or magnetic fields.

Quantities may have components arranged in a multiple array like T_{ik}, which may, for instance, be a set of nine components arranged in a 3×3 matrix. Let, upon a rotation of the coordinate system, these components transform linearly to $T_{ik}' = \sum_{mn} a_{im} a_{kn} T_{mn}$. This transformation property defines a *tensor* of *rank* 2. Examples of rank-two tensors are the moment of inertia of a system, which depends on the distribution of masses in space, or the electric quadrupole moment, which depends on the spatial distribution of electric charges. A high symmetry of the object under study may strongly reduce the number of coefficients a_{ik} needed, as does a good choice of the coordinate system. Scalars and vectors are tensors of rank 0 and 1, respectively. Tensors of higher rank can be constructed by going to higher orders of the multipole expansion, or by going beyond linear response.

D. Dubbers, H.-J. Stöckmann, *Quantum Physics: The Bottom-Up Approach*, Graduate Texts in Physics,
DOI 10.1007/978-3-642-31060-7_17, © Springer-Verlag Berlin Heidelberg 2013

In the course of the twentieth century, tensor algebra was developed to a high degree of refinement and is indispensable in modern physics and engineering. Tensor algebra permits the description of the interactions of various systems in a coordinate-free manner, with results depending only on relative coordinates. A simple example is the work $W = \boldsymbol{F} \cdot \boldsymbol{x}$ done on a system, which can be written either in terms of its Cartesian components $W = \sum_i F_i x_i$ in a given coordinate system, or in a coordinate-free manner as $W = Fx\cos\theta$, with the angle θ between the two vectors. Here two vectors of rank 1 are contracted to a scalar of rank 0. More generally, one can contract two tensors of equal or different ranks to tensor products of lower rank, or from two such tensors one can construct tensor products of the same or higher rank.

For interacting *quantum* systems, one wants to have the same techniques at hand. This generally is considered to be a thorny subject avoided in many textbooks on quantum mechanics. A down-to-earth approach to *irreducible tensors* in quantum physics will be presented in this and the following chapters. The formulae derived have no tedious summations over unobserved magnetic quantum numbers m, linked to a specific quantization axis z, but will depend only on angular momentum quantum numbers j, l, s, etc. of the system's constituents and on the ranks L of the tensors involved.

17.2 Properties of Irreducible Tensors

17.2.1 Definition of Irreducible Tensors

In (16.40), we showed that for arbitrary rotations the $|lm\rangle$, the eigenfunctions for L_z and $\mathbf{L}^2$, are transformed into linear combinations of $|lm'\rangle$ where m' may range from l to $-l$, while l does not change,

$$\mathrm{R}(\alpha,\beta,\gamma)|lm\rangle = \sum_{m'} D^{(l)}_{m'm}(\alpha,\beta,\gamma)|lm'\rangle. \tag{17.1}$$

On the other hand, one can show that it is not possible to find *smaller* subsets of eigenfunctions with this property, or expressed in other words, for an arbitrary rotation you really need *all* of the $|lm'\rangle$, $m' = -l, \ldots, l$ in the sum! When expressing these findings in the terms of group theory, one says that, with respect to the group of all rotations, the $|lm'\rangle$ are the components of an *irreducible tensor* of *rank* l.

The statement that there are no smaller irreducible subsets is no longer true if the symmetry is reduced. Let us take as an example a hydrogen atom in an external magnetic field along the z-axis. The hydrogen atom is symmetric with respect to *all* rotations, but in the presence of the magnetic field, only rotations about the z-axis remain as symmetry operations. For this situation, we obtain from (17.1), in particular,

$$\mathrm{R}(\alpha)|lm\rangle = \sum_{m'} D^{(l)}_{m'm}(\alpha,0,0)|lm'\rangle = \mathrm{e}^{-\mathrm{i}m\alpha}|lm\rangle. \tag{17.2}$$

Now no larger sets of $|lm'\rangle$ are needed any longer, on the contrary: Each $|lm\rangle$ is mapped, up to a phase factor, onto itself.

The concept of irreducibility is closely related to the degeneracy of eigenvalues. The fact that for each l there are $2l+1$ irreducible components means that in systems with full rotational symmetry each state with angular momentum quantum number l is $(2l+1)$-fold degenerate. And if there is only a rotational symmetry about the z-axis left, where each $|lm\rangle$ by itself forms an irreducible tensor of rank 1, the degeneracy is lifted. Exactly this is observed in the Zeeman effect (see Sect. 2.2).

Definition (17.1) for irreducible *tensor functions* $|lm\rangle$ is now transferred to irreducible *tensor operators*

$$\mathrm{R}(\alpha,\beta,\gamma)\mathrm{T}_{LM} = \sum_{M'} D^{(L)}_{M'M}(\alpha,\beta,\gamma)\mathrm{T}_{LM'}. \tag{17.3}$$

We shall see that any operator may be expanded in terms of irreducible tensor operators which, because of their simple rotation behavior, are particularly easy to handle.

Our notation: From now on, we write the integer numbers L and M that indicate the rank of an operator in upper case letters, and call T_L the set of $2L+1$ such operators. We write the integer quantum numbers l and m of the system's angular momentum eigenfunctions in lower case letters. To cover also half-integer spin, we shall from now on write $|j,m\rangle$ for these functions, with m shortly for m_j.

In shorthand matrix notation, omitting the dependences on the Euler angles, (17.3) reads

$$\mathrm{T}'_L = \mathrm{D}^{(L)}\mathrm{T}_L. \tag{17.4}$$

Equations (17.3) or (17.4) state that (i) the application of an arbitrary rotation to a set of tensors T_{LM} transfers each member of the set into a linear combination of generally all members of the set, and that (ii) it is not possible to find smaller subsets with this property.

The three components of a vector are an example of such an irreducible set.

A tensor of rank 0 is called *scalar*. It is invariant with respect to any rotation,

$$\mathrm{R}(\alpha,\beta,\gamma)\mathrm{T}_{00} = \mathrm{T}_{00}. \tag{17.5}$$

To check whether a given set of tensor operators is irreducible, one writes out the left-hand side of (17.3) as

$$\mathrm{T}'_L = \mathrm{R}\mathrm{T}_L = \mathrm{e}^{-\mathrm{i}\alpha\mathrm{L}_z/\hbar}\mathrm{e}^{-\mathrm{i}\beta\mathrm{L}_y/\hbar}\mathrm{e}^{-\mathrm{i}\gamma\mathrm{L}_z/\hbar}\mathrm{T}_L\mathrm{e}^{\mathrm{i}\gamma\mathrm{L}_z/\hbar}\mathrm{e}^{\mathrm{i}\beta\mathrm{L}_y/\hbar}\mathrm{e}^{\mathrm{i}\alpha\mathrm{L}_z/\hbar}, \tag{17.6a}$$

and looks whether the transformed T_{LM} obey (17.3), wherein coefficients $D^{(L)}_{MM'}$ are taken from (16.41b). For example, for a rotation about z, T_{LM} must obey the operator equation

$$\mathrm{e}^{-\mathrm{i}\gamma \mathrm{L}_z/\hbar}\mathrm{T}_{LM}\mathrm{e}^{\mathrm{i}\gamma \mathrm{L}_z/\hbar} = \mathrm{e}^{-\mathrm{i}\gamma M}\mathrm{T}_{LM}. \tag{17.6b}$$

17.2.2 A More Practical Definition

In general, (17.3) with (17.6a) is difficult to handle. There is, however, an alternative definition of irreducible tensor operators, which is more convenient for practical purposes. To find out, we first apply an infinitesimal rotation about the z-axis by an angle ε. For this case, one obtains from (17.3), with R from (17.6a) and $\mathrm{D}^{(L)}$ from (16.41b), up to terms linear in ε

$$\mathrm{T}_{LM} - \mathrm{i}\frac{\varepsilon}{\hbar}[\mathrm{L}_z, \mathrm{T}_{LM}] = \sum_{M'}\langle LM'|1 - \mathrm{i}\frac{\varepsilon}{\hbar}\mathrm{L}_z|LM\rangle \mathrm{T}_{LM'}, \tag{17.7}$$

whence it follows that

$$[\mathrm{L}_z, \mathrm{T}_{LM}] = \sum_{M'}\langle LM'|\mathrm{L}_z|LM\rangle \mathrm{T}_{LM'}. \tag{17.8}$$

For symmetry reasons, analogous relations hold for L_x and L_y, and thus also for the raising and lowering operators $\mathrm{L}_\pm = \mathrm{L}_x \pm \mathrm{i}\mathrm{L}_y$,

$$[\mathrm{L}_\pm, \mathrm{T}_{LM}] = \sum_{M'}\langle LM'|\mathrm{L}_\pm|LM\rangle \mathrm{T}_{LM'}. \tag{17.9}$$

Inserting the matrix elements from (16.20a) and (16.25) into the right-hand sides of (17.8) and (17.9), we get

$$[\mathrm{L}_z, \mathrm{T}_{LM}] = \hbar M \mathrm{T}_{LM}, \tag{17.10a}$$

$$[\mathrm{L}_\pm, \mathrm{T}_{LM}] = \hbar\sqrt{L(L+1) - M(M\pm 1)}\,\mathrm{T}_{L,M\pm 1}. \tag{17.10b}$$

We derived these equations from (17.3). Conversely, it is also possible to derive (17.3) from (17.10a), (17.10b). Both definitions are thus equivalent.

When constructing an irreducible tensor it is sufficient to check that the commutation relations of (17.10a), (17.10b) are fulfilled.

17.2.3 Simple Examples

Let us illustrate this with some examples.

1. From (17.10a), a scalar T_{00} has $[L_z, T_{00}] = 0$ and $[L_\pm, T_{00}] = 0$. Therefore, the square of the angular momentum $\mathbf{L}^2$ is a scalar, too, because it has the same commutation relations as T_{00}, cf. (16.19),

$$\left[L_z, \mathbf{L}^2\right] = 0, \qquad \left[L_\pm, \mathbf{L}^2\right] = 0. \tag{17.11}$$

2. A tensor of rank 1 can be constructed from the components (L_x, L_y, L_z) of angular momentum to give the irreducible (L_{-1}, L_0, L_1), with components $L_0 = L_z$, $L_{\pm1} = \mp\sqrt{\frac{1}{2}}L_\pm$, or

$$T_{10} = L_z/\hbar, \tag{17.12a}$$

$$T_{1\pm1} = \mp\sqrt{\frac{1}{2}}L_\pm/\hbar. \tag{17.12b}$$

   The irreducible tensor components $L_{\pm1}$ are closely related to the raising and lowering operators $L_\pm = L_x \pm iL_y$. Be aware of the $\mp$ (not $\pm$!) sign on the right-hand side of the second equation. The proof follows immediately from the commutation relations (16.15a)–(16.15c) which may be rewritten in terms of L_0 and $L_{\pm1}$ as

$$[L_+, L_1] = 0, \qquad [L_z, L_1] = \hbar L_1, \qquad [L_-, L_1] = \sqrt{2}\hbar L_0, \tag{17.13a}$$

$$[L_+, L_0] = \sqrt{2}\hbar L_1, \qquad [L_z, L_0] = 0, \qquad [L_-, L_0] = \sqrt{2}\hbar L_{-1}, \tag{17.13b}$$

$$[L_+, L_{-1}] = \sqrt{2}\hbar L_0, \qquad [L_z, L_{-1}] = -\hbar L_{-1}, \qquad [L_-, L_{-1}] = 0. \tag{17.13c}$$

   These are exactly the relations (17.10a), (17.10b) that define an irreducible tensor, with $T_{10} = L_0$, $T_{1\pm1} = L_{\pm1}$ for the case $L = 1$.
3. The same can be done for any vector, e.g., for the Cartesian coordinates of the radius vector $\boldsymbol{r} = (x, y, z)$,

$$r_0 = z, \qquad r_{\pm1} = \mp\sqrt{\frac{1}{2}}(x \pm iy). \tag{17.14}$$

   With the help of polar coordinates, this may alternatively be expressed as

$$r_0 = r\cos\theta, \qquad r_{\pm1} = \mp\sqrt{\frac{1}{2}}\sin\theta e^{\pm i\varphi}, \tag{17.15}$$

   or, using the spherical harmonic, as

$$r_M = \sqrt{\frac{4}{3}\pi}\, r Y_{1M}(\theta, \varphi), \quad M = -1, 0, 1. \tag{17.16}$$

4. An important second rank tensor can be constructed from the components of angular momentum, according to

$$\mathrm{T}_{20} = \left(3\mathrm{L}_z^2 - \mathbf{L}^2\right)/\hbar^2, \tag{17.17a}$$

$$\mathrm{T}_{2\pm1} = \mp\sqrt{\frac{3}{2}}(\mathrm{L}_\pm\mathrm{L}_z + \mathrm{L}_z\mathrm{L}_\pm)/\hbar^2, \tag{17.17b}$$

$$\mathrm{T}_{2\pm2} = \sqrt{\frac{3}{2}}\mathrm{L}_\pm^2/\hbar^2, \tag{17.17c}$$

   which can again be verified via the commutation relations (17.10a), (17.10b) for this tensor. Note that for $\mathrm{T}_{2\pm1}$ the product of $\mathrm{L}_\pm$ and L_z on the right-hand side is symmetrized. This is needed whenever products of noncommuting operators are involved. When T_{LM} is constructed from the components of angular momentum $\mathbf{L}$, we indicate this by writing $\mathrm{T}_{LM}(l)$, where L denotes the rank of the tensor, and l the angular momentum quantum number of the system under study.
5. In the very same way, we may construct a second rank tensor from the Cartesian components x, y, z of $\boldsymbol{r}$. Again the result may be expressed in terms of spherical harmonics,

$$\mathrm{T}_{2M}(r) = \sqrt{\frac{8\pi}{15}}r^2 Y_{2M}(\theta,\varphi), \quad M = -2,\ldots,2. \tag{17.18}$$

   In this case, the symmetrization is not needed since x, y, z, considered as operators, are commuting.

All irreducible tensors have the same structure, and in practice one needs to know only a relatively small set of tensor operators, namely T_{1M} from (17.12a), (17.12b), and T_{2M} from (17.17a)–(17.17c). Therein the components L_z and $\mathrm{L}_\pm$ of the angular momentum may be replaced by the corresponding components of any other vector operator.

Such a replacement can be a position vector r_M, see (17.16), a total angular momentum $\mathbf{J}$ (see below), a gradient vector $\boldsymbol{\nabla}$, etc. Instead of $\mathrm{T}_{LM}(l)$, one then writes

$$\mathrm{T}_{LM}(r), \qquad \mathrm{T}_{LM}(j), \qquad \mathrm{T}_{LM}(\nabla),$$

generally $\mathrm{T}_{LM}(A)$ for operator A.

17.3 The Coupling of Irreducible Tensors

The defining equations (17.3) or (17.10a), (17.10b) allow one to check whether a tensor is irreducible, but they give no prescription how to construct such a tensor.

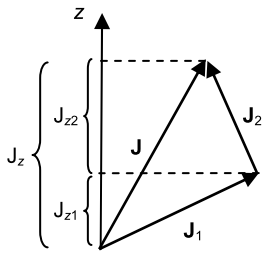


Fig. 17.1 Coupling of two angular momenta $\boldsymbol{J}_1$ and $\boldsymbol{J}_2$ to a total angular momentum $\boldsymbol{J}$

This can be done by building up tensors of higher rank from tensors of low rank. Let us first recall the rules governing the coupling of angular momenta, shortly treated for the case of two spin-$\frac{1}{2}$ particles in Sect. 8.2.

17.3.1 The Coupling of Angular Momenta

Angular momenta are omnipresent in nuclei, atoms and molecules, and there are several different types. Already in the simplest of all atoms, the ordinary hydrogen atom, we meet three of them: (i) the orbital angular momentum from the electron, (ii) the spin of the electron, and (iii) the spin of the proton. In molecules and in heavily deformed nuclei, there may be further angular momenta from collective rotations. This brings the problem of coupling two angular momenta, for instance, $\mathbf{J}_1$ and $\mathbf{J}_2$, to a total angular momentum

$$\mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2. \tag{17.19}$$

Classically, one finds the following relations, see Fig. 17.1:

1. The angular momenta are coupled vectorially, and it holds

$$\big||\mathbf{J}_1| - |\mathbf{J}_2|\big| \le |\mathbf{J}| \le |\mathbf{J}_1 + \mathbf{J}_2|. \tag{17.20}$$

   This is the triangle rule, and we shall abbreviate it by the notation $\Delta(\mathbf{J}_1\mathbf{J}_2\mathbf{J})$ in the following.
2. The *z-compo*nents of the angular momenta couple algebraically,

$$\mathrm{J}_z = \mathrm{J}_{1z} + \mathrm{J}_{2z}. \tag{17.21}$$

   The same, of course, is true for the other components of $\mathbf{J}$.

In *quantum* physics, the components of $\mathbf{J}$ obey the commutation rules (16.15a)–(16.15c) if this is true for the components of $\mathbf{J}_1$ and $\mathbf{J}_2$. Hence it must be possible to construct eigenfunctions $|jm\rangle$ for the total angular momentum from the eigenfunctions $|j_1m_1\rangle$ and $|j_2m_2\rangle$ of $\mathbf{J}_1$ and $\mathbf{J}_2$, respectively. We recall the ansatz of (8.11)

$$|jm\rangle = \sum_{m_1m_2} \langle j_1m_1j_2m_2|jm\rangle |j_1m_1\rangle |j_2m_2\rangle, \tag{17.22}$$

where the $\langle j_1 m_1 j_2 m_2 | jm\rangle$ are the Clebsch–Gordan coefficients. They have to be determined such that the $|jm\rangle$ defined by (17.22) are normalized eigenfunctions of $\mathbf{J}$. Details can be found in textbooks on angular momenta in quantum mechanics.

Most important for practical purposes are the two selection rules

$$\langle j_1 m_1 j_2 m_2 | jm\rangle \neq 0 \quad \text{only if} \quad \begin{cases} m = m_1 + m_2, \\ \triangle(j_1 j_2 j), \end{cases} \tag{17.23}$$

where the triangle rule $\triangle(j_1 j_2 j)$ now has to be interpreted in terms of quantum numbers as $|j_1 - j_2| \leq j \leq j_1 + j_2$. Equations (17.23) are the quantum mechanical analogue of the corresponding classical relations (17.20) and (17.21).

The Clebsch–Gordan coefficients obey the following symmetry relations:

$$\langle j_1 m_1 j_2 m_2 | jm\rangle = (-1)^{j_1+j_2-j} \langle j_1, -m_1, j_2, -m_2 | j, -m\rangle \tag{17.24a}$$

$$= (-1)^{j_1+j_2-j} \langle j_2 m_2 j_1 m_1 | jm\rangle \tag{17.24b}$$

$$= (-1)^{j_1-m_1} \sqrt{\frac{2j+1}{2j_2+1}} \langle j_1, m_1, j, -m | j_2, -m_2\rangle. \tag{17.24c}$$

Further symmetry relations are obtained by combinations of the given ones. The most frequently occurring Clebsch–Gordan coefficients are given in Appendix A.6.

The Clebsch–Gordan coefficients obey the two orthogonality relations

$$\sum_{m_1 m_2} \langle j_1 m_1 j_2 m_2 | j' m'\rangle \langle j_1 m_1 j_2 m_2 | jm\rangle = \delta_{jj'} \delta_{mm'}, \tag{17.25a}$$

$$\sum_{jm} \langle j_1 m_1' j_2 m_2' | jm\rangle \langle j_1 m_1 j_2 m_2 | jm\rangle = \delta_{m_1 m_1'} \delta_{m_2 m_2'}. \tag{17.25b}$$

These equations are a consequence of the corresponding orthogonality relations for the eigenfunctions of $\mathbf{J}_1$, $\mathbf{J}_2$, and $\mathbf{J}$. We note that these equations hold for a given set of quantum numbers m_1, m_2, m_1', m_2', with total magnetic quantum number $m = m_1 + m_2 = m_1' + m_2'$ fixed by (17.23). Thus summation in (17.25a) over m_2 and in (17.25b) over m is purely formal, that is, one only sums over m_1 and j, respectively. In fact, both (17.25a) and (17.25b) express the orthogonality condition

$$\mathrm{C}_m^\dagger \mathrm{C}_m = \mathrm{C}_m \mathrm{C}_m^\dagger = 1, \tag{17.26}$$

with the Clebsch–Gordan matrices $\mathrm{C}_m(j_1 \otimes j_2)$ of Sect. 8.2.

Using the above orthogonality conditions, one can write the inverse expansion of (17.22) as

$$|j_1 m_1\rangle |j_2 m_2\rangle = \sum_{jm} |jm\rangle \langle jm | j_1 m_1 j_2 m_2\rangle. \tag{17.27}$$

17.3.2 General Tensor Coupling

Irreducible tensor operators obey the same relations (17.3) as the corresponding functions (17.1). Therefore, (17.22) still holds if one or both of the kets on its right-hand side are replaced by operators. The application of an irreducible tensor, e.g., to a wave function, yields again a wave function,

$$|jm\rangle = \text{const.} \times \sum_{M_1 m_2} \langle L_1 M_1 j_2 m_2 | jm\rangle \mathrm{T}_{L_1 M_1} |j_2 m_2\rangle, \tag{17.28}$$

where the unspecified constant has to be fixed by the normalization condition for $|jm\rangle$. The application of an irreducible tensor to another irreducible tensor generates the new irreducible tensor

$$\mathrm{T}_{LM}(A, B) = \sum_{M_1 M_2} \langle L_1 M_1 L_2 M_2 | LM\rangle \mathrm{T}_{L_1 M_1}(A) \mathrm{T}_{L_2 M_2}(B). \tag{17.29}$$

In this case, the constant may be fixed to one, as long as there is no normalization prescription for the irreducible tensors. We shall come back to this point in Sect. 19.2.1. Quite often the inverse of (17.29) will be needed, i.e., the analogue of (17.27),

$$\mathrm{T}_{L_1 M_1}(A) \mathrm{T}_{L_2 M_2}(B) = \sum_{LM} \langle L_1 M_1 L_2 M_2 | LM\rangle \mathrm{T}_{LM}(A, B). \tag{17.30}$$

The arguments A and B are introduced in (17.29) to indicate that the components of the tensor are constructed from the components of an operator A. In the present context, A usually means an angular momentum quantum number, but, as stated before, it may indicate any other vector operator.

Expression (17.29) is called a tensor product and is written in shorthand notation as

$$\mathrm{T}_L(A, B) = \left[\mathrm{T}_{L_1}(A) \times \mathrm{T}_{L_2}(B)\right]_L. \tag{17.31}$$

Depending on the angular momenta involved, the tensor product is either symmetric or antisymmetric with respect to commutation of its factors,

$$\left[\mathrm{T}_{L_2}(B) \times \mathrm{T}_{L_1}(A)\right]_L = (-1)^{L_1+L_2-L} \left[\mathrm{T}_{L_1}(A) \times \mathrm{T}_{L_2}(B)\right]_L. \tag{17.32}$$

This is a direct consequence of the symmetry relation (17.24b) for the Clebsch–Gordan coefficients.

17.3.3 Some Special Cases

A number of tensor products are of particular importance:

1. Coupling of two tensors of rank L to a scalar.
   For this case, (17.29) yields (we replace $L_1 = L_2$ by L).

$$\mathrm{T}_{00}(A, B) = \sum_M \langle L, M, L, -M|00\rangle \mathrm{T}_{LM}(A)\mathrm{T}_{L,-M}(B). \tag{17.33}$$

   With $\langle L, M, L, -M|00\rangle = (-1)^{L-M}/(2L+1)^{1/2}$ from (A.29) with (A.28) this yields, using the complex conjugation property $\mathrm{T}^\dagger_{LM} = (-1)^M \mathrm{T}_{L,-M}$ (analogous as known from the spherical harmonics),

$$\mathrm{T}_{00}(A, B) = \frac{(-1)^L}{\sqrt{2L+1}} \sum_M \mathrm{T}^\dagger_{LM}(A)\mathrm{T}_{LM}(B). \tag{17.34}$$

   Next we discuss the coupling of two tensors of rank 1. Because of the triangle rule, the rank of the tensor product can take only the values 0, 1, 2. We start with $L = 0$:
2. Coupling of two tensors of rank 1 to a scalar.
   For this situation, we obtain as a special case of (17.34)

$$\mathrm{T}_{00}(A, B) = -\sqrt{\frac{1}{3}} \sum_M \mathrm{T}^\dagger_{1M}(A)\mathrm{T}_{1M}(B). \tag{17.35}$$

   Generalizing (17.12a), (17.12b), we write the tensor components as

$$\mathrm{T}_{10}(A) = \mathrm{A}_z, \tag{17.36a}$$

$$\mathrm{T}_{1,\pm1}(A) = \mp\sqrt{\frac{1}{2}}(\mathrm{A}_x \pm \mathrm{i}\mathrm{A}_y), \tag{17.36b}$$

   plus the corresponding relations for $\mathrm{T}_{1M}(B)$. We then get, up to a constant factor, the ordinary scalar product

$$\mathrm{T}_{00}(A, B) = -\sqrt{\frac{1}{3}}\mathbf{A} \cdot \mathbf{B}. \tag{17.37}$$

   Well known examples are the hyperfine Hamiltonian $\mathrm{H}_{\mathrm{IJ}} = A\mathbf{I} \cdot \mathbf{J}$ from Sect. 8.3, or the spin-orbit Hamiltonian

$$\mathrm{H}_{\mathrm{LS}} = a\mathbf{L} \cdot \mathbf{S} \tag{17.38}$$

   that induces the splitting of the D lines of the sodium atom mentioned at the end of Sect. 2.2.
3. Coupling of two tensors of rank 1 to a vector.

For this case, (17.31) yields

$$\mathrm{T}_{10}(A,B)=\sqrt{\frac{1}{8}}(\mathrm{A}_{-}\mathrm{B}_{+}-\mathrm{A}_{+}\mathrm{B}_{-}), \tag{17.39a}$$

$$\mathrm{T}_{1\pm1}(A,B)=\frac{1}{2}(\mathrm{A}_{z}\mathrm{B}_{\pm}-\mathrm{A}_{\pm}\mathrm{B}_{z}), \tag{17.39b}$$

where we have used the Clebsch–Gordan coefficients (Appendix A.6)

$$\langle 1,\pm1,1,\mp1|1,0\rangle=\mp\sqrt{\frac{1}{2}}=\langle 1,\pm1,1,0|1,\pm1\rangle, \tag{17.40a}$$

$$\langle 1,0,1,\pm1|1,\pm1\rangle=\mp\frac{1}{2}. \tag{17.40b}$$

Equations (17.39a), (17.39b) take a more suggestive form if we rewrite not only the factors on the right of (17.29), but also the tensor product $\mathrm{T}_1(A,B)$ in terms of its Cartesian components C_x, C_y, C_z (which may be noncommuting operators)

$$\mathrm{T}_{10}(A,B)=\mathrm{C}_z, \tag{17.41a}$$

$$\mathrm{T}_{1,\pm1}(A,B)=\mp\sqrt{\frac{1}{2}}(\mathrm{C}_x\pm\mathrm{i}\mathrm{C}_y). \tag{17.41b}$$

With this, relations (17.39a), (17.39b) read

$$\mathrm{C}_x=\mathrm{i}\sqrt{\frac{1}{2}}(\mathrm{A}_y\mathrm{B}_z-\mathrm{A}_z\mathrm{B}_y), \tag{17.42a}$$

$$\mathrm{C}_y=\mathrm{i}\sqrt{\frac{1}{2}}(\mathrm{A}_z\mathrm{B}_x-\mathrm{A}_x\mathrm{B}_z), \tag{17.42b}$$

$$\mathrm{C}_z=\mathrm{i}\sqrt{\frac{1}{2}}(\mathrm{A}_x\mathrm{B}_y-\mathrm{A}_y\mathrm{B}_x), \tag{17.42c}$$

or, in shorthand notation,

$$\mathbf{C}=\mathrm{i}\sqrt{\frac{1}{2}}\mathbf{A}\times\mathbf{B}. \tag{17.43}$$

This is, up to a factor, nothing but the ordinary vector product. Scalar and vector products thus show up as special cases of the tensor product (17.31).

4. Coupling of two tensors of rank 1 to a tensor of rank 2.

For this case, (17.29) yields

$$\begin{aligned}\mathrm{T}_{20}(A,B)&=\sqrt{\frac{1}{6}}\left(2\mathrm{A}_z\mathrm{B}_z-\frac{1}{2}\mathrm{A}_{+}\mathrm{B}_{-}-\frac{1}{2}\mathrm{A}_{-}\mathrm{B}_{+}\right)\\&=\sqrt{\frac{1}{6}}(3\mathrm{A}_z\mathrm{B}_z-\mathbf{A}\cdot\mathbf{B}),\end{aligned} \tag{17.44a}$$

$$\mathrm{T}_{2\pm1}(A,B) = \mp\frac{1}{2}(\mathrm{A}_{\pm}\mathrm{B}_z + \mathrm{A}_z\mathrm{B}_{\pm}), \tag{17.44b}$$

$$\mathrm{T}_{2\pm2}(A,B) = \frac{1}{2}\mathrm{A}_{\pm}\mathrm{B}_{\pm}, \tag{17.44c}$$

where we used the coefficients from Appendix A.6. We shall see that this tensor describes the dipole–dipole interaction between spins if the corresponding spin operators are inserted for A and B. For $A = B = L$, we recover, up to a factor, the rank 2 tensor product already introduced in (17.17a)–(17.17c). This tensor will be important for the description of the interaction of quadrupole moments with external electric field gradients, see Sect. 18.2.

17.4 The Wigner–Eckart Theorem

The irreducible tensor formalism exhibits its full power if matrix elements $\langle j'm'|\mathrm{T}_{LM}|jm\rangle$ of tensor operators in angular momentum representation are needed. The application of the tensor operator T_{LM} on the wave function $|jm\rangle$ may be expanded as

$$\mathrm{T}_{LM}|jm\rangle \propto \sum_{j''m''} |j''m''\rangle\langle jmLM|j''m''\rangle. \tag{17.45}$$

This is just the inverse of (17.28), where the orthogonality of the Clebsch–Gordan coefficients has been used. With this expression, we get for the tensor matrix element

$$\langle j'm'|\mathrm{T}_{LM}|jm\rangle \propto \sum_{j''m''} \langle j'm'|j''m''\rangle\langle jmLM|j''m''\rangle. \tag{17.46}$$

Because of the orthogonality $\langle j'm'|j''m''\rangle = \delta_{j'j''}\delta_{m'm''}$ of the irreducible tensor functions, the only nonzero term in the sum is the term $j'' = j'$ and $m'' = m'$, yielding

$$\langle j'm'|\mathrm{T}_{LM}|jm\rangle \propto \langle jmLM|j'm'\rangle, \tag{17.47}$$

where the constant of proportionality depends on the angular momentum quantum numbers involved, but not on the m quantum numbers.

Usually (17.47) is written in the following form:

$$\langle j'm'|\mathrm{T}_{LM}|jm\rangle = \frac{1}{\sqrt{2j'+1}}\langle j'||\mathrm{T}_L||j\rangle\langle jmLM|j'm'\rangle. \tag{17.48}$$

This is the *Wigner–Eckart theorem*, where the *reduced matrix element* $\langle j'||\mathrm{T}_L||j\rangle$ is independent of m, m', and M.

The selection rules for the Clebsch–Gordan coefficients (17.23) map immediately onto the corresponding selection rules for the tensor matrix element,

$$\langle j'm'|\mathrm{T}_{LM}|jm\rangle \neq 0 \quad \text{only if} \quad \begin{cases} m' = m + M, \\ \Delta(Ljj'). \end{cases} \tag{17.49}$$

An important consequence of the Wigner–Eckart theorem is that for $j = j'$ all matrix elements of tensors of rank $L > 2j$ vanish.

This means that for spin $s = \frac{1}{2}$ states the only existing tensors are scalars and vectors. This is the very origin of the fact that spin-$\frac{1}{2}$ systems allow such an easy and suggestive interpretation, as has been illustrated for many examples in the preceding chapters.

The Wigner–Eckart theorem allows, for instance, the calculation of the *relative* strengths in spectral line multiplets using exclusively the Clebsch–Gordan coefficients, see Sect. 18.3. No explicit knowledge of atomic or nuclear wave functions is needed for this task. Only for the *absolute* line intensities, which are encoded in the reduced matrix element, there is no alternative to an explicit calculation on the basis of the wave functions of the states involved.

This page intentionally left blank

Chapter 18
Electromagnetic Multipole Interactions

Abstract Irreducible tensor products are used to describe multipole interactions of spin systems exposed to static external fields, in particular the magnetic dipole interaction, and the electric quadrupole interaction of a nucleus with the gradient of a crystalline electric field. Then selection rules and relative line strengths for electromagnetic transitions in atoms are derived from the Wigner–Eckart theorem.

18.1 Static Magnetic Interactions

Because of the isotropy of space, any direction may be selected as the z-axis of the coordinate system, though there may be directions which are better suited than others, for instance, if there is an applied external magnetic field. This means that the Hamiltonian describing the interaction must be invariant with respect to arbitrary rotations, in other words, the Hamiltonian must be a scalar. Hence it must be possible to express it in terms of sums over tensor scalar products. We have met already repeatedly one example, namely, the interaction of a magnetic moment with an external magnetic field $\mathrm{H_M} = -\boldsymbol{\mu} \cdot \boldsymbol{B}$. In Sect. 17.3, we have already calculated how the scalar product can be expressed in terms of an irreducible tensor product, (17.35) and (17.37),

$$\mathrm{H_M} = -\sum_M \mathrm{T}^{\dagger}_{1M}(\mu)\mathrm{T}_{1M}(B). \tag{18.1}$$

Another important magnetic interaction is the dipole–dipole interaction between neighboring magnetic moments $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$,

$$\mathrm{H_{dipole}} = \frac{8\pi}{3}(\boldsymbol{\mu}_1 \cdot \boldsymbol{\mu}_2)\delta(\boldsymbol{r}) + \frac{3(\boldsymbol{r} \cdot \boldsymbol{\mu}_1)(\boldsymbol{r} \cdot \boldsymbol{\mu}_2) - (\boldsymbol{\mu}_1 \cdot \boldsymbol{\mu}_2)r^2}{r^5}, \tag{18.2}$$

where $\boldsymbol{r}$ is the distance vector between the two dipoles, and $r = |\boldsymbol{r}|$ its magnitude. The first term is called the *Fermi contact term*. It is important for the dipole–dipole interaction of a nuclear spin I with the angular momentum of the electronic shell. Because of the δ-function this term is nonzero for s-electrons with $l = 0$ only. Since this term is, up to a factor, just the scalar product between the involved dipoles, where we already know how to rewrite it as a tensor product, there is no need

D. Dubbers, H.-J. Stöckmann, *Quantum Physics: The Bottom-Up Approach*, Graduate Texts in Physics,
DOI 10.1007/978-3-642-31060-7_18, © Springer-Verlag Berlin Heidelberg 2013

to discuss it further. Another example of a scalar interaction between magnetic moments has already been given in Sect. 8.3 on the hyperfine interaction. The dipole–dipole interaction between nuclear moments is of particular importance in solid-state NMR, where it is responsible for the intrinsic line widths, as stated in Sect. 11.3.

To reexpress the second term in (18.2) in terms of a tensor product, we first write the various scalar products $\boldsymbol{r} \cdot \boldsymbol{\mu}_1$, etc. in terms of tensor products as above, and then write the appearing products $\mathrm{T}^{\dagger}_{LM}(\mu_1)\mathrm{T}_{LM}(\mu_2)$, etc. using the coupling formula (17.30). The result is

$$3(\boldsymbol{r} \cdot \boldsymbol{\mu}_1)(\boldsymbol{r} \cdot \boldsymbol{\mu}_2) - (\boldsymbol{\mu}_1 \cdot \boldsymbol{\mu}_2)r^2 = \sqrt{\frac{3}{2}} \sum_M \mathrm{T}^{\dagger}_{2M}(r)\mathrm{T}_{2M}(\mu_1, \mu_2), \tag{18.3}$$

where $\mathrm{T}_{2M}(r)$ and $\mathrm{T}_{2M}(\mu_1, \mu_2)$ are the tensors given in (17.18) and (17.44a)–(17.44c), respectively.

18.2 Static Electric Interactions

Next we discuss the interaction between an atomic nucleus and external electric fields with main emphasis on the electrostatic quadrupole interaction Hamiltonian H_{Q}, a case of interest in NMR spectroscopy met already in Sect. 7.1.3 for the special case $I = 1$.

18.2.1 Multipole Expansion of Electrostatic Energy

The electrostatic interaction between two charge densities is

$$W_{\mathrm{el}} = \int \frac{\rho_N(\boldsymbol{r}_1)\rho_e(\boldsymbol{r}_2)}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2. \tag{18.4}$$

In the present context, $\rho_N(\boldsymbol{r}_1)$ means the charge density of the nucleus, whereas $\rho_e(\boldsymbol{r}_2)$ is the charge density of the electronic shell and/or surrounding external charges producing the external electric field. For $\boldsymbol{r}_1 \neq \boldsymbol{r}_2$, we may expand $|\boldsymbol{r}_1 - \boldsymbol{r}_2|^{-1}$ as

$$\frac{1}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} = 4\pi \sum_{LM} \frac{1}{2L+1} \frac{r_<^L}{r_>^{L+1}} Y_{LM}(\theta_1, \varphi_1) Y^*_{LM}(\theta_2, \varphi_2), \tag{18.5}$$

where $r_<$ and $r_>$ are the smaller and the larger of the two distances $r_1 = |\boldsymbol{r}_1|$ and $r_2 = |\boldsymbol{r}_2|$, respectively, and θ_1, φ_1 and θ_2, φ_2 are the polar angles of $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ with respect to an external coordinate system; see, e.g., Jackson (1967). For s-electrons

with $l = 0$, there is an extra contribution because of their nonvanishing charge density at the site of the nucleus. Excluding this case, one obtains from (18.4)

$$W_{\mathrm{el}} = \sum_{LM} \mathrm{T}^*_{LM}(e)\mathrm{T}_{LM}(N), \tag{18.6}$$

with, because of $r_1(N) \leq r_2(e)$,

$$\mathrm{T}_{LM}(N) = \sqrt{4\pi/(2L+1)} \int r^L Y_{LM}(\theta, \varphi)\rho_N(\boldsymbol{r})\, \mathrm{d}^3 r, \tag{18.7a}$$

$$\mathrm{T}_{LM}(e) = \sqrt{4\pi/(2L+1)} \int \frac{1}{r^{L+1}} Y_{LM}(\theta, \varphi)\rho_e(\boldsymbol{r})\, \mathrm{d}^3 r. \tag{18.7b}$$

Expression (18.7a) describing the charge distribution of the nucleus is classical, but it can be turned into its quantum mechanical equivalent by substituting $\rho_N(\boldsymbol{r})$ with $eZ\psi^*_{Im_I}\psi_{Im_I}$, where Z is the nuclear charge and $\psi_{Im_I} = |Im_I\rangle$ (with nuclear spin quantum numbers I and m_I) is the nuclear wave function. We thus obtain

$$\mathrm{T}_{LM}(N) = \sqrt{4\pi/(2L+1)}eZ\langle Im_I|r^L Y_{LM}(\theta, \varphi)|Im'_I\rangle \tag{18.8}$$

as the quantum mechanical equivalent of (18.7a). Equation (18.6), together with (18.7b) and (18.8), provides us with the wanted expansion of the electric interaction in terms of tensor scalar products.

18.2.2 Electric Quadrupole Interaction

Let us now discuss the different terms in somewhat more detail. For $L = 0$ we have

$$\mathrm{T}_{00}(N) = eZ\langle Im_I|Im_I\rangle = eZ, \tag{18.9a}$$

$$\mathrm{T}_{00}(e) = \int \frac{1}{r}\rho_e(\boldsymbol{r})\, \mathrm{d}^3 r, \tag{18.9b}$$

the latter being the potential generated by the surrounding charges. Not surprisingly, the leading term in the expansion (18.6) is the Coulomb interaction of the charge of the nucleus with the charges in the environment.

The second term with $L = 1$ vanishes. This is due to the fact that parity is a good quantum number in nuclear physics, which means that nuclear wave functions $|Im_I\rangle$ are either even or odd with respect to a change of the sign of $\boldsymbol{r}$. Since $Y_{1M}(\theta, \varphi)$ is odd with respect to the parity operation, the integral (18.7a) yields zero for $L = 1$, both for even and odd wave functions. The same argument holds for all odd values of L.

The $L = 2$ electric quadrupole term is the one of main interest. The nonvanishing higher order terms $L = 4, 6, \ldots$ are usually too small to be detected. We separately discuss the nuclear and the electron part of the interaction.

First the nuclear part: According to the Wigner–Eckart theorem, matrix elements of the type entering the right-hand side of (18.8) are the same for all irreducible tensors of a given rank L, up to a factor not depending on the M quantum numbers. This allows us to reexpress (18.8) for $L = 2$ as

$$\mathrm{T}_{2M}(N) = \sqrt{\frac{4\pi}{5}} eZ \frac{\langle II|r^2 Y_{20}(\theta,\varphi)|II\rangle}{\langle II|\mathrm{T}_{20}(I)|II\rangle} \langle Im_I|\mathrm{T}_{2M}(I)|Im'_I\rangle, \tag{18.10}$$

where $T_{2M}(I)$ is given by (17.44a)–(17.44c) with $A = B = I$. Any other tensor of rank two $\mathrm{T}_{2M}(A)$ would have done as well on the right-hand side, but it is most practical to use $\mathrm{T}_{2M}(I)$ since for this tensor all matrix elements can be easily calculated. In particular, we obtain for the matrix element in the denominator

$$\langle II|\mathrm{T}_{20}(I)|II\rangle = \sqrt{\frac{1}{6}}\big[3I^2 - I(I+1)\big] = \sqrt{\frac{1}{6}} I(2I-1). \tag{18.11}$$

The matrix element in the numerator can be expressed in terms of the nuclear quadrupole moment (this is why these steps had been performed!). It is defined as

$$Q = Z\langle II|3z^2 - r^2|II\rangle = 4\sqrt{\frac{\pi}{5}} Z\langle II|r^2 Y_{20}(\theta,\varphi)|II\rangle. \tag{18.12}$$

This is the quantum mechanical version of definition (7.6) of the quadrupole moment. Collecting the results, we obtain

$$\mathrm{T}_{2M}(N) = eQ \frac{\sqrt{6}}{2I(2I-1)} \langle Im_I|\mathrm{T}_{2M}(I)|Im'_I\rangle = \langle Im_I|\mathrm{Q}_M|Im'_I\rangle, \tag{18.13}$$

where the Q_M are the components of the *quadrupole tensor* defined by

$$\mathrm{Q}_M = eQ \frac{\sqrt{6}}{2I(2I-1)} \mathrm{T}_{2M}(I). \tag{18.14}$$

Next the electron part: For the $L = 2$ term of the interaction, the integrand $\rho_e(\boldsymbol{r})$ in (18.9b) may be transformed as follows. We start with the identity

$$r^{-3} Y_{2M}(\theta,\varphi) = \sqrt{\frac{5}{4\pi}} \mathrm{T}_{2M}(\nabla) r^{-1}, \tag{18.15}$$

where $\mathrm{T}_{2M}(\nabla)$ is the tensor given in (17.44a)–(17.44c) with L replaced by ∇. The relation is checked by direct verification. Inserting (18.15) into the right-hand side of (18.7b) for $L = 2$, we obtain

$$\mathrm{T}_{2M}(e) = \sqrt{\frac{1}{6}} \int \big[\mathrm{T}_{2M}(\nabla) r^{-1}\big] \rho_e(\boldsymbol{r})\, \mathrm{d}^3 r. \tag{18.16}$$

This may alternatively be written as

$$\mathrm{T}_{2M}(e) = \sqrt{\frac{1}{6}}\mathrm{T}_{2M}(\nabla) \int |\boldsymbol{r} - \bar{\boldsymbol{r}}|^{-1} \rho_e(\bar{\boldsymbol{r}})\, \mathrm{d}^3\bar{r}|_{r=0}. \tag{18.17}$$

The integral turns out to be the potential generated by the external charges,

$$\varphi(\boldsymbol{r}) = \int |\boldsymbol{r} - \bar{\boldsymbol{r}}|^{-1} \rho_e(\bar{\boldsymbol{r}})\, \mathrm{d}^3\bar{r}. \tag{18.18}$$

We have thus obtained

$$\mathrm{T}_{2M}(e) = \sqrt{\frac{1}{6}}\mathrm{T}_{2M}(\nabla)\varphi(\boldsymbol{r})\Big|_{r=0} = q_M. \tag{18.19}$$

The tensor $\mathrm{T}_{2M}(e) = q_M$ is the *electric field gradient (efg) tensor*. It is obtained from the second derivatives of the potential and thus from the first derivatives of the components of the electric field.

The irreducible components q_M of the efg tensor are given explicitly by

$$q_0 = \frac{1}{6}(2\varphi_{zz} - \varphi_{xx} - \varphi_{yy}), \tag{18.20a}$$

$$q_{\pm1} = \mp\frac{1}{2}\sqrt{\frac{1}{6}}(\varphi_{zx} \pm \mathrm{i}\varphi_{zy}), \tag{18.20b}$$

$$q_{\pm2} = \frac{1}{2}\sqrt{\frac{1}{6}}(\varphi_{xx} - \varphi_{yy} \pm 2\mathrm{i}\varphi_{xy}), \tag{18.20c}$$

where we introduced the abbreviation

$$\varphi_{x_i x_j} = \frac{\partial^2\varphi}{\partial x_i \partial x_j}\bigg|_{r=0}. \tag{18.21}$$

Since the efg tensor in the Cartesian representation is symmetric, it can be transformed to its principal axes. If this is done, the off diagonal elements of the tensor vanish. A further constraint is given by the Poisson relation for the potential

$$\Delta\varphi = \varphi_{xx} + \varphi_{yy} + \varphi_{zz} = 0. \tag{18.22}$$

Note that there is no inhomogeneous term on the right-hand side since φ_e by definition only includes the contributions form the external charges. Again s-electrons with their nonvanishing charge density at the site of the nucleus are an exception, giving rise to an inhomogeneous term $-4\pi\rho_e(0)$. Discarding this possible complication, the efg tensor in principal axis representation is characterized by just two

parameters. It is common practice to introduce the notation

$$eq = \varphi_{zz}, \tag{18.23a}$$

$$\eta = (\varphi_{xx} - \varphi_{yy})/\varphi_{zz}, \tag{18.23b}$$

where eq is called just the *electric field gradient*, and η *asymmetry parameter*. Ordering the principal axes according to

$$|\varphi_{zz}| \geq |\varphi_{yy}| \geq |\varphi_{xx}|, \tag{18.24}$$

the parameter η can take values between 0 and 1 only. Now (18.20a)–(18.20c) take a particularly simple form

$$q_0 = \frac{1}{2}eq, \qquad q_{\pm 1} = 0, \qquad q_{\pm 2} = \frac{1}{2}\sqrt{\frac{1}{6}}\eta eq. \tag{18.25}$$

With (18.13) and (18.19), we obtain for the $L = 2$ term of expansion (18.6)

$$W_{\text{el}}^{(2)} = \sum_M \mathrm{T}_{2M}^*(e)\mathrm{T}_{2M}(N) = \langle Im_I | \sum_M q_M^* \mathrm{Q}_M | Im_I' \rangle = \langle Im_I | \mathrm{H_Q} | Im_I' \rangle, \tag{18.26}$$

where $\mathrm{H_Q}$ is the Hamiltonian for the quadrupolar interaction

$$\mathrm{H_Q} = \sum_M q_M^* \mathrm{Q}_M. \tag{18.27}$$

In principal axes orientation, the electric quadrupole Hamiltonian is

$$\mathrm{H_Q^0} = \frac{e^2 q\, Q}{4I(2I-1)\hbar}\left[3\mathrm{I}_z^2 - \mathbf{I}^2 + \frac{1}{2}\eta\left(\mathrm{I}_+^2 + \mathrm{I}_-^2\right)\right], \tag{18.28}$$

with eq and η from (18.25). The quantity $e^2 q\, Q$ is called the *quadrupole coupling constant*.

There are a number of conditions to be met to observe a quadrupole interaction. First of all, the lattice symmetry must be not too high. It is immediately clear form the discussion above that at sites with cubic symmetry, the quadrupole interaction must vanish. The same is true for sites with tetrahedral symmetry. From the definition of the asymmetry parameter, it is furthermore clear that η is zero, if the z-axis has a three- or four-fold rotational symmetry. And the nuclear angular momentum must be at least $I = 1$ for the matrix elements of $\mathrm{H_Q}$ not to vanish, see (17.49).

18.3 Selection Rules for Electromagnetic Transitions

Before the development of quantum mechanics, the existence of spectral lines posed an unsolved riddle. It was Bohr who proposed that the energy levels in an atom are

quantized, and that spectral lines correspond to electromagnetic transitions between two levels (see Sect. 10.1). It was soon noticed that there are *selection rules* meaning that not all possible combinations of energy levels are really observed in the spectra. And there was a systematics in the line strengths of the spectral line: The intensity ratio of the D_1 and the D_2 lines in Na, for example, was 1:2, and the same was true for the corresponding transitions in the other alkali atoms. By means of the Wigner–Eckart theorem, these questions can be answered in a systematic way.

The calculation of the emission probability for electromagnetic radiation is rather involved since vector fields have to be expanded obeying a number of constraints from the Maxwell equations. Details can be found, for instance, in the monograph of Blatt and Weisskopf (1952), or in the book of Rose (1957).

In the far field limit, the expansion is seen to consist of electric and magnetic multipole radiation components with amplitudes

$$(a_{\mathrm{E}})_{LM} \propto \langle j_f m_f | r^L Y_{LM}(\theta,\varphi) | j_i m_i \rangle, \tag{18.29a}$$

$$(a_{\mathrm{M}})_{LM} \propto \langle j_f m_f | \nabla \left[r^L Y_{LM}(\theta,\varphi) \right] \cdot \mathbf{L} | j_i m_i \rangle, \tag{18.29b}$$

where f and i refer to the final and initial state, respectively.

Provided the respective matrix elements do not vanish, the amplitudes of the multipole or 2^L-pole radiation decrease from order to order by a factor of the order of kR, where R characterizes the dimensions of the sources. This typically means a factor of about 1000, both for radiation from the electronic shell, and from the nucleus. Furthermore, electric 2^{L+1}-pole radiation is comparable in amplitude with magnetic 2^L-pole radiation.

Due to the Wigner–Eckart theorem, the relative transition amplitudes in (18.29a), (18.29b), being matrix elements of irreducible tensors, can be expressed in terms of Clebsch–Gordan coefficients,

$$(a_{\mathrm{E}})_{LM}, (a_{\mathrm{M}})_{LM} \propto \langle j_i m_i LM | j_f m_f \rangle. \tag{18.30}$$

Hence the *angular momentum* selection rules for radiative transitions are

$$(a_{\mathrm{E}})_{LM}, (a_{\mathrm{M}})_{LM} \neq 0 \quad \text{only if} \quad \begin{cases} M = m_f - m_i, \\ \Delta(j_f L j_i). \end{cases} \tag{18.31}$$

In this respect, electric and magnetic multipole radiation behave identically. But there is a difference with respect to parity. If $\pi_i = (-1)^{l_i}$ and $\pi_f = (-1)^{l_f}$ are the parities of initial and final state, respectively, one notices from inspection of the matrix elements (18.29a), (18.29b) the following:

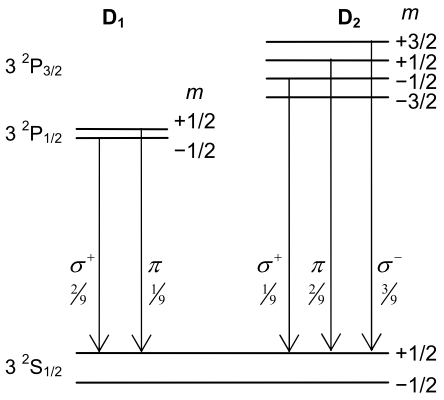


Fig. 18.1 The Zeeman effect in the D_1, D_2 lines of the optical emission spectrum of the Na atom with wavelengths 589.6 nm and 589.0 nm. Indicated are only the optical transitions to the $m = +\frac{1}{2}$ level of the ground state, the polarizations $\sigma^{\pm}$ or π of the emitted photons, and their relative intensities, as calculated from the Wigner–Eckart theorem. The g_j-factors of the $^2S_{1/2}$, $^2P_{1/2}$, and $^2P_{3/2}$ states are 2, 2/3, and 4/3, respectively

The *parity selection rules* are

$$\left.\begin{array}{ll} \pi_i \pi_f = (-1)^L & \text{for electric} \\ \pi_i \pi_f = (-1)^{L+1} & \text{for magnetic} \end{array}\right\} \text{ multipole radiation.} \tag{18.32}$$

As an example, we calculate the relative line strengths of the D_1, D_2 doublet in the optical spectrum of sodium, see Fig. 18.1. The transition is from an excited $^2P_{1/2}$ or $^2P_{3/2}$ state of the Na atom to its ground state $^2S_{1/2}$, in the usual $^{2s+1}l_j$ notation with S for $l = 0$, P for $l = 1$, D for $l = 2$, etc. The two excited states have quantum numbers $s_i = \frac{1}{2}$, $l_i = 1$, and $j_i = l_i \pm s_i = 1/2, 3/2$ for spin, orbital, and total angular momentum, respectively. The spin-orbit Hamiltonian $H_{LS} = a\mathbf{L} \cdot \mathbf{S}$, (17.38), shifts the energy eigenvalues of each state by

$$E_{LS} = \frac{1}{2} a\hbar^2 \big[j(j+1) - l(l+1) - s(s+1) \big], \tag{18.33}$$

in analogy to the hyperfine case (8.13). The $^2S_{1/2}$ ground state has quantum numbers $s_f = \frac{1}{2}$, $l_f = 0$, and $j_f = \frac{1}{2}$. The emitted photons have spin quantum number $|l_i - l_f| = 1$ of dipole radiation, and energies near 2.1 eV, split by $\Delta E_{LS} = \frac{1}{2} a\hbar^2 = 2.1$ meV.

With $\pi_i = -1$ and $\pi_f = 1$, we recognize from the selection rules (18.32) that the D lines correspond to electric dipole transitions. According to (18.30), their relative transition strengths are given by $|\langle l_i m_i L M | l_f m_f \rangle|^2$. Transitions with photon magnetic quantum number $M = m_i - m_f = 0$ correspond to the absorption or emission of linearly polarized light. They are called π-transitions. Transitions with $M = \pm 1$

are induced by circularly polarized light and are called $\sigma^{\pm}$ transitions, as introduced in Sect. 4.2.

For demonstration we explicitly calculate the strength of the transition between the excited substate $|j_i m_i\rangle = |3/2, -1/2\rangle$ and the ground substate $|j_f m_f\rangle = |1/2, 1/2\rangle$. From the M selection rule (18.31) it follows that $M = 1$, and we have a σ^+ transition. The spin wave functions for the ground and excited states factorize in terms of an orbital angular momentum part $\langle l, m_l|$ and a spinor part $\langle s, m_s|$, with the appropriate Clebsch–Gordan coefficients $\langle l_f m_{lf} s_f m_{sf} | j_f m_f\rangle$ and $\langle l_i m_{li} s_i m_{si} | j_i m_i\rangle$, in the limit of low magnetic field, as

$$\langle j_f m_f| = \langle \tfrac{1}{2}, +\tfrac{1}{2}| = \langle 0, 0|\langle \tfrac{1}{2}, +\tfrac{1}{2}| = \langle l_f, m_{lf}|\langle s_f, m_{sf}|, \tag{18.34a}$$

$$|j_i m_i\rangle = |\tfrac{3}{2}, -\tfrac{1}{2}\rangle = \sqrt{\frac{2}{3}}|1, 0\rangle|\tfrac{1}{2}, -\tfrac{1}{2}\rangle + \sqrt{\frac{1}{3}}|1, -1\rangle|\tfrac{1}{2}, +\tfrac{1}{2}\rangle. \tag{18.34b}$$

Inserting these two decompositions into the matrix element $\langle j_f m_f | r Y_{11}(\theta, \varphi) | j_i m_i\rangle$, we obtain

$$\begin{aligned}\langle \tfrac{1}{2}, +\tfrac{1}{2}|rY_{11}(\theta,\varphi)|\tfrac{3}{2}, -\tfrac{1}{2}\rangle &= \sqrt{\frac{2}{3}}\langle \tfrac{1}{2}, +\tfrac{1}{2}|\tfrac{1}{2}, -\tfrac{1}{2}\rangle\langle 00|rY_{11}(\theta,\varphi)|10\rangle \\ &\quad + \sqrt{\frac{1}{3}}\langle \tfrac{1}{2}, +\tfrac{1}{2}|\tfrac{1}{2}, +\tfrac{1}{2}\rangle\langle 00|rY_{11}(\theta,\varphi)|1, -1\rangle \\ &= \sqrt{\frac{1}{3}}\langle 00|rY_{11}(\theta,\varphi)|1, -1\rangle, \end{aligned} \tag{18.35}$$

where we used the orthogonality of the spinor wave functions and the fact that $rY_{11}(\theta, \varphi)$ only acts on the spatial coordinates of orbital angular momentum. The remaining matrix element in the last equation can be evaluated by means of the Wigner–Eckart theorem to $\sqrt{\frac{1}{3}}\langle 1, -1, 1, 1|00\rangle\langle 0||rY_1(\theta, \varphi)||1\rangle$, which gives

$$\langle \tfrac{1}{2}, +\tfrac{1}{2}|rY_{11}(\theta,\varphi)|\tfrac{3}{2}, -\tfrac{1}{2}\rangle = \frac{1}{3}\langle 0||rY_1(\theta,\varphi)||1\rangle. \tag{18.36}$$

In the same way, we can calculate the matrix elements for all possible transitions. All of them are, up to a factor, proportional to the same reduced matrix element $\langle 0||rY_1(\theta, \varphi)||1\rangle$. Thus the squares of these factors correspond directly to the *relative* line intensities, as indicated in Fig. 18.1. The reduced matrix element is needed only if the *absolute* intensities are wanted. This is a really hard job to do, though, since for this purpose the electronic wave functions are needed.

The different absorption and emission properties for linearly and circularly polarized light offer the opportunity to manipulate the occupation numbers of the involved states. Let us assume that we excite Na vapor in a resonance vessel by σ^+ light at the wave length of the D_1 line. Let us further assume that the spectral width of the light is sufficiently large to cover all Zeeman sublevels. What will happen? In

the absorption process, the m quantum number is increased by one, but in the subsequent emission both σ^+ and π transitions are possible. As a consequence, the populations of the $m_i = m_f = +\frac{1}{2}$ spin up levels are increased, both in the ground and the excited state, and the populations of the spin down levels are correspondingly depleted. After some cycles, all atoms will have accumulated in the spin up states and will be fully polarized. This technique is called *optical pumping* (Cohen-Tannoudji and Kastler 1966), which essentially is based on angular momentum transfer from the photon field to the atoms.

References

Blatt, J.M., Weisskopf, V.F.: Theoretical Nuclear Physics. Wiley, New York (1952)
Cohen-Tannoudji, C., Kastler, A.: Optical pumping. Prog. Opt. **5**, 1 (1966)
Jackson, J.D.: Classical Electrodynamics. Wiley, New York (1967)
Rose, M.E.: Elementary Theory of Angular Momentum. Wiley, New York (1957)

Chapter 19
The Generalized Spin Precession Equation

Abstract We had shown that every two-level system can be mapped onto an effective spin-$\frac{1}{2}$ system, and its Hamiltonian onto an effective magnetic interaction. Furthermore, we had learned that there is a one-to-one correspondence between the precession equation for the quantum mechanical expectation values of the spin components and the classical equations of motion of the spinning top. These ideas are extended to arbitrary quantum numbers of the particles and their multipole interactions. We shall arrive at a generalized precession equation, but beyond spin-$\frac{1}{2}$ the correspondence to classical mechanics is lost. A special bra-ket notation for tensor operators is introduced that makes calculations simpler than is found in textbooks on quantum mechanics.

19.1 The Density Matrix

19.1.1 Definition of the Density Matrix

In Sect. 6.5, we had treated the 2×2 density matrix in an intuitive way. In the following, we put this on a sound basis. Starting with the Heisenberg equation for the matrix representation A of an observable,

$$\dot{\mathrm{A}} = \frac{\mathrm{i}}{\hbar}[\mathrm{H}, \mathrm{A}], \tag{19.1}$$

we derived in Sect. 10.1 the time dependent Schrödinger equation

$$\dot{\psi} = -\frac{\mathrm{i}}{\hbar}\mathrm{H}\psi. \tag{19.2}$$

In bra-ket notation, the matrix elements of A read

$$A_{nm} = \langle n|\mathrm{A}|m\rangle. \tag{19.3}$$

The elements A_{nm} may, for instance, be an integral over spatial variables of type (16.26). The set of the $|n\rangle$ may be an arbitrary orthogonal basis system. A convenient choice is to use the eigenfunctions of H, but this is not mandatory. The

D. Dubbers, H.-J. Stöckmann, *Quantum Physics: The Bottom-Up Approach*, Graduate Texts in Physics,
DOI 10.1007/978-3-642-31060-7_19, © Springer-Verlag Berlin Heidelberg 2013

Schrödinger equation (19.2) and the Heisenberg equation (19.1) are mathematically equivalent and may be mutually transformed into each other, see Sect. 10.1. Both equations suffer from the same drawback—neither wave functions nor matrix elements of observables are directly available from the experiment, they are not even unambiguously defined: For the wave function there is the arbitrariness of a phase factor, and for the matrix elements an arbitrariness of an orthogonal transformation, corresponding to a change of the basis system.

The only quantities which are really accessible are expectation values, i.e., averages either over a large ensemble, such as in NMR, where the signal is an average over typically 10^{20} nuclei, or over many measurements as, e.g., in a double slit experiment. Here the Heisenberg equation (19.1), in contrast to the Schrödinger equation, has the advantage that it can be changed into an equation for observables, as we want to prove in the following, just by taking on both sides the ensemble average,

$$\langle \dot{\mathrm{A}} \rangle = \frac{\mathrm{i}}{\hbar} \big\langle [\mathrm{H}, \mathrm{A}] \big\rangle, \tag{19.4}$$

where the expectation value $\langle \mathrm{A} \rangle$ is defined as

$$\langle \mathrm{A} \rangle = \overline{\langle \psi | \mathrm{A} | \psi \rangle}. \tag{19.5}$$

The bar denotes the ensemble average or the average over a large number of measurements, and $|\psi\rangle$ is the wave function describing the system. $|\psi\rangle$ may be expanded in terms of the basis functions,

$$|\psi\rangle(t) = \sum_n a_n(t) |n\rangle, \tag{19.6}$$

where $a_n(t)$ is the amplitude of the nth state as a function of t. Since the total probability to find the system in any state must be one, we have the normalization

$$\sum_n \big| a_n(t) \big|^2 = 1. \tag{19.7}$$

Using expansion (19.6), the expectation value (19.5) may be written as

$$\langle \mathrm{A} \rangle = \sum_{nm} \overline{a_n a_m^*} A_{nm}. \tag{19.8}$$

With

$$\rho_{nm} = \overline{a_n a_m^*}, \tag{19.9}$$

it follows from (19.5)

$$\langle \mathrm{A} \rangle = \sum_{nm} A_{nm} \rho_{mn}. \tag{19.10}$$

The sum over m is a matrix multiplication, and the sum over n is a trace, hence

$$\langle \mathrm{A} \rangle = \mathrm{Tr}(\mathrm{A}\rho). \tag{19.11}$$

Matrix ρ is called the *density matrix*. Equation (19.10) shows that the expectation value of A is nothing but a weighted average over the matrix elements A_{nm} of the corresponding operator, with the elements ρ_{mn} of ρ as weight factors, as in (19.10).

Knowledge of the density matrix ρ is sufficient to calculate the expectation values of any operator, i.e., the result of any measurement.

19.1.2 The Liouville Equation of Motion

To obtain an equation of motion for ρ, we start by inserting (19.6) with $|n\rangle = \psi_n$ into the Schrödinger equation (19.2) and obtain

$$\sum_n \dot{a}_n \psi_n = -\frac{\mathrm{i}}{\hbar} \sum_n \mathrm{H} a_n \psi_n . \tag{19.12}$$

Multiplying both side by ψ_m and making use of the orthogonality of the eigenfunctions, if applicable by integration over spatial variables, we get

$$\dot{a}_n = -\frac{\mathrm{i}}{\hbar} \sum_m H_{nm} a_m . \tag{19.13}$$

In matrix notation, the equation of motion for the evolution coefficients a_n reads

$$\dot{\mathrm{a}} = -\frac{\mathrm{i}}{\hbar} \mathrm{H a}. \tag{19.14}$$

This equation may be interpreted as a matrix representation of the Schrödinger equation.

Therein, H is the Hamiltonian matrix and $\mathrm{a} = (a_1, a_2, \ldots, a_n, \ldots)^{\mathrm{T}}$. From (19.9) and (19.13), we obtain the equation of motion for the density matrix as follows:

$$\begin{aligned} \dot{\rho}_{nm} &= \overline{\dot{a}_n a_m^*} + \overline{a_n \dot{a}_m^*} \\ &= \frac{\mathrm{i}}{\hbar} \sum_k \left(-\overline{H_{nk} a_k a_m^*} + \overline{a_n H_{mk}^* a_k^*} \right) \\ &= \frac{\mathrm{i}}{\hbar} \sum_k \left(-H_{nk} \overline{a_k a_m^*} + H_{mk}^* \overline{a_n a_k^*} \right), \end{aligned} \tag{19.15a}$$

$$\dot{\rho}_{nm} = \frac{\mathrm{i}}{\hbar} \sum_k (\rho_{nk} H_{km} - H_{nk} \rho_{km}), \tag{19.15b}$$

where we used the hermiticity of the Hamiltonian $H_{mk}^* = H_{km}$.

In matrix notation, the time evolution of the density operator is

$$\dot{\rho} = \frac{\mathrm{i}}{\hbar}[\rho, \mathrm{H}]. \tag{19.16}$$

This is the equation of motion we looked for. It is known as the *Liouville equation* and will be of central importance for everything to follow.

The trace of any commutator vanishes, $\mathrm{Tr}[\mathrm{A}, \mathrm{B}] = \mathrm{Tr}(\mathrm{AB}) - \mathrm{Tr}(\mathrm{BA}) = 0$, because of the commutativity property of the trace. Therefore, taking the trace on both sides of (19.16), we get

$$\mathrm{Tr}(\dot{\rho}) = 0, \tag{19.17a}$$

$$\mathrm{Tr}(\rho) = \mathrm{const} = 1. \tag{19.17b}$$

This means that the sum over the occupation probabilities of all states is constant and unity as it should be. We have just recovered the normalization condition (19.7).

Next we multiply both sides of (19.16) by A and take the trace,

$$\langle \dot{\mathrm{A}} \rangle = \mathrm{Tr}(\mathrm{A}\dot{\rho}) = \frac{\mathrm{i}}{\hbar}\,\mathrm{Tr}\big([\rho, \mathrm{H}]\mathrm{A}\big) = \frac{\mathrm{i}}{\hbar}\,\mathrm{Tr}\big(\rho[\mathrm{H}, \mathrm{A}]\big), \tag{19.18a}$$

or

$$\langle \dot{\mathrm{A}} \rangle = \frac{\mathrm{i}}{\hbar}\big\langle [\mathrm{H}, \mathrm{A}] \big\rangle. \tag{19.18b}$$

We recovered the Heisenberg equation (19.1), but now for the expectation values of the observables, in contrast to the matrices entering (19.1).

As an example, let us take again the precession of a magnetization in an external magnetic field. With the Hamiltonian $\mathrm{H_M} = -\boldsymbol{\mu} \cdot \boldsymbol{B} = \gamma \mathbf{J} \cdot \boldsymbol{B}$, we obtain from (19.4), taking for A the components of $\mathbf{J}$,

$$\langle \dot{\mathbf{J}} \rangle = \frac{\mathrm{i}}{\hbar}\gamma \big\langle [\mathbf{J} \cdot \boldsymbol{B}, \mathbf{J}] \big\rangle. \tag{19.19}$$

With $[\mathbf{J}, \mathbf{J}] = \mathrm{i}\hbar\mathbf{J}$ from (16.16), this can be written in the same way as in Sect. 10.3,

$$\langle \dot{\mathbf{J}} \rangle = \gamma \langle \mathbf{J} \times \boldsymbol{B} \rangle, \tag{19.20a}$$

$$\dot{M} = \gamma M \times \mathbf{B}, \tag{19.20b}$$

with the magnetization $\boldsymbol{M} = -n\gamma \langle \mathbf{J} \rangle$ from (6.24).

We recovered the classical equation of motion describing the precession of a spin polarization in an external magnetic field. For a magnetic interaction, there is no difference between the classical and the quantum mechanical treatment. For spin-$\frac{1}{2}$ systems, this equivalence was shown in Sect. 10.3, but nowhere in the present derivation did the value of the angular momentum quantum number enter explicitly. The statement is therefore true for arbitrary angular momenta.

19.2 Some Preparative Steps

19.2.1 Normalized Irreducible Tensor Operators

Next we are going to expand the density matrix in terms of irreducible tensor operators. To this end, a number of preparatory steps are needed.

Like for any operator A, the trace of a tensor operator T_{LM} is the sum over its diagonal matrix elements

$$\mathrm{Tr}(\mathrm{T}_{LM}) = \sum_n \langle n|\mathrm{T}_{LM}|n\rangle \tag{19.21}$$

with respect to the basis states $|n\rangle$. The irreducible tensor operators are orthogonal with respect to the trace operation,

$$\mathrm{Tr}\big(\mathrm{T}^{\dagger}_{L'M'}\mathrm{T}_{LM}\big) = \delta_{LL'}\delta_{MM'}\,\mathrm{Tr}\,|\mathrm{T}_L|^2, \tag{19.22}$$

where $\mathrm{Tr}\,|\mathrm{T}_L|^2$ is independent of the numbers M, M', as it is indicated by the notation. The proof of this statement follows essentially the same steps as the corresponding proof for the irreducible tensor functions. Since it is somewhat technical, it is transferred to Appendix A.1. If in addition the condition $\mathrm{Tr}\,|\mathrm{T}_L|^2 = 1$ holds, the tensor operators are called *normalized*. We introduce hatches to characterize normalized operators,

$$\hat{\mathrm{T}}_{LM} = \frac{\mathrm{T}_{LM}}{\sqrt{\mathrm{Tr}\,|\mathrm{T}_L|^2}}. \tag{19.23}$$

From (19.22) one obtains, in particular, for $L' = 0$

$$\mathrm{Tr}(\mathrm{T}_{LM}) = 0 \quad \text{for } L \neq 0. \tag{19.24}$$

Hence all irreducible tensor operators, with the exception of scalars, are traceless. We have seen in the preceding chapter that all interactions of atoms or nuclei with external fields can be expressed in terms of irreducible tensor operators. This has the consequence that, though degeneracies of eigenvalues are lifted by the interactions, the *center of gravity* of the state is not changed. This is, in particular, true for the magnetic interaction between a magnetic moment and an external magnetic field,

as well as for the electric interaction between a nuclear quadrupole moment and an external electric field gradient; see Sects. 18.1 and 18.2.

A quantity repeatedly needed is the matrix representation of the normalized irreducible tensor operators $\hat{\mathrm{T}}_{LM}(j)$ in the $|jm\rangle$ basis. For a given angular momentum quantum number j, this representation is $(2j+1)$-dimensional with matrix elements $\langle jm'|\hat{\mathrm{T}}_{LM}(j)|jm\rangle$. The Wigner–Eckart theorem (17.48) yields for this quantity

$$\langle jm'|\hat{\mathrm{T}}_{LM}(j)|jm\rangle = \frac{1}{\sqrt{2j+1}}\langle j||\hat{\mathrm{T}}_L(j)||j\rangle\langle jmLM|jm'\rangle. \tag{19.25}$$

The reduced matrix element is

$$\langle j||\hat{\mathrm{T}}_L(j)||j\rangle = \sqrt{2L+1}. \tag{19.26}$$

The derivation is given in Appendix A.2. The desired matrix representation of the normalized tensors thus is

$$\langle jm'|\hat{\mathrm{T}}_{LM}(j)|jm\rangle = \sqrt{\frac{2L+1}{2j+1}}\langle jmLM|jm'\rangle. \tag{19.27}$$

The list of normalized irreducible tensor elements with $L \leq 2$ is given in Appendix A.7. All tensors elements of type $\hat{\mathrm{T}}_{LM}(A)$ for arbitrary vector operator A discussed in this book appear on this list.

19.2.2 A Bra-Ket Notation for Tensor Operators

The very close similarity between the definition of the irreducible tensor operators $\hat{\mathrm{T}}$ of (17.10a), (17.10b),

$$[\mathrm{L}_z, \hat{\mathrm{T}}_{LM}] = \hbar M \hat{\mathrm{T}}_{LM}, \tag{19.28a}$$

$$[\mathrm{L}_\pm, \hat{\mathrm{T}}_{LM}] = \hbar\sqrt{L(L+1) - M(M\pm 1)}\,\hat{\mathrm{T}}_{L,M\pm 1}, \tag{19.28b}$$

and the analogous relation for the tensor functions $|l, m\rangle$, see (16.20a) and (16.25),

$$\mathrm{L}_z|lm\rangle = \hbar m|lm\rangle, \tag{19.29a}$$

$$\mathrm{L}_\pm|lm\rangle = \hbar\sqrt{l(l+1) - m(m\pm 1)}\,|l, m\pm 1\rangle, \tag{19.29b}$$

can be made even more suggestive by introducing a new notation for the tensor operators. To this end, we define tensor bras and kets according to

$$|LM) = \hat{\mathrm{T}}_{LM}(j), \tag{19.30a}$$

$$(LM| = \hat{\mathrm{T}}^\dagger_{LM}(j). \tag{19.30b}$$

At the same time, we assume that an operator A acts on the tensor bras and kets according to

$$\mathrm{A}|LM) = \big[\mathrm{A}, \hat{\mathrm{T}}_{LM}(j)\big], \tag{19.31a}$$

$$(LM|\mathrm{A} = \big[\hat{\mathrm{T}}^{\dagger}_{LM}(j), \mathrm{A}\big]. \tag{19.31b}$$

With these definitions (19.28a), (19.28b) and (19.29a), (19.29b) become identical even with respect to notation. One only has to replace the bras and kets in (19.29a), (19.29b) by the corresponding tensor bras and kets to recover (19.28a), (19.28b) for the tensor operators.

We define the tensor matrix element of an operator A as

$$\begin{aligned}(L'M'|\mathrm{A}|LM) &= \mathrm{Tr}\big\{\hat{\mathrm{T}}^{\dagger}_{L'M'}(j)\big[\mathrm{A}, \hat{\mathrm{T}}_{LM}(j)\big]\big\} \\ &= \mathrm{Tr}\big\{\big[\hat{\mathrm{T}}^{\dagger}_{L'M'}(j), \mathrm{A}\big]\hat{\mathrm{T}}_{LM}(j)\big\},\end{aligned} \tag{19.32}$$

where we have taken the trace (19.21) over the basis states $|j, m\rangle$. The last identity holds because of the commutativity property of the trace. It is thus irrelevant whether operator A acts on the bra or on the ket. The orthogonality relation (19.22) for the tensor operators may now be expressed as

$$(LM|L'M') = \delta_{LL'}\delta_{MM'}. \tag{19.33}$$

From the completeness relation for the irreducible tensor operators follows

$$\sum_{LM}(L_1M_1|\mathrm{A}|LM)(LM|\mathrm{B}|L_2M_2) = (L_1M_1|\mathrm{AB}|L_2M_2). \tag{19.34}$$

In the interpretation of the product AB, care has to be taken. This can be seen if we follow the steps leading to (19.34) in detail, by writing down all matrix elements from (19.32) explicitly:

$$\begin{aligned}&\sum_{LM}(L_1M_1|\mathrm{A}|LM)(LM|\mathrm{B}|L_2M_2) \\ &\quad= \sum_{LM}\mathrm{Tr}\big\{\big[\hat{\mathrm{T}}^{\dagger}_{L_1M_1}(j), \mathrm{A}\big]\hat{\mathrm{T}}_{LM}(j)\big\}\,\mathrm{Tr}\big\{\hat{\mathrm{T}}^{\dagger}_{LM}(j)\big[\mathrm{B}, \hat{\mathrm{T}}_{L_2M_2}(j)\big]\big\} \\ &\quad= \sum_{LM\mu_1\mu_2\mu_3\mu_4}\langle j\mu_1|\big[\hat{\mathrm{T}}^{\dagger}_{L_1M_1}(j), \mathrm{A}\big]|j\mu_2\rangle\langle j\mu_2|\hat{\mathrm{T}}_{LM}(j)|j\mu_1\rangle \\ &\qquad\times\langle j\mu_3|\hat{\mathrm{T}}^{\dagger}_{LM}(j)|j\mu_4\rangle\langle j\mu_4|\big[\mathrm{B}, \hat{\mathrm{T}}_{L_2M_2}(j)\big]|j\mu_3\rangle \\ &\quad= \mathrm{Tr}\big\{\big[\hat{\mathrm{T}}^{\dagger}_{L_1M_1}(j), \mathrm{A}\big]\big[\mathrm{B}, \hat{\mathrm{T}}_{L_2M_2}(j)\big]\big\} \\ &\quad= \mathrm{Tr}\big\{\hat{\mathrm{T}}^{\dagger}_{L_1M_1}(j)\big[\mathrm{A}, \big[\mathrm{B}, \hat{\mathrm{T}}_{L_2M_2}(j)\big]\big]\big\}.\end{aligned} \tag{19.35}$$

In the third step, we used the equality

$$\sum_{LM}\langle jm_1|\hat{\mathrm{T}}^{\dagger}_{LM}(j)|jm_2\rangle\langle jm_2'|\hat{\mathrm{T}}_{LM}(j)|jm_1'\rangle = \delta_{m_1m_1'}\delta_{m_2m_2'} \tag{19.36}$$

that follows from the Wigner–Eckart theorem (17.48) with (19.16) for the irreducible tensor operators and the orthogonality of the Clebsch–Gordan coefficients, and in the last steps we used again (19.32).

Comparison of the last equation (19.35) with (19.34) shows that the application of the product AB on $|LM)$ does not mean $[\mathrm{AB}, \hat{\mathrm{T}}_{LM}(j)]$, but has to be interpreted as the successive application of this expression, namely

$$\mathrm{AB}|LM) = \big[\mathrm{A}, \big[\mathrm{B}, \hat{\mathrm{T}}_{LM}(j)\big]\big]. \tag{19.37}$$

To illustrate the usefulness of this definition, the action of an exponential operator onto a tensor ket $\mathrm{e}^{\mathrm{A}}|LM)$ shall be calculated. Expanding the exponential into its Taylor series, one obtains in a number of steps

$$\begin{aligned}
\mathrm{e}^{\mathrm{A}}|LM) &= \sum_{n=0}^{\infty} \frac{1}{n!} \mathrm{A}^n |LM) \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \big[\underbrace{\mathrm{A}, [\mathrm{A}, \ldots,}_{n \text{ times } [\mathrm{A},} \hat{\mathrm{T}}_{LM} \underbrace{\,]\ldots]}_{n \text{ times}]} \\
&= \sum_{n,m} \frac{1}{n!} \mathrm{A}^n \hat{\mathrm{T}}_{LM} \frac{1}{m!} (-\mathrm{A})^m .
\end{aligned} \tag{19.38}$$

The equivalence of the last two expressions is easily checked by a term-by-term inspection of the powers of A in both expressions. We have thus arrived at the result

$$\mathrm{e}^{\mathrm{A}}|LM) = \mathrm{e}^{\mathrm{A}} \hat{\mathrm{T}}_{LM} \mathrm{e}^{-\mathrm{A}}. \tag{19.39}$$

The action of an exponential operator onto a tensor ket thus means nothing but sandwiching the corresponding tensor operator with the exponential operator!

Because of the one-to-one correspondence of (19.28a), (19.28b) and (19.29a), (19.29b), each proof for the tensor functions is automatically true for the tensor operators, too, and vice versa. In particular, we can apply directly the Wigner–Eckart theorem to calculate the matrix element with tensor bra-kets (see (19.32)) for the case that A is a normalized irreducible tensor operator

$$(LM|\hat{\mathrm{T}}_{L_2 M_2}(j)|L_1 M_1) = \frac{1}{\sqrt{2L+1}} (L||\hat{\mathrm{T}}_{L_2}(j)||L_1) \langle L_1 M_1 L_2 M_2 | LM \rangle. \tag{19.40}$$

The calculation of the reduced matrix element is lengthy and is given in Appendix A.3. The result is

$$\begin{aligned}
(L||\hat{\mathrm{T}}_{L_1}(j)||L_2) = \big[(-1)^{L_1+L_2-L} - 1\big](-1)^{2j+L} \\
\times \sqrt{(2L_1+1)(2L_2+1)(2L+1)} \begin{Bmatrix} L_1 & L_2 & L \\ j & j & j \end{Bmatrix}.
\end{aligned} \tag{19.41}$$

Values of the reduced matrix elements for $L = 1$ and 2 are given in Appendix A.8.

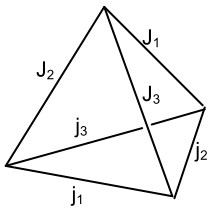


Fig. 19.1 Tetrahedron to visualize the symmetries and triangular relations of the $6j$ symbol

The quantity in curly brackets is called a $6j$ *symbol*. A $6j$ symbol is invariant with respect to the exchange of two columns, as well as the simultaneous exchange of any two quantum numbers in the upper row (call them J_1, J_2, J_3) with the two numbers below them in the lower row (call them j_1, j_2, j_3). Furthermore, these entries obey the triangular relations

$$\triangle(J_1 J_2 J_3), \qquad \triangle(J_1 j_2 j_3), \qquad \triangle(j_1 J_2 j_3), \qquad \triangle(j_1 j_2 J_3). \tag{19.42}$$

Symmetry and triangular relations may be visualized by means of a tetrahedron formed of the six angular momenta (see Fig. 19.1). The symmetry relations correspond to rotations and reflections of the tetrahedron, and the triangular relations mean that the tetrahedron can be geometrically constructed by the six vectors $j_1, \ldots, J_3$. More details on the $6j$ symbols are given in standard monographs on angular momenta in quantum mechanics; see, e.g., the books by Edmonds (1957), Fano and Racah (1959), Rose (1957), Thompson (2004), and Blum (1996).

There are numerous selection rules resulting from relations (19.42). The perhaps most important one, $L \leq 2j$, was met already in the context of the Wigner Eckart theorem: for a tensor of rank L to exist, the angular momentum state of the system under study must at least be $L/2$.

19.3 The Irreducible Components of the Density Matrix

19.3.1 Definition of the Statistical Tensors

Since the number of possible states in an atom or nucleus is infinite, the rank of the density matrix strictly speaking, too, is unlimited. However, often only a small number of states are relevant. In nuclear magnetic or in electron spin resonance, for instance, usually only one angular momentum state j with its m substates is involved, leading to rank $2j + 1$ of the corresponding density submatrix. As long as all other states are so far away in energy that their coupling to the state in question can be neglected, as is mostly the case for light or radiofrequency irradiation at or near resonance, we may restrict discussion to these $2j + 1$ states $|jm\rangle$. We denote the corresponding density matrix by $\rho(j)$.

We expand the density matrix ρ in terms of normalized irreducible tensor operators

$$\rho(j) = \sum_{LM} \rho_{LM} \hat{\mathrm{T}}^{\dagger}_{LM}(j). \tag{19.43}$$

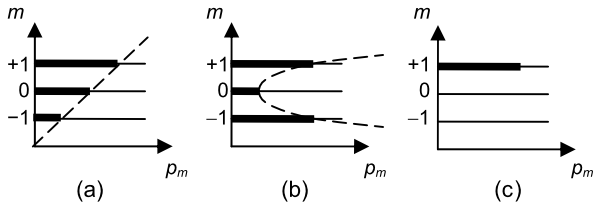


Fig. 19.2 Spin-1 levels (**a**) with a pure longitudinal vector polarization $\rho_{10} = (p_1 - p_{-1})/\sqrt{3}$, with population p_m linear in magnetic quantum number m; (**b**) with a pure longitudinal alignment $\rho_{20} = (1 - 3p_0)/\sqrt{6}$, with p_m quadratic in m; (**c**) with a mixture of both

It is easy to show that the $\hat{\mathrm{T}}_{LM}(j)$ form a complete set. Since L is restricted to $L \leq 2j$, and since for each rank L there are $2L + 1$ components, the total number of terms on the right-hand side of (19.43) is $\sum_{L=0}^{2j}(2L + 1) = (2j + 1)^2$. This is exactly the number of independent components of ρ. Equation (19.43) expresses the $(2j + 1)^2$ matrix elements of ρ in terms of the $(2j + 1)^2$ tensor strengths ρ_{LM}.

We invert this equation by multiplying both sides of (19.43) by $\hat{\mathrm{T}}_{LM}$, taking the trace, and using the orthogonality (19.22) and (19.23) of the tensor operator, and obtain

$$\rho_{LM} = \mathrm{Tr}(\rho \hat{\mathrm{T}}_{LM}). \tag{19.44}$$

With (19.11) we see:

For a given system, the components ρ_{LM} of the density matrix in irreducible representation, frequently called *statistical tensors*, are the expectation values of the normalized tensor operators

$$\rho_{LM} = \langle \hat{\mathrm{T}}_{LM} \rangle. \tag{19.45}$$

Again, for a given rank L we can assemble the elements ρ_{LM} into one column that we call ρ_L. We call a statistical tensor ρ_1 of rank 1 a (*vector-*)*polarization*, and ρ_2 of rank 2 an *alignment*. The notation is not uniform in the literature. Atomic physicists call the ρ_L *state multipoles* and use the name *orientation* for ρ_1.

19.3.2 Simple Examples of Statistical Tensors

For the purpose of illustration, let us consider three particularly simple cases frequently found in experiments. To keep the discussion simple, we take the simplest situation where all these cases listed below can be realized, namely, an ensemble of particles with angular momentum $j = 1$ with substates $m = 0, \pm 1$.

1. Linear occupation of the substates (see Fig. 19.2(a)).
   Here the occupation probabilities p_m for the three substates are

$$p_1 = \frac{1}{3} + a, \quad p_0 = \frac{1}{3}, \quad p_{-1} = \frac{1}{3} - a. \tag{19.46}$$

   This holds for a thermal Boltzmann distribution at high temperature $k_B T \gg E$ (see the end of Sect. 6.4), and after various polarizing nuclear reactions. The resulting irreducible density matrix is diagonal and may be expressed as

$$\rho = \frac{1}{3}\mathrm{I} + a\mathrm{J}_z/\hbar = \frac{1}{\sqrt{3}}\hat{\mathrm{T}}_{00} + a\sqrt{2}\,\hat{\mathrm{T}}_{10}, \tag{19.47}$$

   with identity I, and with $\hat{\mathrm{T}}_{LM}$ from Appendix A.7. Thus ρ contains the irreducible components

$$\rho_{00} = \langle \hat{\mathrm{T}}_{00} \rangle = \sqrt{\frac{1}{3}}, \tag{19.48a}$$

$$\rho_{10} = \langle \hat{\mathrm{T}}_{10} \rangle = a\sqrt{2} = \sqrt{\frac{1}{2}}(p_{+1} - p_{-1}). \tag{19.48b}$$

   All other components are zero. ρ_{10} is related to our previously defined longitudinal polarization $P_z = \langle \mathrm{J}_z \rangle/\hbar$ as $\rho_{10} = P_z/\sqrt{2}$.
2. Quadratic occupation (see Fig. 19.2(b)).
   Here the occupation numbers are

$$p_1 = \frac{1}{3} + a, \quad p_0 = \frac{1}{3} - 2a, \quad p_{-1} = \frac{1}{3} + a, \tag{19.49}$$

   yielding for the density matrix

$$\rho = \frac{1}{3}\mathrm{I} + a\left(3\mathrm{J}_z^2 - \mathrm{J}^2\right)/\hbar^2 = \sqrt{\frac{1}{3}}\hat{\mathrm{T}}_{00} + a\sqrt{6}\,\hat{\mathrm{T}}_{20}. \tag{19.50}$$

   Now the only nonvanishing components of the density matrix are

$$\rho_{00} = \sqrt{\frac{1}{3}}, \tag{19.51a}$$

$$\rho_{20} = a\sqrt{6} = \sqrt{\frac{1}{6}}(1 - 3p_0). \tag{19.51b}$$

3. Only one state is occupied (see Fig. 19.2(c)).
   Here the occupation numbers are, for example,

$$p_1 = 1, \quad p_0 = 0, \quad p_{-1} = 0. \tag{19.52}$$

   This is the situation found, for instance, in optical pumping with circularly polarized light (see the end of Sect. 18.3). Here the only nonvanishing irreducible components are ρ_{00}, ρ_{10}, and ρ_{20}.

19.4 The Liouville Equation for the Statistical Tensors

The Liouville equation (19.16) is a coupled differential equation system for the $(2j+1)^2$ components of the density matrix. Already for moderate values of j a tremendous effort is required to solve it by brute force methods. Here the irreducible tensors exhibit their full potential. The derivation presented here follows essentially Fano (1964), see also Dubbers (1976). Consequent use of the tensor operator matrix elements introduced in Sect. 19.2.2 (see (19.32)) and of the Wigner–Eckart theorem, though, makes the derivation much more straightforward than in the mentioned textbooks on angular momentum.

We start by taking the time derivative on both sides of (19.44) and inserting (19.16)

$$\dot{\rho}_{LM} = \mathrm{Tr}(\dot{\rho}\hat{\mathrm{T}}_{LM}) = \frac{\mathrm{i}}{\hbar}\,\mathrm{Tr}\big([\rho, \mathrm{H}]\hat{\mathrm{T}}_{LM}\big). \tag{19.53}$$

Inserting the irreducible expansion (19.43) of the density matrix on the right-hand side gives

$$\dot{\rho}_{LM} = \frac{\mathrm{i}}{\hbar}\sum_{L_2M_2}\rho_{L_2M_2}\,\mathrm{Tr}\big(\big[\hat{\mathrm{T}}^{\dagger}_{L_2M_2}, \mathrm{H}\big]\hat{\mathrm{T}}_{LM}\big), \tag{19.54}$$

or, using the bra-ket notation of (19.32),

$$\dot{\rho}_{LM} = \frac{\mathrm{i}}{\hbar}\sum_{L_2M_2}\rho_{L_2M_2}(L_2M_2|\mathrm{H}|LM). \tag{19.55}$$

We then expand H in terms of normalized irreducible tensor operators,

$$\mathrm{H} = \hbar\sum_{L_1M_1}\Omega_{L_1M_1}\hat{\mathrm{T}}^{\dagger}_{L_1M_1}(j). \tag{19.56}$$

In Sects. 18.1 and 18.2, it had been stated that in nearly all practical applications the sum over L_1 is restricted to $L_1 = 1$ (magnetic dipole interaction) and $L_1 = 2$ (electric quadrupole interaction). In the next chapter, we shall see how the irreducible interaction frequencies $\Omega_{L_1M_1}$ are connected to the dipole and quadrupole interaction parameters developed in earlier chapters. The dipole–dipole interaction, with H depending on two angular momenta j_1 and j_2, will not be treated here. Inserting expression (19.56) into (19.55) yields

$$\dot{\rho}_{LM} = \mathrm{i}\sum_{L_1M_1L_2M_2}\Omega_{L_1M_1}\rho_{L_2M_2}(L_2M_2|\hat{\mathrm{T}}^{\dagger}_{L_1M_1}(j)|LM). \tag{19.57}$$

Using the Wigner–Eckart theorem (19.40), the tensor matrix element is evaluated to

$$\begin{aligned}(L_2M_2|\hat{\mathrm{T}}^{\dagger}_{L_1M_1}(j)|LM) &= (LM|\hat{\mathrm{T}}_{L_1M_1}(j)|L_2M_2)^*\\ &= \frac{1}{\sqrt{2L+1}}(L||\hat{\mathrm{T}}_{L_1}(j)||L_2)\langle L_2M_2L_1M_1|LM\rangle.\end{aligned} \tag{19.58}$$

It follows

$$\dot{\rho}_{LM} = \mathrm{i}\frac{1}{\sqrt{2L+1}} \sum_{L_1 L_2} (L||\hat{\mathrm{T}}_{L_1}(j)||L_2) \sum_{M_1 M_2} \Omega_{L_1 M_1} \rho_{L_2 M_2} \langle L_2 M_2 L_1 M_1 | L M \rangle. \tag{19.59}$$

Using the compact notation (17.31) for the tensor product, we obtain the *generalized spin precession equation* for the statistical tensors:

$$\dot{\rho}_L = \mathrm{i} \sum_{L_1 L_2} c_j(L_1 L_2 L)[\Omega_{L_1} \times \rho_{L_2}]_L. \tag{19.60}$$

The coefficients therein are

$$c_j(L_1 L_2 L) = -\frac{1}{\sqrt{2L+1}}(L||\hat{\mathrm{T}}_{L_1}(j)||L_2), \tag{19.61}$$

with $(L||\mathrm{T}_{L_2}(j)||L_1)$ taken from (19.41) and using (17.32). The most frequently used $c_j(L_1 L_2 L)$ are given in Appendix A.8.

From the selection rules for the Clebsch–Gordan coefficients and the $6j$ symbols and from the factor $[1-(-1)^{L_1+L_2-L}]$ in (19.41), we deduce that $\dot{\rho}_{LM}$ is coupled by $\Omega_{L_1 M_1}$ to $\rho_{L_2 M_2}$ only if

$$L_1 + L_2 - L \text{ odd}; \tag{19.62a}$$

$$M_1 + M_2 = M; \tag{19.62b}$$

$$\Delta(L_1 L_2 L); \tag{19.62c}$$

$$L_1, L_2, L \leq 2j. \tag{19.62d}$$

These far reaching selection rules are the harvest of the irreducible tensor formalism. In many cases, the original coupled differential equation system for the $(2j+1)^2$ matrix elements of the density matrix ρ will separate into a number of decoupled differential equations for the irreducible components of ρ, each of them much easier to solve than for instance in the angular momentum representation of (19.15b) with its multiple summations over unobserved substates. Note that, though the derivation in preceding sections is demanding for readers not familiar with the tensor formalism, the final equation (19.60) has a simple structure well suited for practical applications. Examples will be given in the chapters to follow.

References

Blum, K.: Density Matrix Theory and Applications, 2nd edn. Plenum Press, New York (1996)

Dubbers, D.: Nuclear reorientation in static and radio-frequency electro-magnetic fields. Z. Phys. A **276**, 245–259 (1976)
Edmonds, A.R.: Angular Momentum in Quantum Mechanics. Princeton University Press, Princeton (1957)
Fano, U.: Precession equation of a spinning particle in nonuniform fields. Phys. Rev. **133**, B828–B830 (1964)
Fano, U., Racah, G.: Irreducible Tensorial Sets. Academic Press, New York (1959)
Rose, M.E.: Elementary Theory of Angular Momentum. Wiley, New York (1957)
Thompson, W.J.: Angular Momentum. Wiley-VCH, Weinheim (2004)

Chapter 20
Reorientation in Static Electromagnetic Fields

Abstract Several applications of the generalized precession equation are presented. First, the magnetic spin precession equation is extended to arbitrary ranks of the polarization tensor. Then it is shown that a transverse electric quadrupole interaction always couples polarization tensors of different rank. A special case of a mixed magnetic dipole–electric quadrupole interaction is solved. To arrive at the time average solution one only needs to invert a matrix of coupling coefficients.

20.1 Magnetic Dipole Precession

For illustration we start with the example discussed in Sect. 10.3 and at the end of Sect. 19.1, namely, the precession of a nuclear ensemble of spin **I** under the influence of the magnetic dipole Hamiltonian (2.9)

$$\mathrm{H}_{\mathrm{magn}} = -\frac{\mu}{\hbar I}\mathbf{I}\cdot\boldsymbol{B} = -\frac{\mu}{I}\sum_{M_1}\mathrm{T}^{\dagger}_{1M_1}(I)\mathrm{T}_{1M_1}(B), \tag{20.1}$$

or, using the notation of (19.56),

$$\mathrm{H}_{\mathrm{magn}} = \hbar\sum_{M_1}\Omega_{1M_1}\hat{\mathrm{T}}^{\dagger}_{1M_1}(I), \tag{20.2}$$

with

$$\Omega_{1M_1} = -\frac{\mu}{\hbar I}\sqrt{\frac{1}{3}I(I+1)(2I+1)}\mathrm{T}_{1M_1}(B), \tag{20.3}$$

and with $\mathrm{T}_{1M}(B)$ taken from (17.36a), (17.36b). To better keep track, we use the same summation index M_1 as in (19.56). In the simplest case $I = 1$, we have $\Omega_{1M} = -\sqrt{2}(\mu/\hbar)B_M$, i.e., $\Omega_{10} = -\sqrt{2}(\mu/\hbar)B_z$, and $\Omega_{1\pm1} = \mp(\mu/\hbar)(B_x \pm \mathrm{i}B_y)$.

Because of the selection rules (19.62a)–(19.62d), the sum (19.60) contains only one term

$$\dot{\rho}_L = \mathrm{i}c_I(1LL)[\Omega_1 \times \rho_L]_L. \tag{20.4}$$

The tensor cross product on the right is reduced, up to a factor, to the ordinary vector product, as was shown in (17.43). We have generalized the $L = 1$ magnetic

D. Dubbers, H.-J. Stöckmann, *Quantum Physics: The Bottom-Up Approach*, Graduate Texts in Physics,
DOI 10.1007/978-3-642-31060-7_20, © Springer-Verlag Berlin Heidelberg 2013

precession equation

$$\dot{\boldsymbol{P}} = \boldsymbol{P} \times \boldsymbol{\omega}_{\mathrm{L}} \tag{20.5}$$

derived in a more direct approach before, see (19.20a) or (10.34b), to arbitrary degrees L of polarization, and recovered the special $L = 1$ case from the general case (20.4), which reads explicitly

$$\dot{\rho}_{LM} = \mathrm{i}\frac{g\mu_{\mathrm{B}}}{\hbar}\Bigg[-\sqrt{\frac{1}{2}\big[L(L+1) - M(M-1)\big]}B_{+1}\rho_{L,M-1} + MB_0\rho_{L,M} + \sqrt{\frac{1}{2}\big[L(L+1) - M(M+1)\big]}B_{-1}\rho_{L,M+1}\Bigg]. \tag{20.6}$$

Therein L runs from 0 to $2I$, and we used $\mu/\hbar I = g\mu_{\mathrm{B}}/\hbar$ from (2.7) to emphasize that this magnetic precession equation is independent of spin I.

We recognize the selection rules for the ranks L, M of polarization

$$\Delta L = 0, \qquad \Delta M = 0, \pm 1 \quad \text{for magnetic interactions,} \tag{20.7}$$

which follow directly from the corresponding selection rules of the Clebsch–Gordan coefficients.

Instead of one original system of equations (19.16) for the $(2I+1)^2$ Cartesian components of the density operator ρ, we have $2I+1$ separate equation systems for each degree L of the statistical tensors. For the special case $B_1 = B_{-1} = 0$ for which the magnetic field is parallel to the z-axis, the system is even completely decoupled,

$$\dot{\rho}_{LM} = \mathrm{i}M\gamma_1 B_0 \rho_{LM} = \mathrm{i}M\omega_0\rho_{LM}. \tag{20.8}$$

This holds for arbitrary angular momenta. These equations can be integrated immediately

$$\rho_{LM}(t) = \mathrm{e}^{\mathrm{i}M\omega_0 t}\rho_{LM}(0). \tag{20.9}$$

We see that each tensor component ρ_{LM} precesses with the angular frequency $M\omega_0$ about the z-axis, where $\omega_0 = \gamma_I B_0$ is the Larmor angular frequency. Hence Larmor precession in an external magnetic field takes place in the same way for all values of spin and for all degrees of tensor polarization, and is equivalent to a rotation operation of the entire system.

20.2 Electric Quadrupole Reorientation

Next we discuss the reorientation of a nucleus with spin $I > \frac{1}{2}$ and nonvanishing quadrupole moment Q in an electric field gradient (in zero magnetic field). The

Hamiltonian (18.27) for the quadrupole interaction may be rewritten as

$$\mathrm{H}_Q = \hbar \sum_{M_1} \Omega_{2M_1} \hat{\mathrm{T}}^{\dagger}_{2M_1}(I), \quad \text{with} \tag{20.10}$$

$$\Omega_{2M} = \frac{eQ}{\hbar} \sqrt{\frac{1}{20} \frac{(2I+1)(2I+2)(2I+3)}{2I(2I-1)}} q_M \tag{20.11}$$

using a_2 from (A.35) and q_M from (18.20a)–(18.20c). In this case, the L selection rules reduce the sum in the equation system (19.60) to two terms,

$$\dot{\rho}_L = \mathrm{i}\big\{c_j(2, L-1, L)[\Omega_2 \times \rho_{L-1}]_L + c_j(2, L+1, L)[\Omega_2 \times \rho_{L+1}]_L\big\}. \tag{20.12}$$

This illustrates the selection rules for the degrees L, M of polarization

$$\Delta L = 1, \qquad \Delta M = 0, \pm 1, \pm 2 \quad \text{for electric interactions.} \tag{20.13}$$

20.3 Reorientation in Mixed Magnetic and Electric Fields

Let us next consider the case of a combined magnetic dipole and electric quadrupole interaction for the special case $I = 1$. The precession equations (19.16) read, with $c_1(111) = -1$, $c_1(122) = -\sqrt{3}$, $c_1(221) = \sqrt{5}$, and $c_1(212) = -\sqrt{3}$ from Appendix A.8,

$$\dot{\rho}_1 = -\mathrm{i}[\Omega_1 \times \rho_1]_1 + \mathrm{i}\sqrt{5}[\Omega_2 \times \rho_2]_1, \tag{20.14a}$$

$$\dot{\rho}_2 = -\mathrm{i}\sqrt{3}[\Omega_2 \times \rho_1]_2 - \mathrm{i}\sqrt{3}[\Omega_1 \times \rho_2]_2. \tag{20.14b}$$

For those readers who prefer to write out matrices, we combine the three components of polarization ρ_1 and the five components of alignment ρ_2 into one column vector and write

$$\begin{pmatrix} \dot{\rho}_1 \\ \dot{\rho}_2 \end{pmatrix} = \begin{pmatrix} \mathrm{M}_{11} & \mathrm{Q}_{12} \\ \mathrm{Q}_{21} & \mathrm{M}_{22} \end{pmatrix} \begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix}, \tag{20.15}$$

where the submatrices M_{11} and M_{22} have purely magnetic components linear in the frequencies Ω_{1M}, and the submatrices Q_{12} and Q_{21} have purely quadrupolar components linear in the frequencies Ω_{2M}.

In accordance with rule (19.62a), only an electric quadrupole interaction can change an initial polarization $\rho_1(0)$ into an alignment $\rho_2(t)$ (and vice versa). In contrast, a magnetic dipole interaction merely rotates tensor components of the same rank.

We explicitly study the case of a magnetic field B_0 parallel to the z-axis, and a nonaxially symmetric efg tensor with one main axis also parallel to z, that is, using (18.18) and (7.8), with Hamiltonian

$$\mathrm{H} = \omega_0 \mathrm{I}_z + \frac{\omega_\mathrm{Q}}{\hbar}\left[3\mathrm{I}_z^2 - \mathbf{I}^2 + \frac{1}{2}\eta\left(\mathrm{I}_+^2 + \mathrm{I}_-^2\right)\right]. \tag{20.16}$$

Consider again the easiest case $I = 1$. The corresponding quadrupole interaction frequencies are $\Omega_{20} = \sqrt{6}\omega_\mathrm{Q}$, $\Omega_{2\pm1} = 0$, $\Omega_{2\pm2} = \eta\omega_\mathrm{Q}$. Let us further assume that in the beginning there was only a dipolar polarization $\rho_{10}(0) \neq 0$. We then obtain the following system of differential equations as a special case of (20.14a), (20.14b),

$$\dot{\rho}_{10} = -\mathrm{i}\eta\sqrt{2}\omega_\mathrm{Q}(\rho_{22} - \rho_{2-2}), \tag{20.17a}$$

$$\dot{\rho}_{2\pm2} = \mp\mathrm{i}(-2\omega_0\rho_{2\pm2} + \eta\sqrt{2}\omega_\mathrm{Q}\rho_{10}). \tag{20.17b}$$

This is a closed system of differential equations for $\rho_{10}, \rho_{22}, \rho_{2-2}$. The equation systems for the other components of ρ are not of interest since we assumed that in the beginning only ρ_{10} differs from zero. In the general case, ρ_{2M} is also coupled to $\rho_{1M}, \rho_{3M}, \rho_{4M}$, etc. For $I = 1$, however, these tensor components vanish due to the selection rule $L \leq 2I$.

Instead of the original Liouville equation with its $(2I + 1)^2 = 9$ variables, we are left with an equation system with only three variables. But one has to admit that for the case of arbitrary orientations of the efg tensor this advantage is lost. In this case, each component of ρ is coupled to every other by the combined magnetic plus electric interaction, with the only exception of ρ_{00}, and there is no longer a technical advantage with the tensor formalism.

20.4 Time Average Results

Often the transients of the polarization are not directly accessible but only their time averages

$$\bar{\rho}_{LM} = \frac{1}{\tau}\int_0^\infty \rho_{LM}(t)\mathrm{e}^{-t/\tau}\,\mathrm{d}t, \tag{20.18}$$

where τ typically is the lifetime of the state under observation. The time average of $\dot{\rho}_{LM}$,

$$\bar{\dot{\rho}}_{LM} = \frac{1}{\tau}\left[\bar{\rho}_{LM} - \rho_{LM}(0)\right], \tag{20.19}$$

is obtained from (20.18) through integration by parts. Applied to the $I = 1$ equation system (20.17a), (20.17b), this gives

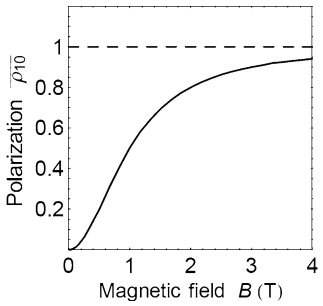


Fig. 20.1 A transverse electric- quadrupole interaction destroys the time average dipolar polarization $\bar{\rho}_{10} = \sqrt{\frac{1}{2}}\bar{P}_z$. However, a strong magnetic field decouples the quadrupole interaction and reestablishes the polarization. Shown is a decoupling curve with $\eta\omega_Q = \omega_0$ at $B = 1$ T

$$\frac{1}{\tau}\left[\bar{\rho}_{10} - \rho_{10}(0)\right] = -\mathrm{i}\eta\sqrt{2}\omega_Q(\bar{\rho}_{22} - \bar{\rho}_{2-2}), \tag{20.20a}$$

$$\frac{1}{\tau}\bar{\rho}_{2\pm2} = \mp\mathrm{i}(2\omega_0\bar{\rho}_{2\pm2} + \eta\sqrt{2}\omega_Q\bar{\rho}_{10}), \tag{20.20b}$$

for the case that only $\rho_{10}(0)$ is different from zero. The system of linear differential equations (20.17a), (20.17b) is turned into a system of ordinary linear equations that can be solved analytically. Readers familiar with the technique will have recognized that the change from transient to time average polarization is a Laplace transform.

For $\bar{\rho}_{10}$ we obtain, in particular,

$$\bar{\rho}_{10} = \frac{1 + 4\omega_0^2\tau^2}{1 + 4\tau^2(\omega_0^2 + \eta^2\omega_Q^2)}\rho_{10}(0). \tag{20.21}$$

In the limit $\omega_0\tau \gg 1$, with many precession periods during one lifetime, this simplifies to

$$\bar{\rho}_{10} = \frac{\omega_0^2}{\omega_0^2 + \eta\omega_Q^2}\rho_{10}(0). \tag{20.22}$$

How can we interpret this result? As a function of the external magnetic field $B_0 = \omega_0/\gamma_I$, we observe a *decoupling curve*, see Fig. 20.1: For a low magnetic field, the angular momentum states are mixed by the quadrupole interaction, and the original linear polarization is lost. For a high magnetic field, the quadrupole interaction is decoupled, and the full original polarization $\rho_{10}(0)$ is preserved. Such decoupling curves are generic and are observed in many different contexts. For a rotationally symmetric efg along the external magnetic field, i.e., for asymmetry parameter $\eta = 0$, the field dependence vanishes, and the initial polarization is unchanged. This can be generalized to the following:

A rotationally symmetric interaction cannot change a polarization that is rotationally symmetric about the same axis.

20.5 Angular Distribution of Radiation

When a polarized ensemble of particles emits secondary radiation, the angular distribution of this radiation can be calculated with the tools developed in the preceding chapters. We give two examples, first the electron angular distribution in the β-decay of polarized nuclei or particles, and second the photon angular distribution in atomic light scattering.

Nuclei and particles can be polarized in many ways. A steady state method is:

1. Atomic and nuclear polarization of a paramagnetic ensemble of atoms or of nuclei in condensed matter, in thermal equilibrium in a high external and/or internal magnetic field at low temperature, see (6.28) or (6.29), which sometimes is called the *brute force* method. A variant of this method is dynamic nuclear polarization, in which a more easily achievable electron polarization in a solid is transferred onto the nuclei.

Off-equilibrium methods involving particle reactions include:

2. Atomic polarization by optical pumping, in which the angular momentum of photons is transferred onto the atoms, as mentioned at the end of Sect. 18.3.
3. Muon polarization in parity-violating weak-interaction pion decay (see Sect. 4.1).
4. Neutron polarization by neutron scattering off magnetized materials.
5. Nuclear polarization by the capture of polarized neutrons in unpolarized nuclei.

In this last case, the activated nuclei acquire a pure polarization ρ_1 with a linear substate population, as shown in Fig. 19.2(a) for the special case $j = 1$. The reason of course is that neutrons as spin-$\frac{1}{2}$ particles have a maximum polarization rank $L_1 = 2 \times \frac{1}{2} = 1$ from (19.62d). Unpolarized nuclei of arbitrary spin have rank $L_2 = 0$. Coupling of both ranks gives $L = L_1 \pm L_2 = 1$ from (19.62c), which gives a pure polarization ρ_1 with linear occupation numbers.

20.5.1 Asymmetric β-Decay

Our first case is nuclear β-decay in which an electron and an electron–antineutrino are emitted, both having spin $\frac{1}{2}$. Let the β-instable nuclei have a pure polarization ρ_{10} with respect to a quantization axis z, created, for instance, by one of the methods listed above.

In β-decay, a fraction of this polarization is carried away by the emitted particles. As we saw in Sect. 4.1 for the case of muon β-decay, parity violation of the weak interaction leads to an electron polarization opposite to their direction of flight, called helicity, of size $h = -\upsilon_e/c$ for an electron emitted with velocity υ_e. As a consequence of their helicity, the decay electrons emitted from polarized nuclei have an asymmetric angular distribution in order to guarantee that angular momentum is the same before and after decay.

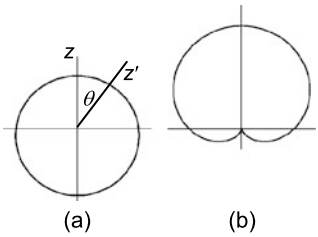


Fig. 20.2 Asymmetric distributions, see (20.25), in the β-decay of fully polarized neutrons in the limit of helicity $|h| = |\upsilon|/c \to 1$ (**a**) of the decay electrons with $A = -0.12$, (**b**) of the decay antineutrinos with $A = 1$

This β-decay asymmetry is proportional both to the polarization $P_z = \sqrt{2}\rho_{10}$ of the β-instable nuclei and to the helicity $h = -\upsilon_e/c$ of the electrons, and disappears if either one is zero. The electron emission rate under polar angle $\theta = 0$, i.e., in direction z of nuclear polarization, then is $W_\beta(0)\,\mathrm{d}\Omega = (1 + AhP_z)\,\mathrm{d}\Omega$, or

$$W_\beta(0)\,\mathrm{d}\Omega = \left(1 + \frac{\upsilon_e}{c} A\sqrt{2}\rho_{10}\right)\mathrm{d}\Omega\,, \tag{20.23}$$

normalized to $W_\beta(0) = 1$ for $\rho_{10} = 0$. The asymmetry parameter A can be calculated from weak-interaction theory, but its exact value is not of importance for our argument. For the $j_i = 1$ to $j_f = 0$ β^--decay of ^{60}Co, for example, in which parity violation was detected more than fifty years ago, $A = -1$.

Equation (20.23) gives the emission rate only into the z direction, and we want to deduce from it the full angular distribution of the decay electrons. For longitudinal nuclear polarization ρ_{10}, this angular distribution must be rotationally symmetric about z. We know that if a quantum system has polarization ρ_1 with respect to an axis z, then this polarization with respect to some other axis z', under angle θ towards z, is obtained by a rotation operation, for instance, about y, as $\rho'_1 = \mathrm{D}^{(1)}(0, \theta, 0)\rho_1$, see (17.4). Therefore, if we know the emission rate for z, we also know it for any other direction z'. We simply have to replace any ρ_L by the rotated ρ'_L. The polarization component ρ'_{10} along axis z' is

$$\rho'_{10} = d^{(1)}_{00}(\theta)\rho_{10} = \rho_{10}\cos\theta\,, \tag{20.24}$$

with the central matrix element $d^{(1)}_{00}(\theta) = \cos\theta$ of the rank-1 rotation matrix $\mathrm{d}^{(1)}$, see (16.48). The asymmetric electron emission rate (20.23) then changes to

$$W_\beta(\theta)\,\mathrm{d}\Omega = \left(1 + \frac{\upsilon}{c} A\sqrt{2}\rho'_{10}\right)\mathrm{d}\Omega = \left(1 + \frac{\upsilon}{c} A P_z \cos\theta\right)\mathrm{d}\Omega\,, \tag{20.25}$$

see Fig. 20.2 for the angular distributions of electrons and neutrinos. This is the well known parity-violating asymmetric electron angular distribution in the β-decay of polarized nuclei. In the muon spin rotation experiment of Sect. 4.1, it was this angular β-distribution that rotated together with the muon polarization about the magnetic field of the muon storage ring.

20.5.2 Anisotropic Photon Emission

A second example has a much wider field of application, namely the resonant scattering of light by atoms. We want to derive the angular distribution of the photons reemitted after an optical excitation of the atoms.

We have to distinguish between the polarization degrees of the photons and the atoms. In Sect. 12.1, we had learned that photons are spin-1 particles that have, with respect to their line of flight, only their $m_{\rm ph} = \pm 1$ substates populated. From (19.51a) with substate population $p_{0\,\rm ph} = 0$, we then find that the photons carry a longitudinal alignment of size $1/\sqrt{6}$. This alignment is independent of the sign of $m_{\rm ph}$, that is, of the relative sizes of populations $p_{\pm 1\,\rm ph}$, and therefore is independent of the photons' degree of circular polarization. In particular, unpolarized light, which is an incoherent mixture of $m_{\rm ph} = \pm 1$ or $\sigma^{\pm}$ photons, carries the same alignment of size $1/\sqrt{6}$.

By angular momentum conservation, absorption of unpolarized (or circularly polarized light) by unpolarized atoms carries this natural photonic alignment over to the excited atomic state. We treat the special case of an atom with a $j = 0$ ground state and a $j = 1$ excited state. Absorption of unpolarized light then leads to a population probability $p_{\pm 1} = \frac{1}{2}$ of the atomic $m = \pm 1$ substates, i.e., to an atomic alignment $\rho_{20} = 1/\sqrt{6}$ of the same size as that of the photons, and with respect to their axis of incidence.

Linearly polarized π-photons, oscillating along z, are a coherent superposition of $\sigma^{\pm}$ photons, namely the states H or V on the Bloch sphere of Fig. 12.1, with zero z-component of angular momentum. Hence π-photons populate the $m = 0$ substate of the excited atom, $p_0 = 1$, which leads to an alignment $\rho_{20} = -2/\sqrt{6}$.

Next we regard the reemitted photons from the atoms' decay back to their ground state. As the $j = 0$ ground state is unpolarized, all alignment is carried away by the scattered photons. The electromagnetic interaction responsible for this decay is parity conserving. Therefore, the rotationally symmetric angular distribution of the scattered photons must be mirror symmetric about the x–y plane through the atom, and odd ranks L of ρ_{L0} are forbidden, in contrast to (20.25). With $j = 1$ of the emitting state, only the longitudinal components ρ_{00} and ρ_{20} play a role. Let the photon emission rate along z, under $\theta = 0$ be given by

$$W_{\rm ph}(0)\,\mathrm{d}\Omega = (c_0\rho_{00} + c_2\rho_{20})\,\mathrm{d}\Omega\,, \tag{20.26}$$

where $\rho_{00} = 1/\sqrt{3}$ from (19.48a) is independent of the choice of z. The emission probability for the same atomic alignment into a different direction z' then is given by replacing ρ_{20} by $\rho'_{20} = d^{(2)}_{00}(\theta)\rho_{20}$, or

$$W_{\rm ph}(\theta)\,\mathrm{d}\Omega = \big[c_0\rho_{00} + c_2\rho_{20}P_2(\cos\theta)\big]\,\mathrm{d}\Omega\,. \tag{20.27}$$

An atomic alignment $\rho_{20} = 1/\sqrt{6}$ leads to $\sigma^{\pm}$ emission from substates $m = \pm 1$, an atomic alignment $\rho_{20} = -2/\sqrt{6}$ leads to π emission from $m = 0$.

The coefficients c_0 and c_2 are fixed by two requirements. First, photon emission must be isotropic for equal weights of σ^+, σ^-, and π emission from an unpolarized excited state, that is, for all values of θ,

$$W_{\sigma^+} + W_{\sigma^-} + W_\pi = 3c_0\rho_{00} = 1, \tag{20.28}$$

which is fulfilled if $c_0\rho_{00} = 1/3$, or $c_0 = 1/\sqrt{3}$. Second, forward scattering of unpolarized light under $\theta = 0$ by the atoms separately requires for $\sigma^\pm$ light

$$W_{\rm ph}(0) = W_{\sigma^+}(0) + W_{\sigma^-}(0) = 2\left(\frac{1}{3} + c_2\rho_{20}\right) = 1 \tag{20.29}$$

with $\rho_{20} = 1/\sqrt{6}$, which is fulfilled for $c_2\rho_{20} = 1/6$, or $c_2 = 1/\sqrt{6}$. With these values the angular distribution of $\sigma^\pm$ light becomes

$$W_{\sigma^\pm}(\theta)\,{\rm d}\Omega = \frac{1}{4}\big(1 + \cos^2\theta\big)\,{\rm d}\Omega, \tag{20.30}$$

which coincides with the classical emission characteristic of an oscillating current in a ring antenna, see (4.8) and Fig. 4.3(b). The σ-angular distribution reflects the distribution of angular momentum within the atom before the decay, given by the alignment $\rho_{20} \propto \langle {\rm T}_{20}\rangle \propto 3\langle {\rm J}_z^2\rangle - \langle \mathbf{J}^2\rangle$, which becomes $\propto 3\cos^2\theta - 1$, if one associates $\rm J_z$ semiclassically with $J\cos\theta$ and $\mathbf{J}^2$ with J^2.

The angular distribution of the emitted π light is, using (20.28),

$$W_\pi(\theta)\,{\rm d}\Omega = \frac{1}{2} I_0 \sin^2\theta\,{\rm d}\Omega, \tag{20.31}$$

which coincides with the emission characteristic of an electric dipole antenna, see Fig. 4.3(a) and (4.7). In summary:

> The angular distribution of radiation emitted from a polarized ensemble is derived from the elements of the appropriate rotation matrices.

Between polarization (Step 1) and analysis (Step 3) the excited system may, during its finite lifetime, be perturbed by external or internal fields (Step 2), under which the atomic or nuclear states evolve as given by the generalized spin precession equation (19.60), or by (20.15) for the special case of atomic or nuclear spin 1. This is the wide field of *perturbed angular distributions* (PAD) in nuclear physics, which includes the detection of NMR signals via β or γ-emission anisotropies discussed later on. In atomic physics, the same effects are traded under the names of Hanle effect, quantum beats, etc., see Sects. 4.2 and 12.3.

For completeness, we mention also the field of *perturbed angular correlations* (PAC), mainly used in nuclear solid-state physics; see, for instance, Schatz and Weidinger (1996), and references therein. There, instead of photon absorption, one uses

a preceding photon emission to align an intermediate nuclear state in a cascade of two γ-decays. The first photon γ_1 is isotropically emitted from a highly excited unpolarized nuclear state, and carries with it its natural alignment of size $1/\sqrt{6}$. It leaves behind, by angular momentum conservation, an aligned intermediate nuclear state. The second photon γ_2 therefore has an angular distribution that is anisotropic with respect to the direction of emission of γ_1, that is, the emission angles of γ_1 and γ_2 are correlated. During the lifetime of the intermediate state, a time dependent spin rotation of this intermediate alignment can be induced by external fields and measured in a γ–γ coincidence experiment.

Most general formulae for angular distributions and correlations in terms of rotation matrices are given in Chap. 19 of Fano and Racah (1959).

References

Fano, U., Racah, G.: Irreducible Tensorial Sets. Academic Press, New York (1959)
Schatz, G., Weidinger, A.: Nuclear Condensed Matter Physics: Nuclear Methods and Applications. Wiley, New York (1996)

Chapter 21
Reorientation in Time Dependent Fields

Abstract The generalized spin precession equation of the previous chapter is extended to systems subject to periodic time dependent perturbations. Main applications are in laser spectroscopy and nuclear or electron magnetic resonance. We discuss multiple-quantum transitions and dressed-atom effects, and give an outlook on experiments with nonclassical photon fields.

21.1 Radiofrequency Irradiation in a Magnetic Field

Up to this point, we restricted the discussion of the Liouville equation to time independent Hamiltonians. In this case, one ends up with a linear system of differential equations for the components of the density matrix ρ with time independent coefficients, which can be solved by standard techniques. The time average of ρ is obtained from a system of ordinary linear equations for the components of ρ which can be solved by matrix inversion. All these techniques break down when the Hamiltonian becomes time dependent.

To include these cases, we start with the simplest possible model Hamiltonian

$$\mathrm{H} = \omega_0 \mathrm{I}_z + 2\omega_1 \cos(\omega t)\mathrm{I}_x, \tag{21.1}$$

where $\omega_0 = \gamma_I B_0$ and $\omega_1 = \gamma_I B_1$. This is the situation found, for instance, in nuclear magnetic resonance experiments (this is why we use I for the angular momentum quantum number). For the case $I = \frac{1}{2}$, (21.1) was already discussed in Sect. 11.1. In the $s = \frac{1}{2}$ effective-spin model, the calculation can be immediately transferred to an arbitrary two-level system, e.g., to the situation where a light field couples to two atomic states. This approach is now generalized to arbitrary angular momenta.

21.1.1 Density Operator in the Rotating Frame

The standard technique to cope with time dependent systems is to try to find a transformation by which the time dependence is eliminated. For the Hamiltonian equation (21.1), this is achieved by changing to a coordinate system rotating with the

D. Dubbers, H.-J. Stöckmann, *Quantum Physics: The Bottom-Up Approach*, Graduate Texts in Physics,
DOI 10.1007/978-3-642-31060-7_21, © Springer-Verlag Berlin Heidelberg 2013

angular frequency ω about the z-axis. To this end, we define the density operator ρ_R in the rotation frame by

$$\rho_R = e^{-i\omega t I_z/\hbar} \rho e^{i\omega t I_z/\hbar}, \tag{21.2}$$

with the time derivative

$$\dot{\rho}_R = e^{-i\omega t I_z/\hbar} \dot{\rho} e^{i\omega t I_z/\hbar} + i\omega[\rho_R, I_z]. \tag{21.3}$$

Substituting for $\dot{\rho}$ the Liouville equation $\dot{\rho} = \frac{i}{\hbar}[\rho, H]$, see (19.16), we obtain a new Liouville equation

$$\dot{\rho}_R = \frac{i}{\hbar}[\rho_R, H_R], \tag{21.4}$$

where the Hamiltonian in the rotating frame is

$$H_R = \omega I_z + e^{-i\omega t I_z/\hbar} H e^{i\omega t I_z/\hbar}. \tag{21.5}$$

This is still true for arbitrary Hamiltonians. For the special Hamiltonian equation (21.1), we need nothing but the action of the rotation on I_z and I_x. As these are vector operators, this can immediately be written down to

$$I_{zR} = I_z, \qquad I_{xR} = I_x \cos\omega t + I_y \sin\omega t. \tag{21.6}$$

There is not even the need to resort to the rotation matrices.

We insert this into (21.1) and obtain for the Hamiltonian in the rotating frame

$$\begin{aligned} H_R &= e^{-i\omega t I_z/\hbar} H e^{i\omega t I_z/\hbar} \\ &= \omega_0 I_z + 2\omega_1 \cos(\omega t)\big[I_x \cos(\omega t) + I_y \sin(\omega t)\big] \\ &= \omega_0 I_z + \omega_1 I_x + \omega_1\big[I_x \cos(2\omega t) + I_y \sin(2\omega t)\big]. \end{aligned} \tag{21.7}$$

By going into the rotating frame, part of the radio frequency field has become constant in time, whereas another part rotates with twice the angular frequency. This can be understood without any formula: the radio frequency field oscillating linearly along the x-axis may be decomposed into two circular components, one rotating clockwise, the other one counter clockwise with angular frequency ω about the z-axis. In the rotating frame, the component rotating synchronously with the frame becomes constant, whereas the other one now rotates with 2ω. Next we neglect the component rotating in the wrong direction, which is the rotating wave approximation mentioned in Sect. 11.1. This is justified as long as the radiofrequency field is small compared to the static field. In Sect. 21.3, we shall discuss what happens if this assumption has to be abandoned. This approach is exact if a rotating field is applied in the experiment, or if circularly polarized light is used to induce an atomic resonance transition.

21.1.2 Rotating Wave Approximation

Omission of the time dependent part of (21.7) gives the Hamiltonian in the rotating frame

$$\mathrm{H_R} = (\omega_0 - \omega)\mathrm{I}_z + \omega_1 \mathrm{I}_x. \tag{21.8}$$

This corresponds to a static magnetic interaction in a field which is tilted by an angle β against the z-axis, see Fig. 11.1(a). From the figure, we read that β is obtained from

$$\sin\beta = \frac{\omega_1}{\omega_{\mathrm{eff}}}, \qquad \cos\beta = \frac{\omega_0 - \omega}{\omega_{\mathrm{eff}}}, \qquad \omega_{\mathrm{eff}} = \sqrt{(\omega_0 - \omega)^2 + \omega_1^2}, \tag{21.9}$$

with the angular frequency ω_{eff} by which the angular momenta are precessing about the direction of the effective field.

In the last step, we turn the z-axis into the direction of the effective field by a corresponding rotation about the y-axis. The density matrix in the new coordinate system is given by rotating ρ_{R} to

$$\tilde{\rho} = \mathrm{e}^{\mathrm{i}\beta \mathrm{I}_y/\hbar} \rho_{\mathrm{R}} \mathrm{e}^{-\mathrm{i}\beta \mathrm{I}_y/\hbar}. \tag{21.10}$$

This density matrix $\tilde{\rho}$ obeys the Liouville equation

$$\dot{\tilde{\rho}} = \frac{\mathrm{i}}{\hbar}[\tilde{\rho}, \tilde{\mathrm{H}}], \tag{21.11}$$

where

$$\tilde{\mathrm{H}} = \mathrm{e}^{\mathrm{i}\beta \mathrm{I}_y/\hbar} \mathrm{H_R} \mathrm{e}^{-\mathrm{i}\beta \mathrm{I}_y/\hbar}. \tag{21.12}$$

With $\mathrm{H_R}$ from (21.8) this becomes

$$\tilde{\mathrm{H}} = \omega_{\mathrm{eff}} \mathrm{e}^{\mathrm{i}\beta \mathrm{I}_y/\hbar} (\mathrm{I}_z \cos\beta + \mathrm{I}_x \sin\beta) \mathrm{e}^{-\mathrm{i}\beta \mathrm{I}_y/\hbar} = \omega_{\mathrm{eff}} \mathrm{I}_z. \tag{21.13}$$

21.1.3 Statistical Tensors in the Rotating Wave Approximation

We have thus reduced the problem to the interaction of the nuclear spin with a static magnetic field parallel to the z-axis. The above equations for the density operators can then be written in terms of the state multipoles ρ_{LM} by choosing the irreducible representation. The solution for this problem was given in (20.9) and yields for the irreducible components of $\tilde{\rho}$

$$\tilde{\rho}_{LM}(t) = \mathrm{e}^{-\mathrm{i}M\omega_{\mathrm{eff}}t} \tilde{\rho}_{LM}(0). \tag{21.14}$$

What remains to be done is to go back to the laboratory frame by reversing the two applied rotations. The relation between the ρ_{LM} and $\tilde{\rho}_{LM}$ is obtained as follows:

$$\begin{aligned}
\tilde{\rho}_{LM}(t) &= \mathrm{Tr}\big[\tilde{\rho}(t)\hat{\mathrm{T}}_{LM}\big] \\
&= \mathrm{Tr}\big[\mathrm{e}^{\mathrm{i}\beta \mathrm{I}_y/\hbar}\mathrm{e}^{-\mathrm{i}\omega t \mathrm{I}_z/\hbar}\rho(t)\mathrm{e}^{\mathrm{i}\omega t \mathrm{I}_z/\hbar}\mathrm{e}^{-\mathrm{i}\beta \mathrm{I}_y/\hbar}\hat{\mathrm{T}}_{LM}\big] \\
&= \mathrm{Tr}\big[\rho(t)\mathrm{e}^{\mathrm{i}\omega t \mathrm{I}_z/\hbar}\mathrm{e}^{-\mathrm{i}\beta \mathrm{I}_y/\hbar}\hat{\mathrm{T}}_{LM}\mathrm{e}^{\mathrm{i}\beta \mathrm{I}_y/\hbar}\mathrm{e}^{-\mathrm{i}\omega t \mathrm{I}_z/\hbar}\big] \\
&= \mathrm{Tr}\Big[\rho(t)\sum_{M'} D^{(L)}_{M'M}(-\omega t,\beta,0)\hat{\mathrm{T}}_{LM'}\Big] \\
&= \sum_{M'} D^{(L)}_{M'M}(-\omega t,\beta,0)\,\mathrm{Tr}\big[\rho(t)\hat{\mathrm{T}}_{LM'}\big] \\
&= \sum_{M'} D^{(L)}_{M'M}(-\omega t,\beta,0)\rho_{LM'}(t). \qquad (21.15)
\end{aligned}$$

Inverting the last equation, we obtain in a number of further steps

$$\begin{aligned}
\rho_{LM}(t) &= \sum_{M'} D^{(L)*}_{M'M}(-\omega t,\beta,0)\tilde{\rho}_{LM'}(t) \\
&= \sum_{M'} D^{(L)*}_{MM'}(-\omega t,\beta,0)\mathrm{e}^{-\mathrm{i}M'\omega_{\mathrm{eff}}t}\tilde{\rho}_{LM'}(0) \\
&= \sum_{M'} D^{(L)*}_{MM'}(-\omega t,\beta,0)\mathrm{e}^{-\mathrm{i}M'\omega_{\mathrm{eff}}t} \\
&\quad \times \sum_{M''} D^{(L)}_{M''M'}(0,\beta,0)\rho_{LM''}(0) \\
&= \sum_{M'M''} \mathrm{e}^{-\mathrm{i}(M\omega+M'\omega_{\mathrm{eff}})t} d^{(L)*}_{MM'}(\beta)d^{(L)}_{M''M'}(\beta)\rho_{LM''}(0). \qquad (21.16)
\end{aligned}$$

21.1.4 Time Average Results

For the time averaged state multipoles

$$\bar{\rho}_{LM} = \frac{1}{\tau}\int_0^\infty \rho_{LM}(t)\mathrm{e}^{-t/\tau}\,\mathrm{d}t \qquad (21.17)$$

with the lifetime τ of the state under observation, it follows from (21.16)

$$\bar{\rho}_{LM} = \sum_{M'M''} \frac{1}{1+\mathrm{i}(M\omega+M'\omega_{\mathrm{eff}})\tau} d^{(L)*}_{MM'}(\beta)d^{(L)}_{M''M'}(\beta)\rho_{LM''}(0). \qquad (21.18)$$

For the case of static magnetic fields the problem of radio frequency irradiation is thus completely solved.

The influence of radio frequency irradiation can be reduced to a reorientation in a static field in the rotating frame!

For the special case that there is only an initial longitudinal polarization $\rho_{10}(0)$ and all other components $\rho_{LM}(0)$ are zero, we obtain for the time averaged longitudinal polarization

$$\begin{aligned}\bar{\rho}_{10} &= \sum_M \frac{1}{1+\mathrm{i}M\omega_{\mathrm{eff}}\tau}\left[d^1_{0M}(\beta)\right]^2\rho_{10}(0)\\ &= \left[\frac{1}{1+\mathrm{i}\omega_{\mathrm{eff}}\tau}\frac{1}{2}\sin^2\beta+\cos^2\beta+\frac{1}{1-\mathrm{i}\omega_{\mathrm{eff}}\tau}\frac{1}{2}\sin^2\beta\right]\rho_{10}(0). \quad (21.19)\end{aligned}$$

Inserting expressions (21.9) for $\cos\beta$, $\sin\beta$ and ω_{eff}, we obtain

$$\bar{\rho}_{10} = \frac{1+(\omega_0-\omega)^2\tau^2}{1+[(\omega_0-\omega)^2+\omega_1^2]\tau^2}\rho_{10}(0). \quad (21.20)$$

For a long lived state we recover (11.6).

As a function of ω or, alternatively, as a function of $\omega_0 B_0/\gamma_I$, (21.20) describes a Lorentzian-shaped resonance curve with a minimum at the resonance $\omega = \omega_0$. It is remarkable that the spin quantum number I does not enter at all into (21.20). This is another demonstration of the feature that classical and quantum mechanical calculations give the same result as long as only magnetic interactions are involved.

21.2 Multiple Quantum Transitions

The rotating frame method works for a static Hamiltonian that commutes with I_z and a linearly oscillating field applied at a right angle to the z-axis. As a next step, we add to the magnetic field B_z of (21.1) a quadrupole interaction rotationally symmetric about z as well. The Hamiltonian then is

$$\mathrm{H} = \omega_0\mathrm{I}_z + \frac{\omega_\mathrm{Q}}{\hbar}\left(3\mathrm{I}_z^2-\mathbf{I}^2\right)+2\omega_1\cos(\omega t)\mathrm{I}_x. \quad (21.21)$$

Going to the rotating frame and neglecting the counter clockwise rotating component as before, we end up with the static Hamiltonian

$$\mathrm{H_R} = (\omega_0-\omega)\mathrm{I}_z+\omega_1\mathrm{I}_x+\frac{\omega_\mathrm{Q}}{\hbar}\left(3\mathrm{I}_z^2-\mathbf{I}^2\right). \quad (21.22)$$

For spin $I = 1$, the energy matrix is

$$\mathrm{H_R} = \hbar\begin{pmatrix}\omega_0-\omega+\omega_\mathrm{Q} & \omega_1/\sqrt{2} & 0\\ \omega_1/\sqrt{2} & -2\omega_\mathrm{Q} & \omega_1/\sqrt{2}\\ 0 & \omega_1/\sqrt{2} & \omega-\omega_0+\omega_\mathrm{Q}\end{pmatrix}. \quad (21.23)$$

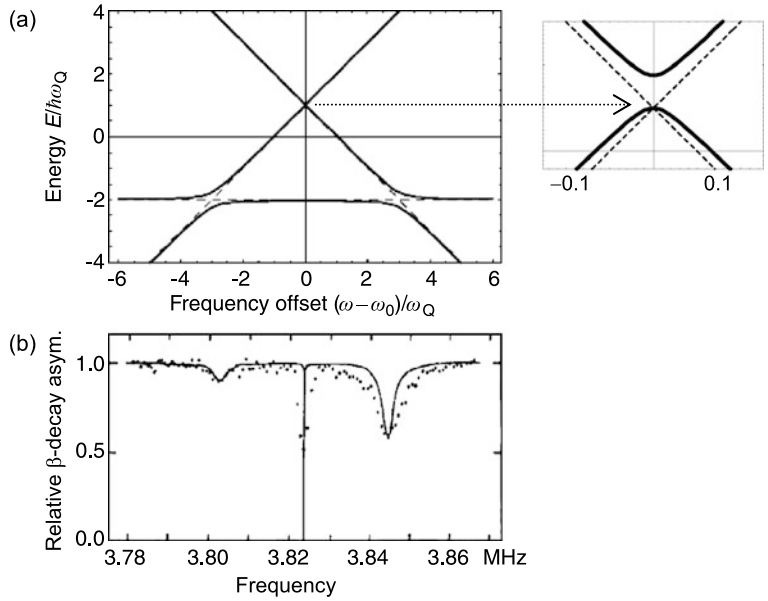


Fig. 21.1 Multiple-quantum NMR transition in ^{12}B ($I = 1$) implanted in a Be single crystal. Same scheme as in Fig. 11.3, but with an additional quadrupole interaction. (**a**) At the level crossings the RF field mixes two states and leads to a reduction of polarization. Resonances at $\omega = \omega_0 \pm 3\omega_Q$ are single quantum transitions, the resonance at $\omega = \omega_0$ is a two-quantum transition (see *enlarged view on the right*). (**b**) Multiple-quantum transitions are detected via the β-decay asymmetry of ^{12}B. Data points are from McDonald and McNab (1974)

The eigenvalues of H_R are obtained from the zeros of the characteristic polynomial

$$\det|x\mathrm{I} - \mathrm{H_R}| = \left[(x-\omega_Q)^2 - (\omega-\omega_0)^2\right](x+2\omega_Q) - \omega_1^2(x-\omega_Q). \qquad (21.24)$$

Figure 21.1(a) shows the energy eigenvalues as a function of ω. In the limit $\omega_1 \to 0$ with no high-frequency field applied, H_R is diagonal. The eigenenergies then are simply the diagonal elements of (21.23)

$$x_1 = \omega_Q - \omega + \omega_0, \qquad x_0 = -2\omega_Q, \qquad x_{-1} = \omega_Q + \omega - \omega_0. \qquad (21.25)$$

In this limit, the eigenenergies cross at the three points ω_0 and $\omega_0 \pm 3\omega_Q$, see the dashed lines in the figure. When the high-frequency field is turned on, the off-diagonal elements of H_R lead to a mixing of the angular momentum states involved in the crossing and thus convert the crossings into avoided crossings. In the process, the occupation numbers of the participating states are equilibrated, and a resonance signal appears in the polarization of the system.

Let us assume an initially purely dipolar polarization $\rho_{10}(0)$. Figure 21.1(b) shows what happens: At the crossings of the terms x_{-1}, x_0 and of the terms x_0, x_{+1} at frequencies $\omega_0 \pm 3\omega_Q$, one observes two broad minima in the time averaged polarization. They correspond to ordinary radio frequency transitions between

the two stated involved. In addition, there is a very narrow deep minimum at ω_0, where the states x_{-1}, and x_{+1} cross, corresponding to a $\Delta m = 2$ transition. Since the matrix element between the terms x_{-1} and x_1 of $\mathrm{H_R}$ is zero (the elements in the outer corners of (21.23)), it is a second order transition with respect of ω_1. It is called a *two-quantum transition*. It corresponds to the simultaneous absorption of two radio-frequency quanta. It is not difficult to calculate $\bar{\rho}_{10}$ exactly, by applying the techniques illustrated in Sect. 20.4. In the present situation, an inversion of a 3×3 matrix is sufficient, not even the eigenenergies are needed for the time averaged quantities.

The data points in Fig. 21.1(b) are from an experiment using the technique of β-NMR (McDonald and McNab 1974). In the experiment, β-instable ^{12}B nuclei ($I = 1$, $T_{1/2} = 20$ ms) were implanted into a single crystal of beryllium. The selection of a proper recoil angle leads to a spin orientation of the implanted nuclei. The details are not relevant here. Due to parity violation, the subsequent β-decay shows a left–right asymmetry, Sect. 20.5.1, that is proportional to the nuclear polarization ρ_{10}, in a similar way as muon polarization was monitored via the muon decay β-asymmetry in Sect. 4.1. Every destruction of the polarization, in the present case due the resonant irradiation of a radio-frequency field, shows up in a corresponding decrease of the β-decay asymmetry. The figure nicely shows the two $\Delta m = 1$ transitions and a sharp $\Delta m = 2$ transition. The depth of the two $\Delta m = 1$ resonances differs, since the polarization produced in the nuclear reaction is not purely dipolar ρ_{10}, but contains an alignment contribution ρ_{20} as well.

21.3 Dressed Atoms

Unfortunately, quite often the rotating frame method is not applicable. If there is *no* rotational symmetry about one axis, if there is more than one high frequency field, or if one cannot neglect the counter-rotating high-frequency component, then one has to proceed in a different manner. Atoms embedded in a strong photon field can be described in the framework of quantum electrodynamics and are called *dressed atoms*. Let us first describe these dressed-atom effects in terms of classical electromagnetic fields.

21.3.1 Dressed Atoms and the Floquet Theorem

We start with the very general time-periodic Hamiltonian

$$\mathrm{H}(t) = \sum_n \mathrm{H}_n \mathrm{e}^{\mathrm{i}n\omega t}, \tag{21.26}$$

where the Fourier coefficients H_n are arbitrary time independent operators. Time periodic systems are usually not treated in beginner's courses, but *spatially-periodic*

systems are well known from solid state physics. There *Bloch's theorem* states that in the spatially periodic potential of a crystal lattice, a stationary wave function can always be written as $\psi_{\boldsymbol{K}}(\boldsymbol{r}) = u_{\boldsymbol{K}}(\boldsymbol{r})\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}$, where $u_{\boldsymbol{K}}(\boldsymbol{r})$ is lattice-periodic, and $\boldsymbol{K}$ is a vector of the reciprocal lattice. The wave function is hence a plane wave of wave vector $\boldsymbol{k}$ whose amplitude is periodically modulated by the underlying lattice.

For time-periodic systems, the equivalent of Bloch's theorem is *Floquet's theorem*, which states that the time dependence of the wave function can be written as $\psi_\omega(t) = u_\omega(t)\ \mathrm{e}^{-\mathrm{i}Et/\hbar}$, where $u_\omega(t)$ is time-periodic with the period $T = 2\pi/\omega$ of the underlying Hamiltonian equation (21.26). It is sufficient to restrict E to the range $0 < E < \hbar\omega$, in analogy to solid state physics where $\boldsymbol{k}$ can be restricted to the first Brillouin zone. This may be interpreted in a hand-waving manner as follows: Since the Hamiltonian is time dependent, E is no longer a conserved quantity, but may increase or decrease in steps of $\hbar\omega$, in the same way as $\boldsymbol{k}$ in the solid-state analogy may increase in steps of $\boldsymbol{K}$. For this reason, E is sometimes called *pseudo energy*. The density matrix may be decomposed as

$$\rho(t) = \sum_n \rho_n \mathrm{e}^{\mathrm{i}(n\omega-\Omega)t}, \tag{21.27}$$

where $\Omega = E/\hbar$. (The operators ρ_n must not be confused with the statistical tensors ρ_L in the irreducible representation of ρ.)

Entering this ρ into the Liouville equation (19.16), we obtain

$$\sum_n (-\mathrm{i}\Omega\,\rho_n + \mathrm{i}n\omega\rho_n)\mathrm{e}^{\mathrm{i}n\omega t} = \frac{\mathrm{i}}{\hbar}\sum_{nm}[\rho_n, \mathrm{H}_m]\mathrm{e}^{\mathrm{i}(n+m)\omega t}. \tag{21.28}$$

Equating the coefficients of the $\mathrm{e}^{\mathrm{i}n\omega t}$ on both sides of the equation, we obtain the following equation system for the ρ_n,

$$-\mathrm{i}\Omega\,\rho_n = -\mathrm{i}n\omega\rho_n + \frac{\mathrm{i}}{\hbar}\sum_k [\rho_{n-k}, \mathrm{H}_k]. \tag{21.29}$$

We turned the original Liouville differential equation for a time dependent Hamiltonian into a system of linear equations with time independent Hamiltonians for the pseudo energy $E = \hbar\Omega$. The price to be paid is high, however: The equation system is unlimited on both sides, since n extends from $-\infty$ to $+\infty$, and we must suitably truncate it.

We come back to the example of Sect. 21.1, namely,

$$\mathrm{H} = \omega_0 \mathrm{I}_z + 2\omega_1 \cos(\omega t)\mathrm{I}_x. \tag{21.30}$$

In this case, $\mathrm{H}(t)$ in (21.26) has three terms

$$\mathrm{H}_0 = \omega_0 \mathrm{I}_z, \qquad \mathrm{H}_{\pm 1} = \omega_1 \mathrm{I}_x. \tag{21.31}$$

Inserting H_0 and $H_{\pm 1}$ into (21.29) we obtain

$$\ldots$$

$$-i\Omega\rho_1 = -i\omega\rho_1 - i\omega_1[\rho_2, I_x/\hbar] - i\omega_0[\rho_1, I_z/\hbar] - i\omega_1[\rho_0, I_x/\hbar], \tag{21.32a}$$

$$-i\Omega\rho_0 = -i\omega_1[\rho_1, I_x/\hbar] - i\omega_0[\rho_0, I_z/\hbar] - i\omega_1[\rho_{-1}, I_x/\hbar], \tag{21.32b}$$

$$-i\Omega\rho_{-1} = i\omega\rho_{-1} - i\omega_1[\rho_0, I_x/\hbar] - i\omega_0[\rho_{-1}, I_z/\hbar] - i\omega_1[\rho_{-2}, I_x/\hbar], \tag{21.32c}$$

$$\ldots$$

The equations for $\dot{\rho}_1$ and $\dot{\rho}_{-1}$ couple with components ρ_2 and ρ_{-2}, respectively. If these terms are neglected, one has a closed system of equations for ρ_1, ρ_0, and ρ_{-1}. This neglect is equivalent to the omission of the rapidly oscillating term in (21.7). One may extend the calculation also to ρ_2 and ρ_{-2}, and so on.

If we are interested in the linear polarization only (i.e., in the statistical tensors of rank 1), these equations reduce to an equation system for the expectation values of the components of the angular momentum operator **I**. Defining

$$\langle I_i \rangle_n = \mathrm{Tr}(\rho_n I_i), \quad i = x, y, z, \tag{21.33}$$

one obtains in the very same way as described repeatedly above a coupled differential equation system for $\langle I_x \rangle_n$, $\langle I_y \rangle_n$, and $\langle I_z \rangle_n$.

21.3.2 Dressed Atoms and Second Quantization

The previous approach ignores the fact that the radiation field, too, is quantized. A proper treatment of this feature needs *second quantization*. We do not want to enter into details here, but rather just sketch the approach. Details are found in textbooks of quantum mechanics like Messiah (1961). In the second quantization approach, the Hamiltonian consists of a term describing the undisturbed system, a term describing the radiation field, and a term describing the coupling between the two,

$$H = H_{\text{system}} + H_{\text{system-radiation}} + H_{\text{radiation}}, \tag{21.34}$$

where

$$H_{\text{radiation}} = \frac{1}{2}\hbar\omega(aa_+ + a_+a) \tag{21.35}$$

is the Hamiltonian of the radiation field. Finally,

$$H_{\text{system-radiation}} = \lambda(a + a_+)I_x \tag{21.36}$$

describes the interaction, where λ is a coupling constant. In photon number representation, the matrix of the photon creation operator a^+ has elements $\sqrt{n}$ on the

superdiagonal, while the annihilation matrix a has the same elements on the subdiagonal. The matrix of the *photon number operator* $H_{radiation}$ is diagonal with $(n+\frac{1}{2})\hbar\omega$ on its diagonal. The coupling constant λ has to be adjusted such that in the limit of large photon number n the classical and the second-quantization treatment of the radiation field yield the same results. $H_{system\text{-}radiation}$ replaces our previous radio frequency part $-2\omega_1 \cos(\omega t) I_x$. At the end of a longer calculation, one obtains exactly the same equation system (21.32a)–(21.32c) as above. One only has to make some adjustments:

First, the combined Hamiltonian equation (21.34) is time independent, meaning that now the total energy is conserved as it should be. It is still true that the energy of H_{system} is *not* conserved and may be increased or decreased in steps of $\hbar\omega$, whenever a photon is absorbed or emitted. But this is compensated for by a corresponding change of the energy of the radiation field.

Second, ρ_n in (21.27) now has to be interpreted as the density matrix for the case that there are exactly n photons in the radiation field. Whenever a photon is absorbed or emitted, ρ_n changes to $\rho_{n\pm1}$. This is exactly what is observed in equation system (21.32a)–(21.32c) for the terms going with ω_1. As a consequence, this equation system, which is open to both sides in the classical treatment of the radiation field, in the second-quantization approach has a bottom, since n cannot take negative values. In fact, this is the only qualitative difference between the two approaches. Whenever one is sufficiently far from the bottom, meaning large photon numbers, both approaches are equivalent.

21.3.3 A Dressed Neutron Experiment

It is instructive to look at the eigenenergy scheme for the case of a dressed particle. This scheme is shown by the dashed lines of Fig. 21.2(a), for the simplest case where $H_{system} = \omega_0 I_z$ is the Zeeman interaction with a static magnetic field, and $I = \frac{1}{2}$. For vanishing coupling $\lambda \to 0$ to the radiation field, one observes, as a function of ω_0, just two sets of parallel lines with opposite slopes, which is the manifestation of the Zeeman effect, Fig. 2.2. In this case, $H_{radiation}$ generates an offset of $\hbar\omega$ for each additional photon.

For nonzero coupling to the radiation field, every second crossing is converted into an anti-crossing. Depending on the photon numbers associated with the terms involved, they correspond to $\Delta n = 1, 3, \ldots$ quantum transitions. Since the absorption or emission of a photon always changes the particle's quantum number m_I by ±1, the crossings with an even Δn are not influenced by the radiation field. Since in the Hamiltonian in first order only terms with $\Delta n = \pm1$ are linked, the level repulsions at the $\Delta n = \pm3, \pm5$, etc. are of higher order and correspondingly smaller. The throughgoing line in Fig. 21.2(a) is calculated by diagonalization of the matrix of (21.34), truncated to photon numbers $n = 18$ to $n = 28$.

The dressed-particle energy scheme can be reconstructed experimentally with a beam of polarized neutrons by irradiating the neutrons with a magnetic RF-field of

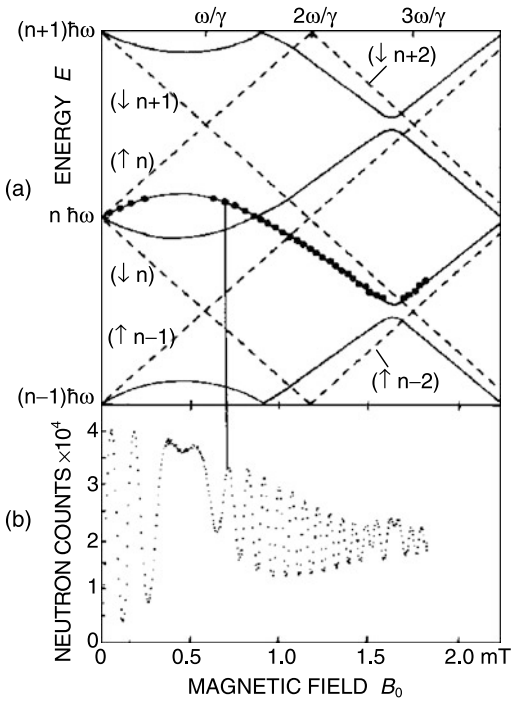


Fig. 21.2 Dressed-neutron level diagram: Reconstruction of the total energy diagram from the spin rotation curve of a polarized neutron beam in a static magnetic field, exposed to radio frequency quanta. The energies are given in units of $\hbar\omega$, the numbers at the ordinate denote the numbers of photons associated with the states. For details, see Muskat et al. (1987)

amplitude B_1 and arbitrary frequency ω. From the measured neutron spin rotation curve $P(B_0)$ shown in Fig. 21.2(b), one can reconstruct the dressed-neutron level scheme shown in Fig. 21.2(a). For details, see Muskat et al. (1987).

The effective g-factor of the dressed neutrons is the slope $g_{\text{eff}} = \partial E/\partial B_0$ of the energy levels in Fig. 21.2 at low field $B_0 \to 0$, see (7.3). Hence by tuning the RF-field amplitude B_1, one can change the dressed-neutron spin precession frequency $\omega_{\text{L}} = g_{\text{eff}}\mu_{\text{B}}B/\hbar$ at will. When we increase the RF field to a critical value $\omega_1/\omega = 2.40$, the slope g_{eff} becomes zero and the measured neutron spin rotation comes to a full stop, and reverses its direction of rotation for still higher B_1 values, as predicted by theory. These dressed particle effects are frequently employed in contemporary atomic physics experiments, but their demonstration with a beam of polarized neutrons has the advantage that it appears in pure form such that no interference effects between in and outgoing light waves have to be invoked.

21.3.4 Outlook on Nonclassical Photon Interactions

In our above examples, second quantization of the radiation field was very helpful for physically understanding a number of peculiar multiple-quantum and dressed-atom effects, but was not indispensable: in the limit of large photon numbers, the photons can as well be treated as a classical electromagnetic radiation field, in which

case the dressed-atom effects can be derived also semiclassically. However, in past years a growing number of experiments in atomic physics have been performed whose outcome cannot be predicted by semiclassical theory, but does require the framework of quantum field theory.

The detailed description of this exciting development is beyond the scope of this book, therefore we give only a short account of it. In experiments on *cavity quantum electrodynamics*, one lets single atoms interact one after the other with single photons (or a precise small number of photons) in a microwave cavity of extremely high quality factor. In this way, one can pursue the atom–photon interplay in much detail; for a review, see Walther et al. (2006). One can even nondestructively probe the presence of a photon in the cavity via the light shift it induces in a passing Rydberg atom, which allows one to "count photons like marbles in a box". In such quantum-nondemolition experiments, the photon is an object of investigation whose evolution can be repeatedly interrogated without destroying it. For an introduction to this subject, see Guerlin et al. (2007) and references therein.

References

Guerlin, C., Bernu, J., Deléglise, S., Sayrin, C., Gleyzes, S., Kuhr, S., Brune, M., Raimond, J.-M., Haroche, S.: Progressive field-state collapse and quantum non-demolition photon counting. Nature **448**, 889–894 (2007)

McDonald, R.E., McNab, T.K.: Double-quantum transitions in the resonance depolarization of ^{12}B. Phys. Rev. Lett. **32**, 1133–1135 (1974)

Messiah, A.: Quantum Mechanics. North-Holland, Amsterdam (1961). Reprint Dover Publications (1999)

Muskat, E., Dubbers, D., Schärpf, O.: Dressed neutrons. Phys. Rev. Lett. **58**, 2047 (1987)

Walther, H., Varcoe, B.T.H., Englert, B.-G., Becker, T.: Cavity quantum electrodynamics. Rep. Prog. Phys. **69**, 1325–1382 (2006)

Chapter 22
Relaxation and Decoherence

Abstract In all solutions of the generalized spin precession equation studied so far, the components of the density matrix were either constant or exhibited an oscillatory behavior. In reality, however, every oscillation will eventually be damped out, and after some time the system will return to its thermal equilibrium, a process called relaxation. A perturbation approach within second order approximation leads to the remarkable result that each rank of tensor polarization decays with its own relaxation time constant. A similar result is obtained with a stochastic relaxation model in which relaxation is due to jump diffusion. The book concludes with a discussion of decoherence and its distinction from relaxation.

22.1 General Features

In NMR, relaxation was phenomenologically taken into account by adding extra terms to the spin precession equation. This leads to Bloch's equations to describe the precession and relaxation of magnetization in an external magnetic field, introduced in Sect. 10.3. The additional terms guarantee that in the limit $t \to \infty$ all oscillations are damped out and the magnetization approaches its thermal equilibrium value M_{z0}.

Due to the additional relaxation terms, the whole process has become irreversible! But how can this be possible, in view of the fact that there is no preferential direction of time in the Liouville equation? Let us illustrate this in a thought experiment: if one follows the evolution of a system forward in time, then reverses the sign of time and follows the time evolution backward in time, one will completely recover the initial state! In spin echo experiments in NMR, as described in Sect. 11.2, this can to some extent be realized experimentally.

The solution of this apparent contradiction comes from the fact that the system one is interested in never is completely isolated. There will be always a coupling to the surroundings, which traditionally is termed the *bath*. The total Hamiltonian will thus look like

$$\mathrm{H} = \mathrm{H_S} + \mathrm{H_{SB}} + \mathrm{H_B}, \tag{22.1}$$

D. Dubbers, H.-J. Stöckmann, *Quantum Physics: The Bottom-Up Approach*, Graduate Texts in Physics,
DOI 10.1007/978-3-642-31060-7_22, © Springer-Verlag Berlin Heidelberg 2013

where $\mathrm{H_S}$ is the Hamiltonian of our system. For the purpose of illustration, we again take the interaction of a nuclear spin with an external magnetic field,

$$\mathrm{H_S} = -\frac{\mu}{\hbar I} B_0 \mathrm{I}_z. \tag{22.2}$$

The Hamiltonian $\mathrm{H_{SB}}$ couples the system to the bath. In NMR, this coupling is provided by the dipole–dipole interaction between the system spin I and the spins of the surrounding nuclei. Again we simplify matters by replacing these dipole fields by classical fields,

$$\mathrm{H_{SB}} = -\frac{\mu}{\hbar I} \mathbf{I} \cdot \boldsymbol{B}_{\mathrm{loc}}, \tag{22.3}$$

thus ignoring the quantum mechanical origin of B_{loc}.

The bath Hamiltonian $\mathrm{H_B}$ is, strictly speaking, the Hamiltonian of the rest of the universe. This includes even the experimental setup; a vexing aspect which shall not be pursued further here. But even if one takes into account *only* the particles of the probe, one ends up with a number of particles of the order of the Avogadro number. It is hopeless to try to solve the resultant Liouville equation exactly. So what to do?

We start with the complete Liouville equation for the system coupled to the environment

$$\dot{\rho} = \frac{\mathrm{i}}{\hbar}[\rho, \mathrm{H_S} + \mathrm{H_{SB}} + \mathrm{H_B}]. \tag{22.4}$$

Here, ρ is the total density matrix including the environment. However, since we treated the local field in (22.3) classically with no back-action from the system, we can limit ρ to the density matrix of the system. This means replacing the bath density matrix by the unity matrix for the bath states, while in reality the state occupation numbers will obey a Boltzmann distribution. Therefore, our approach means essentially a high-temperature approximation. It is possible to do better, though only at the cost of considerable technical complications without getting new insight.

Next we remove the bath Hamiltonian by the transformation

$$\tilde{\rho} = \mathrm{e}^{\mathrm{i}\mathrm{H_B}t/\hbar} \rho \mathrm{e}^{-\mathrm{i}\mathrm{H_B}t/\hbar}. \tag{22.5}$$

How this works is shown in Appendix A.9. Using (A.41) with H_1 replaced by $\mathrm{H_B}$ and H_2 replaced by $\mathrm{H_S} + \mathrm{H_{SB}}(t)$, we find the new Liouville equation

$$\dot{\tilde{\rho}} = \frac{\mathrm{i}}{\hbar}\big[\tilde{\rho}, \mathrm{H_S} + \mathrm{H_{SB}}(t)\big], \tag{22.6}$$

where

$$\mathrm{H_{SB}}(t) = \mathrm{e}^{\mathrm{i}\mathrm{H_B}t/\hbar} \mathrm{H_{SB}}(0) \mathrm{e}^{-\mathrm{i}\mathrm{H_B}t/\hbar}. \tag{22.7}$$

In our classical field approximation, $\mathrm{H_{SB}}(t)$ may be written as

$$\mathrm{H_{SB}}(t) = -\frac{\mu}{\hbar I} \mathbf{I} \cdot \boldsymbol{B}_{\mathrm{loc}}(t), \tag{22.8}$$

where our whole ignorance on the details of the bath is now absorbed by an unknown time dependence of $\boldsymbol{B}_{\mathrm{loc}}(t)$. There are essentially two strategies to cope with this situation, which will be sketched in the next two sections. A really thorough treatment of relaxation would exceed by far the scope of an introductory presentation.

22.2 The Perturbative Approach

The standard approach to treat relaxation is to use a second order expansion of the system-bath interaction. Our starting point is the Liouville equation (22.6),

$$\dot{\rho} = \frac{\mathrm{i}}{\hbar}\big[\rho, \mathrm{H}_0 + \mathrm{H}_1(t)\big], \tag{22.9}$$

with small $\mathrm{H}_1 \ll \mathrm{H}_0$, where we have removed the tilde on ρ and simplified the notation. First, we remove H_0 from the Liouville equation by means of the transformation

$$\hat{\rho} = \mathrm{e}^{\mathrm{i}\mathrm{H}_0 t/\hbar}\rho\mathrm{e}^{-\mathrm{i}\mathrm{H}_0 t/\hbar}. \tag{22.10}$$

This is the so-called *interaction representation*. $\hat{\rho}$ obeys the Liouville equation

$$\dot{\hat{\rho}} = \frac{\mathrm{i}}{\hbar}\big[\hat{\rho}, \hat{\mathrm{H}}_1(t)\big], \tag{22.11}$$

with

$$\hat{\mathrm{H}}_1(t) = \mathrm{e}^{\mathrm{i}\mathrm{H}_0 t/\hbar}\mathrm{H}_1(t)\mathrm{e}^{-\mathrm{i}\mathrm{H}_0 t/\hbar}. \tag{22.12}$$

Note that there are two sources for the time dependence of $\mathrm{e}^{\mathrm{i}\mathrm{H}_0 t/\hbar}\rho\mathrm{e}^{-\mathrm{i}\mathrm{H}_0 t/\hbar}$, one resulting from transformation (22.5) by which we removed the bath Hamiltonian, the other one resulting from (22.12), by which the system Hamiltonian was removed. There is one important difference between these time dependencies: We have no control over the bath variable, meaning that for us the time dependence of $\mathrm{H}_1(t)$ looks just like noise. The system Hamiltonian is under our control, meaning that the second transformation (22.12) produces an oscillatory time dependence exhibiting the spectrum of H_0.

For further calculation, it is convenient to rewrite (22.11) in the irreducible tensor notation, using (19.55),

$$\dot{\hat{\rho}}_{LM} = \frac{\mathrm{i}}{\hbar}\sum_{L'M'}\hat{\rho}_{L'M'}(L'M'|\hat{\mathrm{H}}_1(t)|LM). \tag{22.13}$$

Next we integrate both sides of (22.13) over t,

$$\hat{\rho}_{L'M'}(t) = \hat{\rho}_{L'M'}(0) + \frac{\mathrm{i}}{\hbar}\sum_{L''M''}\int_0^t \mathrm{d}\tau\,\hat{\rho}_{L''M''}(\tau)(L''M''|\hat{\mathrm{H}}_1(\tau)|L'M'), \tag{22.14}$$

where in addition we have changed the indices. We substitute this expression for $\hat{\rho}_{L'M'}$ on the right-hand side of (22.13),

$$\dot{\hat{\rho}}_{LM} = \frac{\mathrm{i}}{\hbar} \sum_{L'M'} \hat{\rho}_{L'M'}(0)(L'M'|\hat{\mathrm{H}}_1(t)|LM) - \frac{1}{\hbar^2} \sum_{L'M'L''M''} \int_0^t \mathrm{d}\tau\, \hat{\rho}_{L''M''}(\tau)(L''M''|\hat{\mathrm{H}}_1(\tau)|L'M')(L'M'|\hat{\mathrm{H}}_1(t)|LM), \tag{22.15}$$

or

$$\dot{\hat{\rho}}_{LM} = \frac{\mathrm{i}}{\hbar} \sum_{L'M'} \hat{\rho}_{L'M'}(0)(L'M'|\hat{\mathrm{H}}_1(t)|LM) - \frac{1}{\hbar^2} \sum_{L'M'} \int_0^t \mathrm{d}\tau\, \hat{\rho}_{L'M'}(t-\tau)(L'M'|\hat{\mathrm{H}}_1(t-\tau)\hat{\mathrm{H}}_1(t)|LM), \tag{22.16}$$

where the completeness of the irreducible tensor operators has been used. In addition, we changed the integration variable from τ to $t - \tau$. To avoid a possible misinterpretation of the equation, we remind ourselves of the operator multiplication convention $\mathrm{AB}|LM) = [\mathrm{A}, [\mathrm{B}, \hat{\mathrm{T}}_{LM}]]$, see (19.37)!

Equation (22.16) is still exact, though to proceed further we have to use approximations. There are various alternatives; however, we just follow the standard route. Essentially, there are two assumptions:

First, it is assumed that $\hat{\mathrm{H}}_1(t)$ varies rapidly with time as compared to $\dot{\hat{\rho}}_{LM}(t)$. This is a reasonable assumption since $\hat{\mathrm{H}}_1(t)$ had been assumed as small, and in the absence of $\hat{\mathrm{H}}_1(t)$, the polarization $\hat{\rho}_{LM}(t)$ would just be constant, $\dot{\hat{\rho}} = 0$ from (22.11). One then averages (22.16) over the time variations of $\hat{\mathrm{H}}_1(t)$,

$$\dot{\hat{\rho}}_{LM} = \frac{\mathrm{i}}{\hbar} \sum_{L'M'} \hat{\rho}_{L'M'}(0)(L'M'|\overline{\hat{\mathrm{H}}_1(t)}|LM) - \frac{1}{\hbar^2} \sum_{L'M'} \int_0^t \mathrm{d}\tau\, \hat{\rho}_{L'M'}(t-\tau)(L'M'|\overline{\hat{\mathrm{H}}_1(t-\tau)\hat{\mathrm{H}}_1(t)}|LM), \tag{22.17}$$

where the bar denotes the time average. Without loss of generality, we may assume that the time average of $\hat{\mathrm{H}}_1(t)$ disappears (if necessary after a proper redefinition of H_0 in (22.9)). Then the first term on the right-hand side vanishes. Furthermore the term $\mathrm{G}(\tau) = \overline{\hat{\mathrm{H}}_1(t-\tau)\hat{\mathrm{H}}_1(t)}$ will decay on a time scale τ_c, the correlation time for the fluctuations of the bath variables. If τ_c is short compared to the time scale on which the changes of $\hat{\rho}_{LM}(t)$ occur, we may replace $t - \tau$ in the argument of $\hat{\rho}_{L'M'}$ on the right-hand side by t, and extend the upper limit of the integration to ∞. We

then get

$$\dot{\hat{\rho}}_{LM} = -\frac{1}{\hbar^2} \sum_{L'M'} \int_0^\infty \mathrm{d}\tau\, \hat{\rho}_{L'M'}(t)(L'M'|\overline{\hat{\mathrm{H}}_1(t-\tau)\hat{\mathrm{H}}_1(t)}|LM). \tag{22.18}$$

Second, $\hat{\mathrm{H}}_1(t)$ is taken up to second order perturbation only. We then may restrict the sum over $(L'M')$ to just one term (LM) to obtain

$$\dot{\hat{\rho}}_{LM} = -\lambda_{LM}\hat{\rho}_{LM}, \tag{22.19}$$

where

$$\lambda_{LM} = \frac{1}{\hbar^2} \int_0^\infty \mathrm{d}\tau (LM|\overline{\hat{\mathrm{H}}_1(t-\tau)\hat{\mathrm{H}}_1(t)}|LM). \tag{22.20}$$

Equation (22.19) constitutes the main result of this section:

> Each tensor component $\hat{\rho}_{LM}(t)$ in the interaction system decays exponentially with its own decay constant λ_{LM}.

Going back to the laboratory frame, we obtain for tensor rank $L = 1$ Bloch's equations, where we can identify λ_0 and $\lambda_{\pm 1}$ with T_1^{-1} and T_2^{-1}, respectively. We did not recover, however, a decay towards a thermal equilibrium in this way, since all tensor components at the very end decay to zero. This is not a deficiency of the approach, but only a consequence of the neglect of the bath details, as discussed above. To get explicit expressions for the relaxation rates from (22.20) for various situations found in experiments still some effort is needed. We refer to standard monographs on NMR, in particular to Abragam (1961). The book does not use the irreducible tensor formalism, however.

22.3 The Stochastic Approach

The other alternative to solve the Liouville equation (22.9) is to take $\mathrm{H}(t) = \mathrm{H}_0 + \mathrm{H}_1(t)$ as piecewise constant. One may imagine, for instance, the diffusive jumps of an atom in a lattice or flips of a molecule between different configurations (for an example, see Fig. 13.1), though many other situations are imaginable. For a given sequence, the Hamiltonian then looks like

$$\mathrm{H}(t) = \begin{cases} \mathrm{H}_0 & \text{if } 0 < t < t_1, \\ \mathrm{H}_1 & \text{if } t_1 < t < t_2, \\ \mathrm{H}_2 & \text{if } t_2 < t < t_3, \\ \dots \end{cases} \tag{22.21}$$

where the H_n are taken from a finite set. Since $\mathrm{H}(t)$ is piecewise constant, ρ may be obtained by a successive integration of the Liouville equation. For t between t_n and t_{n+1}, one obtains

$$\rho(t) = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\mathrm{H}_n(t-t_n)}\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\mathrm{H}_{n-1}(t_n-t_{n-1})}\cdots\mathrm{e}^{-\mathrm{i}\mathrm{H}_0 t_1/\hbar}\rho(0) \times \mathrm{e}^{\mathrm{i}\mathrm{H}_0 t_1/\hbar}\mathrm{e}^{\mathrm{i}\mathrm{H}_1(t_2-t_1)/\hbar}\cdots\mathrm{e}^{\mathrm{i}\mathrm{H}_n(t-t_n)/\hbar}. \tag{22.22}$$

This expression has to be averaged over all possible jump sequences. To this end, we divide the time t into N intervals of length $\Delta t = t/N$. Let $P(\alpha_0, \alpha_1, \ldots, \alpha_N)$ be the probability of H to take the choice H_{α_0} in the interval $0 < t < \Delta t$, the choice H_{α_1} in the interval $\Delta t < t < 2\Delta t$, etc. Then we obtain from (22.22)

$$\rho(t) = \sum_{\alpha_1,\ldots\alpha_N} P(\alpha_0, \ldots, \alpha_N)\mathrm{e}^{-\mathrm{i}\mathrm{H}_{\alpha_N}\Delta t/\hbar}\cdots\mathrm{e}^{-\mathrm{i}\mathrm{H}_{\alpha_1}\Delta t/\hbar}\rho(0) \times \mathrm{e}^{\mathrm{i}\mathrm{H}_{\alpha_1}\Delta t/\hbar}\cdots\mathrm{e}^{\mathrm{i}\mathrm{H}_{\alpha_N}\Delta t/\hbar}. \tag{22.23}$$

It is suitable to separate out the last summation over α_N and write (22.23) in the form

$$\rho(t) = \sum_{\alpha} \rho_\alpha(t), \tag{22.24}$$

where

$$\rho_\alpha(t) = \sum_{\alpha_1,\ldots,\alpha_{N-1}} P(\alpha_0, \ldots, \alpha_{N-1}, \alpha)\mathrm{e}^{-\mathrm{i}\mathrm{H}_\alpha\Delta t/\hbar}\cdots\mathrm{e}^{-\mathrm{i}\mathrm{H}_{\alpha_1}\Delta t/\hbar}\rho(0) \times \mathrm{e}^{\mathrm{i}\mathrm{H}_{\alpha_1}\Delta t/\hbar}\cdots\mathrm{e}^{\mathrm{i}\mathrm{H}_\alpha\Delta t/\hbar}. \tag{22.25}$$

$\rho_\alpha(t)$ may be looked upon as the density matrix of the subset of the ensemble experiencing at time t just the interaction H_α. To proceed further, we now assume that the jumps are *stationary Markov processes* meaning that (i) the jump probabilities are independent of the history of the preceding jumps, and (ii) the jump probabilities are independent of time.

Then the jump probability $P(\alpha_0, \ldots, \alpha_N)$ factorizes:

$$P(\alpha_0, \ldots, \alpha_{N-1}, \alpha) = \begin{cases} P(\alpha_0, \ldots, \alpha_{N-1})W_{\alpha_{N-1}\alpha}\Delta t & \text{if } \alpha_{N-1} \neq \alpha, \\ P(\alpha_0, \ldots, \alpha_{N-1})(1 - \sum_\beta W_{\alpha\beta}\Delta t) & \text{if } \alpha_{N-1} = \alpha. \end{cases} \tag{22.26}$$

Here $W_{\alpha\beta}$ is the probability per time for a jump from H_α to H_β. The equation expresses the fact that there are two possibilities: For $\alpha_{N-1} \neq \alpha$, there had been a jump in the time interval Δt from $\mathrm{H}_{\alpha_{N-1}}$ to H_α, and for $\alpha_{N-1} = \alpha$, no jump happened to any other H_β. Inserting the factorization (22.26) into (22.25), and performing the

limit $\Delta t \to 0$, we obtain

$$\dot{\rho}_\alpha = \frac{\mathrm{i}}{\hbar}[\rho_\alpha, \mathrm{H}_\alpha] + \sum_\beta (W_{\beta\alpha}\rho_\beta - W_{\alpha\beta}\rho_\alpha). \tag{22.27}$$

Equation (22.27) is the starting point for the stochastic relaxation theory. It is a special case of the Lindblad equation, the most general form of a quantum Markov master equation insuring positive definiteness of the density matrix (see Chap. 3 of the monograph by Breuer and Petruccione (2002) for details). The terms on the right-hand side of (22.27) allow an easy interpretation. The first term describes the reorientation of the ensemble at times when it experiences the Hamiltonian H_α, the next term describes the jumps from H_β to H_α and back. There are a number of monographs on stochastic relaxation theories, among others the books by Dattagupta (1987) and Van Kampen (1981).

As an illustration, we restrict ourselves to the simplest situation where there are jumps between just two possible Hamiltonians. We go back to our previous notation for the Hamiltonian, $\mathrm{H}(t) = \mathrm{H}_0 + \mathrm{H}_1(t)$, where it is now assumed that $\mathrm{H}_1(t)$ jumps stochastically between the two values H_1 and $-\mathrm{H}_1$. Equation (22.27) then reads

$$\dot{\rho}_1 = \frac{\mathrm{i}}{\hbar}[\rho_1, \mathrm{H}_0 + \mathrm{H}_1] - W\rho_1 + W\rho_2, \tag{22.28a}$$

$$\dot{\rho}_2 = \frac{\mathrm{i}}{\hbar}[\rho_2, \mathrm{H}_0 - \mathrm{H}_1] - W\rho_2 + W\rho_1. \tag{22.28b}$$

Introducing the total density matrix $\rho = \rho_1 + \rho_2$ and the difference $\delta = \rho_1 - \rho_2$ as new variables, the equation system reads

$$\dot{\rho} = \frac{\mathrm{i}}{\hbar}\big\{[\rho, \mathrm{H}_0] + [\delta, \mathrm{H}_1]\big\}, \tag{22.29a}$$

$$\dot{\delta} = \frac{\mathrm{i}}{\hbar}\big\{[\delta, \mathrm{H}_0] + [\rho, \mathrm{H}_1]\big\} - 2W\delta. \tag{22.29b}$$

An inspection of these equations shows without any calculation the qualitative behavior of the solution: In the limit of jump rates $W \to 0$, the equations (22.28a), (22.28b) for ρ_1 and ρ_2 are decoupled, resulting in an inhomogeneous superposition of the two sub-ensembles reorienting under the influence of $\mathrm{H}_0 + \mathrm{H}_1$ and $\mathrm{H}_0 - \mathrm{H}_1$, respectively. In the limit $W \to \infty$, on the other hand, (22.29b) shows that δ now decays rapidly to zero, and after a very short time in (22.29a) only the first term on the right-hand side will survive, resulting in a reorientation of the total density matrix ρ in the time averaged Hamiltonian H_0. In this latter limit, the fluctuating part $\pm\mathrm{H}_1$ of the Hamiltonian does not show up any longer in the spectrum, due to motional narrowing as discussed in Sect. 11.3, but it gives rise to relaxation. This is the aspect which shall be explored in some more detail.

To this end, we proceed exactly in the same way as in the preceding section. First, we remove the static Hamiltonian H_0 by a transformation into the interaction

system,

$$\dot{\hat{\rho}} = \frac{\mathrm{i}}{\hbar}\big[\hat{\delta}, \hat{\mathrm{H}}_1(t)\big], \tag{22.30a}$$

$$\dot{\hat{\delta}} = \frac{\mathrm{i}}{\hbar}\big[\hat{\rho}, \hat{\mathrm{H}}_1(t)\big] - 2W\hat{\delta}, \tag{22.30b}$$

where the hatched quantities are defined exactly as in the previous section (see (22.10) and (22.12)). The second equation may be integrated with respect to δ,

$$\hat{\delta}(t) = \hat{\delta}(0) + \frac{\mathrm{i}}{\hbar}\mathrm{e}^{-2Wt}\int_0^t \mathrm{d}\tau\, \mathrm{e}^{2W\tau}\big[\hat{\rho}(\tau), \hat{\mathrm{H}}_1(\tau)\big]. \tag{22.31}$$

Plugging this expression for $\hat{\delta}$ into (22.30a), we get

$$\dot{\hat{\rho}}(t) = \frac{\mathrm{i}}{\hbar}\big[\hat{\delta}(0), \mathrm{H}_1(t)\big] - \frac{1}{\hbar^2}\int_0^t \mathrm{d}\tau \mathrm{e}^{-2W\tau}\big[\big[\hat{\rho}(t-\tau), \hat{\mathrm{H}}_1(t-\tau)\big], \hat{\mathrm{H}}_1(t)\big]. \tag{22.32}$$

If we rewrite this equation for $\hat{\rho}$ into an equation system for the $\hat{\rho}_{LM}$, we more or less recover (22.16) of the last section, with two differences:

1. In the first term on the right, $\delta(0)$ appears, in contrast to $\rho(0)$ in (22.16). Since this term will disappear in the next step anyway, this difference is of no importance.
2. In (22.32), in contrast to (22.16), the time dependence of $\mathrm{H}_1(t)$ is due to the transformation to the interaction system only, there are no uncontrolled fluctuations from the coupling to the bath. Instead, there is the extra exponential factor $\mathrm{e}^{-2W\tau}$, taking care of the jump process.

Applying the same approximations as in the last section, we again arrive at (22.19),

$$\dot{\hat{\rho}}_{LM} = -\lambda_{LM}\hat{\rho}_{LM}. \tag{22.33}$$

Though now we have an explicit expression for the relaxation rate

$$\lambda_{LM} = \frac{1}{\hbar^2}\int_0^\infty \mathrm{d}\tau\, \mathrm{e}^{-2W\tau}(LM|\hat{\mathrm{H}}_1(t-\tau)\hat{\mathrm{H}}_1(t)|LM). \tag{22.34}$$

This is the great advantage of the stochastic approach, as compared to the perturbative one. For the case that again H_0 describes just the interaction with an external magnetic field $\mathrm{H}_0 = \omega_0 \mathrm{I}_z$, the integration can be done explicitly,

$$\lambda_{LM} = \frac{1}{\hbar^2}\sum_{M'}\int_0^\infty \mathrm{d}t\, \mathrm{e}^{-2W\tau}(LM|\mathrm{e}^{-\mathrm{i}\omega_0\mathrm{I}_z(t-\tau)/\hbar}\mathrm{H}_1\mathrm{e}^{\mathrm{i}\omega_0\mathrm{I}_z(t-\tau)/\hbar}|LM')$$
$$\times (LM'|\mathrm{e}^{-\mathrm{i}\omega_0\mathrm{I}_z t/\hbar}\mathrm{H}_1\mathrm{e}^{\mathrm{i}\omega_0\mathrm{I}_z t/\hbar}|LM)$$

$$= \frac{1}{\hbar^2} \sum_{M'} \int_0^\infty \mathrm{d}\tau \, \mathrm{e}^{-[\mathrm{i}(M'-M)\omega_0 + 2W]\tau} (LM|\mathrm{H}_1|LM')(LM'|\mathrm{H}_1|LM)$$

$$= \frac{1}{\hbar^2} \sum_{M'} \frac{(LM|\mathrm{H}_1|LM')(LM'|\mathrm{H}_1|LM)}{\mathrm{i}(M'-M)\omega_0 + 2W}. \tag{22.35}$$

Let us now assume that H_1 corresponds to a local magnetic field in the x-direction, changing its sign stochastically,

$$\mathrm{H}_1 = \omega_x \mathrm{I}_x = \frac{1}{2}\omega_x(\mathrm{I}_+ + \mathrm{I}_-). \tag{22.36}$$

Inserting this expression into (22.35), explicit formulas for all λ_{LM} are obtained. For $\lambda_{10} = T_1^{-1}$ we have, in particular,

$$T_1^{-1} = \omega_x^2 \frac{2W}{(2W)^2 + \omega_0^2} = \omega_x^2 \frac{\tau_c}{1 + (\omega_0 \tau_c)^2}, \tag{22.37}$$

where $\tau_c = 1/(2W)$. This equation was first derived by Bloembergen, Pound and Purcell (1948) in a purely phenomenological approach, and is therefore known as BPP model.

In many cases, the time constant τ_c obeys an *Arrhenius relation*,

$$\tau_c = \tau_c^0 \mathrm{e}^{E_\mathrm{A}/k_\mathrm{B}T}. \tag{22.38}$$

E_A is the *activation energy* for the process relevant for the fluctuations. For diffusive jumps, E_A is just the energy barrier which has to be overcome in the process.

As a function of temperature λ_{10} therefore shows two different limiting cases,

$$T_1^{-1} \sim \begin{cases} \tau_c & \text{for } \omega_0 \tau_c \ll 1, \\ \frac{1}{\omega_0^2 \tau_c} & \text{for } \omega_0 \tau_c \gg 1. \end{cases} \tag{22.39}$$

A plot of $\ln(1/T_1)$ versus the reciprocal temperature, called an *Arrhenius plot*, thus will first show a linear increase with increasing $1/T$, and after having attained its maximum, a linear decrease with slopes given by $\pm E_A/k_\mathrm{B}$, see Fig. 22.1. These general features of the temperature dependence of T_1^{-1}, in particular, the limiting cases (22.39) are stable and do not depend on the details of the jump model. Therefore, this model usually gives a good description of the experimental data, though the two-site model is valid only in exceptional cases. A detailed analysis, in particular, of the region close to the maximum offers the chance to get some information on the true jump model, see Heitjans et al. (1985) for an example.

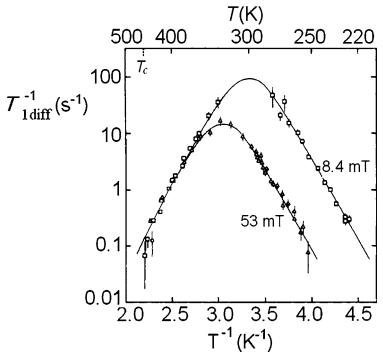


Fig. 22.1 Diffusion induced spin lattice relaxation rate $1/T_{\text{diff}}$ of ^{8}Li in ^{7}Li as a function of reciprocal temperature $1/T$, for two values of the external magnetic field. The *full curves* represent a fit based on the encounter model for self-diffusion via mono- and divacancies assuming purely dipolar relaxation. From Heitjans et al. (1985)

22.4 Decoherence

After 80 years of experience, there are no longer any doubts that quantum mechanics within its limits of validity is the correct theory to describe the world on the microscopic level of atoms and nuclei. And still there remains a debate from the very beginning, which has not been settled until today, on the "meaning" of quantum mechanics. The main problem is the act of measuring. Quantum mechanics, for example, gives the probability density to find an electron at a given time at a given site. But in the moment a detector performs a click, we know with certainty that the electron had been just there and nowhere else. What is really happening in this moment? The term "collapse of the wave function" has been coined to describe the situation, though this is nothing but a euphemistic way to say "We do not know what is really going on".

One may argue "Who cares?". Quantum mechanics is a theory making statements on statistical averages only, and it works. The question for "the meaning of" is irrelevant in this context. One may even argue that this is not a physical question, and turn to more interesting subjects.

It was Schrödinger who illustrated with his maltreated cat that there is a real problem. He proposed to lock the cat in a box together with a flask filled with hydrocyanic acid, and a radioactive substance with a low activity, say, with a 50 percent probability of one decay per hour. In the moment a radioactive decay is registered in a Geiger counter, the flask is destroyed thus killing the cat. After one hour the box is opened again. Common sense tells us that with 50 percent probability the cat is still alive, however, with 50 percent probability the cat will be dead.

What is the quantum mechanical treatment of the same situation? Since there is a quantum mechanical probability of the decay to happen or not, the wave function for the combined system sample-counter is a superposition of the two states $|n, 0\rangle$ and $|n-1, 1\rangle$, where n is the number of radioactive nuclei, and figures 0, 1 refer to the state of the counter (no click, or a click). This maps directly onto a quantum

mechanical probability of the cat to be dead or alive. Hence we may identify the sample-counter states $|n, 0\rangle$ and $|n-1, 1\rangle$ with the corresponding cat states $|\text{alive}\rangle$ and $|\text{dead}\rangle$, respectively. In the beginning, the cat is in the state $|\text{alive}\rangle$, however, with time passing there are increasing admixtures of the state $|\text{dead}\rangle$, while the weight of the $|\text{alive}\rangle$ state decreases. Just before the reopening of the box the wave function of the cat is a coherent superposition of a dead cat and a living cat with equal weights,

$$\psi_{\text{cat}} = \big(|\text{dead}\rangle + |\text{alive}\rangle\big)/\sqrt{2}. \tag{22.40}$$

In the moment of the reopening, corresponding to the act of measuring, the state collapses immediately with equal probability into one of the two alternatives $|\text{dead}\rangle$ or $|\text{alive}\rangle$. Common sense is reluctant to accept such an scenario, all the more so as the cat experiences a completely different situation: It probably will not like the situation of being locked in a dark box, but if it is lucky it will leave the box after an hour alive, without any awareness of having been half dead meanwhile. And if it is unlucky, it will be killed in a moment meaning that at any given time the cat is either in the state $|\text{dead}\rangle$ or the state $|\text{alive}\rangle$, but at no time experiences a coherent superposition!

So what is going wrong? The solution to this problem had been proposed by H.D. Zeh in a number of papers in the 1970s. He showed that macroscopic quantum mechanical coherences cannot survive for long times since due to coupling to the environment interferences are damped out very rapidly. The reader is referred to the monograph by Joos et al. (1996) for details and the relevant literature. In particular, Chap. 3 presents a very useful introduction into the fundamental ideas. The starting point is the same Hamiltonian

$$\mathrm{H} = \mathrm{H_S} + \mathrm{H_{SB}} + \mathrm{H_B} \tag{22.41}$$

that we have already used to describe relaxation processes, see Sect. 22.1. Here $\mathrm{H_S}$ is the Hamiltonian of the system we are interested in, $\mathrm{H_B}$ the bath Hamiltonian describing the surroundings, and $\mathrm{H_{SB}}$ the Hamiltonian for the interaction between system and surroundings. The total density matrix ρ_{SB} for the system coupled to the surroundings obeys the Liouville equation

$$\dot{\rho}_{\mathrm{SB}} = \frac{\mathrm{i}}{\hbar}[\rho_{\mathrm{SB}}, \mathrm{H_S} + \mathrm{H_{SB}} + \mathrm{H_B}]. \tag{22.42}$$

The system density matrix ρ_{S} that we are really interested in is obtained from ρ_{SB} by taking the average over the bath, i.e., one has to trace out the bath degrees of freedom

$$\rho_{\mathrm{S}} = \mathrm{Tr_B}(\rho_{\mathrm{SB}}). \tag{22.43}$$

By this trace operation, the unitary evolution of ρ_{SB} is turned into a subunitary evolution for ρ_{S}. This is essentially the same path we followed to derive, e.g., the Bloch equations, introduced originally phenomenologically to describe relaxation in NMR experiments, and one may wonder where there is the difference between

relaxation and decoherence. In fact, for microscopic systems, there is *no* difference. In NMR, for example, decoherence and transversal relaxation mean just the same, namely, the decay of the transversal magnetization to zero with time constant T_2. But we shall see that for a macroscopic system decoherence times can become extremely short, much shorter than any relaxation time ever observed, and here it makes sense to use different terms for the two phenomena though the underlying mechanism is the same in both cases.

This shall be illustrated by means of a model discussed in a paper by Strunz et al. (2003). The explicit forms of the system and the bath Hamiltonian will turn out to be irrelevant, but let us imagine that $\mathrm{H_S}$ is just the ordinary one-particle Hamiltonian in one dimension with a kinetic and potential energy term, and that the bath Hamiltonian $\mathrm{H_B}$ is provided by the thermal radiation of the surroundings. For the interaction Hamiltonian, we take the form

$$\mathrm{H_{SB}} = \mathrm{x}\mathrm{F_B}, \tag{22.44}$$

where x is the coordinate, and the force operator $\mathrm{F_B}$ depends on the bath variables only. Proceeding exactly as in the relaxation section we remove system and bath Hamiltonian by means of the substitution

$$\hat{\rho}_{\mathrm{SB}} = \mathrm{e}^{\mathrm{i}\mathrm{H_S}t/\hbar}\mathrm{e}^{\mathrm{i}\mathrm{H_B}t/\hbar}\rho_{\mathrm{SB}}\mathrm{e}^{-\mathrm{i}\mathrm{H_B}t/\hbar}\mathrm{e}^{-\mathrm{i}\mathrm{H_S}t/\hbar}. \tag{22.45}$$

$\hat{\rho}_{\mathrm{SB}}$ obeys the Liouville equation

$$\dot{\hat{\rho}}_{\mathrm{SB}} = \frac{\mathrm{i}}{\hbar}\big[\hat{\rho}_{\mathrm{SB}}, \hat{\mathrm{H}}_{\mathrm{SB}}(t)\big]. \tag{22.46}$$

Since $\mathrm{H_S}$ in (22.45) acts only on x, and $\mathrm{H_B}$ only on $\mathrm{F_B}$, $\hat{\mathrm{H}}_{\mathrm{SB}}(t)$ is given by

$$\hat{\mathrm{H}}_{\mathrm{SB}}(t) = \hat{\mathrm{x}}(t)\hat{\mathrm{F}}_{\mathrm{B}}(t), \tag{22.47}$$

with

$$\hat{\mathrm{x}}(t) = \mathrm{e}^{\mathrm{i}\mathrm{H_S}t/\hbar}\mathrm{x}\mathrm{e}^{-\mathrm{i}\mathrm{H_S}t/\hbar}, \tag{22.48}$$

$$\hat{\mathrm{F}}_{\mathrm{B}}(t) = \mathrm{e}^{\mathrm{i}\mathrm{H_B}t/\hbar}\mathrm{F_B}\mathrm{e}^{-\mathrm{i}\mathrm{H_B}t/\hbar}. \tag{22.49}$$

In the discussion of relaxation in the preceding sections, three time scales have appeared: (i) a typical time τ_{sys} for the isolated system, e.g., the precession period about an external magnetic field, (ii) a dissipation time τ_{diss} for the thermal relaxation of the system towards its equilibrium, and (iii) a time τ_{bath} characteristic for the thermalization within the bath. The latter did not appear explicitly before, since we always assumed a thermal equilibrium for the bath, corresponding to an infinitely small τ_{bath}. We thus had the hierarchy $\tau_{\mathrm{bath}} \ll \tau_{\mathrm{sys}} \ll \tau_{\mathrm{diss}}$ of time scales. Now another time scale τ_{dec} for decoherence comes into play. It is not obvious where to arrange it in the hierarchy. The situation usually discussed in literature is the golden rule limit corresponding to $\tau_{\mathrm{bath}} \ll \tau_{\mathrm{sys}} \ll \tau_{\mathrm{dec}} \ll \tau_{\mathrm{diss}}$, meaning that the system undergoes many cycles before decoherence takes place. For this case, the treatment is

more or less the same as we applied in the previous sections to describe relaxation. But following Strunz et al. (2003), we shall discuss another limit here, namely, that the decoherence time is the shortest one of all time scales, even shorter than the bath relaxation time. In this case, we may neglect all time dependencies of $\hat{\mathrm{H}}_{\mathrm{SB}}(t)$, and the Liouville equation (22.46) can be immediately integrated,

$$\hat{\rho}_{\mathrm{SB}}(t) = \mathrm{e}^{-\mathrm{i}x\mathrm{F_B}t/\hbar}\rho_{\mathrm{SB}}(0)\mathrm{e}^{\mathrm{i}x\mathrm{F_B}t/\hbar}. \tag{22.50}$$

A slightly less stringent approximation had been applied in Strunz et al. (2003). On the level of the applied approximations, the difference between $\hat{\rho}_{\mathrm{SB}}(t)$ and $\rho_{\mathrm{SB}}(t)$ in (22.45) vanishes, hence (22.50) holds for $\rho_{\mathrm{SB}}(t)$ as well.

Now let us assume a two-site model for the system, as it is realized, e.g., with the double-well potential for the ammonia molecule, see Sect. 12.4.1. Then we have a 2×2 density matrix

$$\rho_{\mathrm{SB}} = \begin{pmatrix} \rho_{\mathrm{LL}} & \rho_{\mathrm{LR}} \\ \rho_{\mathrm{RL}} & \rho_{\mathrm{RR}} \end{pmatrix}, \tag{22.51}$$

where indices "L" and "R" refer to the left and the right site of Fig. 12.4, respectively. Equation (22.50) with $\mathrm{F_B} = F_{\mathrm{B}}\sigma_z$ gives for the time evolution

$$\rho_{\mathrm{SB}}(t) = \begin{pmatrix} \rho_{\mathrm{LL}}(0) & \rho_{\mathrm{LR}}(0)\mathrm{e}^{-\mathrm{i}\Delta x F_{\mathrm{B}}t/\hbar} \\ \rho_{\mathrm{RL}}(0)\mathrm{e}^{\mathrm{i}\Delta x F_{\mathrm{B}}t/\hbar} & \rho_{\mathrm{RR}}(0) \end{pmatrix}, \tag{22.52}$$

where $\Delta x = x_{\mathrm{L}} - x_{\mathrm{R}}$ is the distance between the two sites.

Now we are in the position to remove the bath variables. Typically, there are a large number of independent contributions to F_{B}. In the case the bath is provided by the thermal background radiation, for example, there is a broad frequency spectrum with each frequency contributing to the interaction. This means that we have to take the average over $\exp(-\mathrm{i}\Delta x F_{\mathrm{B}}t/\hbar)$ with respect to F_{B}. Due to the central limit theorem, it is reasonable to assume that F_{B} is Gaussian distributed, allowing for a straightforward calculation of the average,

$$\left\langle \mathrm{e}^{-\mathrm{i}\Delta x F_{\mathrm{B}}t/\hbar} \right\rangle = \mathrm{e}^{-\langle F_{\mathrm{B}}^2\rangle(\Delta x)^2 t^2/(2\hbar^2)}, \tag{22.53}$$

where the brackets mean the average over the bath variables.

Inserting the result into (22.52), we end up with

$$\rho_{\mathrm{SB}}(t) = \begin{pmatrix} \rho_{\mathrm{LL}}(0) & \rho_{\mathrm{LR}}(0)\exp[-\frac{1}{2}(t/\tau_{\mathrm{dec}})^2] \\ \rho_{\mathrm{RL}}(0)\exp[-\frac{1}{2}(t/\tau_{\mathrm{dec}})^2] & \rho_{\mathrm{RR}}(0) \end{pmatrix}, \tag{22.54}$$

where

$$\tau_{\mathrm{dec}}^{-1} = \frac{1}{\hbar}\sqrt{\langle F_{\mathrm{B}}^2\rangle}|\Delta x|. \tag{22.55}$$

Equation (22.54) constitutes the main result of this section. It shows that the off-diagonal elements of the density matrix in the position representation exhibit a Gaussian decay with the time constant τ_{dec}, and after some decoherence times only an incoherent superposition of the left and the right state will survive.

For a quantitative calculation, assumptions on the interaction have to be made. For the thermal background radiation as the bath, for example, we may substitute for F_{B} the force exerted by the electric field of amplitude E of the background radiation onto the particle, $F_{\mathrm{B}} = NeE$, where N is the number of charged particles participating in the interaction, typically the valence electrons of the atoms. Hence N corresponds approximately to the number of atoms in the particle (or to the number of surface atoms in case only those atoms participate). It follows for the reciprocal decoherence time

$$\tau_{\mathrm{dec}}^{-1} = \frac{Ne}{\hbar}\sqrt{\langle E^2\rangle}|\Delta x|. \tag{22.56}$$

For the background radiation, $\langle E^2\rangle$ can be calculated from Planck's radiation formula:

The variance of the electric field corresponds, up to the factor ε_0, to the volume density of the electromagnetic radiation. For a thermal emitter Planck's radiation formula, on the other hand, yields for this quantity, up to a factor of the order of one, $(k_{\mathrm{B}}T)^4/(\hbar c)^3$, or $\langle E^2\rangle \sim (k_{\mathrm{B}}T)^4/[\varepsilon_0(\hbar c)^3]$, or

$$F_{\mathrm{B}} = e\sqrt{\langle E^2\rangle} \sim \sqrt{4\pi\alpha}(k_{\mathrm{B}}T)^2/\hbar c \approx 10^3 \text{ eV/m}, \tag{22.57}$$

with $k_{\mathrm{B}}T = 0.025$ eV for room temperature, and $\alpha = e^2/\varepsilon_0 4\pi\hbar c = 1/137$. Arranging the factors conveniently, we obtain

$$\tau_{\mathrm{dec}}^{-1} \sim \frac{N}{\hbar}\frac{|\Delta x|}{a_{\mathrm{B}}}F_{\mathrm{B}}a_{\mathrm{B}} \approx 7\times 10^7\, N\frac{|\Delta x|}{a_{\mathrm{B}}}\ \mathrm{s}^{-1}, \tag{22.58}$$

where the Bohr radius $a_{\mathrm{B}} = \hbar c/\alpha m_e c^2 = 0.5\times 10^{-10}$ m has been introduced as a typical order of magnitude for a distance on the atomic scale.

One thus obtains, e.g., for the ammonia molecule at room temperature a decoherence time of the order of 10^{-8} s.

In view of the tunneling rate of 24 GHz found in NH_3 (see Sect. 12.4) and frequencies of 10^{13} s^{-1} of the thermal background radiation, this rate is by far too small to lead to a decoherence of the tunneling process. Indeed, the molecule is found in the experiment either in a symmetric or antisymmetric state, but not with the nitrogen atom in one of the two potential wells. Note that the potential wells are not really needed for this consideration since the potential does not enter at the level of the applied approximation. It is only distance that matters!

The situation changes dramatically if we increase the distance between the two positions or the particle's mass to daily life scales. Then we immediately arrive at

decoherence times smaller by orders of magnitudes then any other relaxation time, be it in the system or the bath.

One should be aware that decoherence does not solve the measurement problem. For Schrödinger's cat, it is still a matter of luck whether it has an escape or not. But decoherence explains why quantum mechanical coherences are not observed in the classical world. There is no such weird situation like a "⟨dead| and ⟨alive|" for the cat, decoherence immediately turns the state into a "⟨dead| or ⟨alive|".

Is this the solution to the cat problem? We have some doubts. For very short times, the coherence still exists; and also for longer times, the coherence still is, though tiny, not equal to zero. From the practical point of view, one might argue that these remaining effects are unobservable, but is this a sufficiently strong argument to discard this point?

References

Abragam, A.: The Principles of Nuclear Magnetism. Oxford University Press, Oxford (1961)

Bloembergen, N., Purcell, E.M., Pound, R.V.: Relaxation effects in nuclear magnetic resonance absorption. Phys. Rev. **73**, 679–709 (1948)

Breuer, H.-P., Petruccione, F.: The Theory of Open Quantum Systems. Oxford University Press, Oxford (2002)

Dattagupta, S.: Relaxation Phenomena in Condensed Matter Physics. Academic Press, New York (1987)

Heitjans, P., Körblein, A., Ackermann, H., Dubbers, D., Fujara, F., Stöckmann, H.-J.: Self-diffusion of solid lithium probed by spin-lattice relaxation of ^{8}Li nuclei. J. Phys. F **15**, 41–54 (1985)

Joos, E., Zeh, H.D., Kiefer, C., Kupsch, J., Stamatescu, I.-O.: Decoherence and the Appearance of a Classical World in Quantum Theory. Springer, Berlin (1996)

Strunz, W.T., Haake, F., Braun, D.: Universality of decoherence for macroscopic quantum superpositions. Phys. Rev. A **67**, 022101(13) (2003)

Van Kampen, N.G.: Stochastic Processes in Physics and Chemistry. North-Holland, Amsterdam (1981)

This page intentionally left blank

Appendices

A.1 The Orthogonality of the Irreducible Tensor Operators

To prove the orthogonality relation (19.22) for irreducible tensors, we start with the relation (17.10a),

$$[\mathrm{L}_z, \mathrm{T}_{LM}] = \hbar M \mathrm{T}_{LM}. \tag{A.1}$$

Taking the adjoint and replacing L, M by L', M', we get

$$\left[\mathrm{T}^{\dagger}_{L'M'}, \mathrm{L}_z\right] = \hbar M' \mathrm{T}^{\dagger}_{L'M'}. \tag{A.2}$$

Multiplying the first equation from the left by $\mathrm{T}^{\dagger}_{L'M'}$, the second from the right by T_{LM}, and taking the difference, we obtain

$$\mathrm{T}^{\dagger}_{L'M'}[\mathrm{L}_z, \mathrm{T}_{LM}] - \left[\mathrm{T}^{\dagger}_{L'M'}, \mathrm{L}_z\right]\mathrm{T}_{LM} = \hbar\big(M - M'\big)\mathrm{T}^{\dagger}_{L'M'}, \mathrm{T}_{LM}. \tag{A.3}$$

Taking the trace on both sides, it follows with the commutativity of the trace operation

$$\big(M - M'\big)\,\mathrm{Tr}\big(\mathrm{T}^{\dagger}_{L'M'}\mathrm{T}_{LM}\big) = 0. \tag{A.4}$$

In a similar way, from the commutation relation $[\mathbf{L}^2, \mathrm{T}_{LM}] = \hbar^2 L(L+1)\mathrm{T}_{LM}$ one obtains

$$\big[L(L+1) - L'\big(L'+1\big)\big]\,\mathrm{Tr}\big(\mathrm{T}^{\dagger}_{L'M'}\mathrm{T}_{LM}\big) = 0. \tag{A.5}$$

From (A.4) and (A.5) we obtain the wanted orthogonality relation

$$\mathrm{Tr}\big(\mathrm{T}^{\dagger}_{L'M'}\mathrm{T}_{LM}\big) = \delta_{LL'}\delta_{ML'}\,\mathrm{Tr}\big(\mathrm{T}^{\dagger}_{LM}\mathrm{T}_{LM}\big). \tag{A.6}$$

Next we show that this expression in fact is independent of M. To this end, we take (17.10b) for L_+,

$$[\mathrm{L}_+, \mathrm{T}_{LM}] = \hbar\sqrt{L(L+1) - M(M+1)}\,\mathrm{T}_{L,M+1}, \tag{A.7}$$

D. Dubbers, H.-J. Stöckmann, *Quantum Physics: The Bottom-Up Approach*, Graduate Texts in Physics,
DOI 10.1007/978-3-642-31060-7, © Springer-Verlag Berlin Heidelberg 2013

and for L_-, the latter one with M replaced by $M+1$,

$$[\mathrm{L}_-, \mathrm{T}_{L,M+1}] = \hbar\sqrt{L(L+1) - M(M+1)}\,\mathrm{T}_{LM}. \tag{A.8}$$

Multiplying the first equation by $\mathrm{T}^\dagger_{L,M+1}$, the adjoint of the second one by T_{LM}, taking the difference and then the trace, we obtain in the same way as above

$$\sqrt{L(L+1) - M(M+1)}\,\mathrm{Tr}\big(\mathrm{T}^\dagger_{L,M+1}\mathrm{T}_{L,M+1} - \mathrm{T}^\dagger_{LM}\mathrm{T}_{LM}\big) = 0, \tag{A.9}$$

or

$$\mathrm{Tr}\big(\mathrm{T}^\dagger_{L,M+1}\mathrm{T}_{L,M+1}\big) = \mathrm{Tr}\big(\mathrm{T}^\dagger_{LM}\mathrm{T}_{LM}\big), \quad -L \le M < L. \tag{A.10}$$

$\mathrm{Tr}(\mathrm{T}^\dagger_{LM}\mathrm{T}_{LM})$ is thus independent of M as stated.

A.2 The Reduced Matrix Element $\langle j||\hat{\mathrm{T}}_L(j)||j\rangle$

The reduced matrix element (19.26) (in angular brackets) is obtained from the normalization condition $\mathrm{Tr}[\hat{\mathrm{T}}^\dagger_{LM}(j)\hat{\mathrm{T}}_{LM}(j)] = 1$. By means of the Wigner–Eckart theorem, this trace can be calculated:

$$\begin{aligned}
\mathrm{Tr}\big[\hat{\mathrm{T}}^\dagger_{LM}(j)\hat{\mathrm{T}}_{LM}(j)\big] &= \sum_{m,m'} \langle jm|\hat{\mathrm{T}}^\dagger_{LM}(j)|jm'\rangle\langle jm'|\hat{\mathrm{T}}_{LM}(j)|jm\rangle \\
&= \sum_{mm'} \big|\langle jm'|\hat{\mathrm{T}}_{LM}(j)|jm\rangle\big|^2 \\
&= (2j+1)^{-1}\big|\langle j||\hat{\mathrm{T}}_L(j)||j\rangle\big|^2 \sum_{mm'} \big|\langle jmLM|jm'\rangle\big|^2 \\
&= (2j+1)^{-1}\big|\langle j||\hat{\mathrm{T}}_L(j)||j\rangle\big|^2 \frac{2j+1}{2L+1} \\
&\quad \times \sum_{mm'} \big|\langle jmjm'|L,-M\rangle\big|^2.
\end{aligned} \tag{A.11}$$

In the last equality, the symmetry relation (17.24c) for the Clebsch–Gordan coefficients was used (with a subsequent substitution of m' by $-m'$). The double sum over the Clebsch–Gordan coefficients equals one because of the orthogonality relation (17.25a). Collecting the results, we get

$$\langle j||\hat{\mathrm{T}}_L(j)||j\rangle = \pm\sqrt{2L+1}. \tag{A.12}$$

The sign by convention is fixed to be positive.

A.3 The Reduced Matrix Element $(L||\hat{\mathbf{T}}_{L_2}(j)||L_1)$

To calculate this reduced matrix element (19.41) (in round brackets), we need an expression for the tensor product of two normalized irreducible tensor operators in terms of another normalized irreducible tensor operator. Our starting point is (17.29). Here we run into the problem that the thus constructed new tensor will not be automatically normalized. Hence we make the ansatz

$$N_j(L_1L_2L)\hat{\mathrm{T}}_{LM}(j) = \sum_{M_1M_2} \langle L_1M_1L_2M_2|LM\rangle \hat{\mathrm{T}}_{L_1M_1}(j)\hat{\mathrm{T}}_{L_2M_2}(j) \tag{A.13}$$

and determine the factor $N_j(L_1L_2L)$ such that $\hat{\mathrm{T}}_{LM}(j)$ is normalized. Multiplying both sides with $\hat{\mathrm{T}}^{\dagger}_{LM}(j)$ and taking the trace, we obtain

$$N_j(L_1L_2L) = \sum_{M_1M_2} \langle L_1M_1L_2M_2|LM\rangle \mathrm{Tr}\big[\hat{\mathrm{T}}^{\dagger}_{LM}(j)\hat{\mathrm{T}}_{L_1M_1}(j)\hat{\mathrm{T}}_{L_2M_2}(j)\big]. \tag{A.14}$$

We know already that the normalization constant is independent on the M quantum number. Therefore, the right-hand side of the equation must be independent of M even though M enters. We may therefore sum the right-hand side over M, meaning nothing but taking the same term $2L+1$ times, and dividing the result by $2L+1$ to compensate for this fact. Writing the trace explicitly in terms of sums over matrix elements, we obtain

$$\begin{aligned} N_j(L_1L_2L) = \frac{1}{2L+1} \sum_{MM_1M_2m_1m_2m_3} & \langle L_1M_1L_2M_2|LM\rangle \langle jm_1|\hat{\mathrm{T}}^{\dagger}_{LM}(j)|jm_2\rangle \\ & \times \langle jm_2|\hat{\mathrm{T}}_{L_1M_1}(j)|jm_3\rangle \langle jm_3|\hat{\mathrm{T}}_{L_2M_2}(j)|jm_1\rangle. \end{aligned} \tag{A.15}$$

Next each tensor matrix element is expressed in terms of Clebsch–Gordan coefficients using the Wigner–Eckart theorem for normalized tensors (19.27) yielding

$$\begin{aligned} N_j(L_1L_2L) = & \sqrt{\frac{(2L_1+1)(2L_2+1)}{(2L+1)(2j+1)^3}} \\ & \times \sum_{MM_1M_2m_1m_2m_3} \langle L_1M_1L_2M_2|LM\rangle \langle jm_2LM|jm_1\rangle \\ & \times \langle jm_3L_1M_1|jm_2\rangle \langle jm_1L_2M_2|jm_3\rangle. \end{aligned} \tag{A.16}$$

It is here where the $6j$ symbol comes into play. It is defined as

$$\begin{Bmatrix} J_1 & J_2 & J_3 \\ j_1 & j_2 & j_3 \end{Bmatrix} = \frac{(-1)^{J_1+J_2+j_1+j_2}}{\sqrt{(2J_3+1)(2j_3+1)}} \sum_{M_1M_2M_3m_1m_3} \langle J_1M_1J_2M_2|J_3M_3\rangle$$

$$\times \langle J_1 M_1 j_3 m_3 | j_2 m_2 \rangle \langle J_2 M_2 j_1 m_1 | j_3 m_3 \rangle$$
$$\times \langle J_3 M_3 j_1 m_1 | j_2 m_2 \rangle. \tag{A.17}$$

Note that there is no sum over m_2, but it can be shown that the term on the right-hand side is, in fact, independent in m_2. The situation is similar as in (A.14), and as above we may sum the expression on the right-hand side over m_2 and divide the result by $2j_2 + 1$ to compensate for this. If this is done, we recognize that, after changing the sequence in the arguments of the Clebsch–Gordan coefficients using the symmetry relations (A.28), the sum over the M's in (A.16) can be expressed in terms of a special $6j$ symbol:

$$N_j(L_1 L_2 L) = (-1)^{2j+L} \sqrt{(2L_1+1)(2L_2+1)} \begin{Bmatrix} L_1 & L_2 & L \\ j & j & j \end{Bmatrix}. \tag{A.18}$$

Taking the inverse of (A.13), we get

$$\hat{\mathrm{T}}_{L_1 M_1}(j) \hat{\mathrm{T}}_{L_2 M_2}(j) = \sum_{LM} N_j(L_1 L_2 L) \langle L_1 M_1 L_2 M_2 | LM \rangle \hat{\mathrm{T}}_{LM}(j). \tag{A.19}$$

This is used to rewrite the commutator on the right-hand side of the definition (19.32) of the tensor matrix element, with $\mathrm{A} = \hat{\mathrm{T}}_{L_2 M_2}$,

$$\begin{aligned}
(LM|\hat{\mathrm{T}}_{L_2 M_2}|L_1 M_1) &= \mathrm{Tr}\big\{\hat{\mathrm{T}}^\dagger_{LM}(j)\big[\hat{\mathrm{T}}_{L_2 M_2}(j), \hat{\mathrm{T}}_{L_1 M_1}(j)\big]\big\} \\
&= \sum_{\bar{L},\bar{M}} \big[(-1)^{L_1+L_2-\bar{L}} - 1\big] N_j(L_1 L_2 L) \langle L_1 M_1 L_2 M_2 | \bar{L}\bar{M} \rangle \mathrm{Tr}\big[\hat{\mathrm{T}}^\dagger_{LM}(j) \hat{\mathrm{T}}_{\bar{L}\bar{M}}(j)\big] \\
&= \big[(-1)^{L_1+L_2-L} - 1\big] N_j(L_1 L_2 L) \langle L_1 M_1 L_2 M_2 | LM \rangle.
\end{aligned} \tag{A.20}$$

In the second equality, we used the symmetry relations for the Clebsch–Gordan coefficients, and in the last equality the orthonormality (19.22) with (19.23) of the $\hat{\mathrm{T}}_{LM}$. Comparison of (A.20) with (19.40) yields immediately expression (19.41) for the reduced matrix element.

A.4 Spherical Harmonics

The Spherical Harmonics for $L = 0, 1, 2$ are

$$l = 0: \quad Y_{00}(\theta, \varphi) = \frac{1}{\sqrt{4\pi}}. \tag{A.21}$$

$$l = 1: \quad Y_{10}(\theta, \varphi) = \sqrt{\frac{3}{4\pi}} \cos\vartheta, \tag{A.22a}$$

$$Y_{1\pm1}(\theta, \varphi) = \mp\sqrt{\frac{3}{8\pi}} \sin\theta \, \mathrm{e}^{\pm \mathrm{i}\varphi}. \tag{A.22b}$$

$$l = 2: \quad Y_{20}(\theta, \varphi) = \sqrt{\frac{5}{16\pi}}\big(3\cos^2\theta - 1\big), \tag{A.23a}$$

$$Y_{2,\pm1}(\theta, \varphi) = \mp\sqrt{\frac{15}{8\pi}}\sin\theta\cos\theta\,\mathrm{e}^{\pm\mathrm{i}\varphi}, \tag{A.23b}$$

$$Y_{2\pm2}(\theta, \varphi) = \sqrt{\frac{15}{32\pi}}\sin^2\mathrm{e}^{\pm2\mathrm{i}\varphi}. \tag{A.23c}$$

A.5 Rotation Matrices

Given below are the elements $\{d^{(j)}_{mm'}\}$ with $m = m' = +j$ in the upper left, and with $m = m' = -j$ in the lower right of the matrix:

$$\mathrm{d}^{(1/2)}(\beta) = \begin{pmatrix} \cos\frac{1}{2}\beta & -\sin\frac{1}{2}\beta \\ \sin\frac{1}{2}\beta & \cos\frac{1}{2}\beta \end{pmatrix}, \tag{A.24}$$

$$\mathrm{d}^{(1)}(\beta) = \begin{pmatrix} \frac{1}{2}(1+\cos\beta) & -\sqrt{\frac{1}{2}}\sin\beta & \frac{1}{2}(1-\cos\beta) \\ \sqrt{\frac{1}{2}}\sin\beta & \cos\beta & -\sqrt{\frac{1}{2}}\sin\beta \\ \frac{1}{2}(1-\cos\beta) & \sqrt{\frac{1}{2}}\sin\beta & \frac{1}{2}(1+\cos\beta) \end{pmatrix}, \tag{A.25}$$

$$\mathrm{d}^{(3/2)}(\beta) = \begin{pmatrix} \cos^3\frac{1}{2}\beta & -\sqrt{3}\sin\frac{1}{2}\beta\cos^2\frac{1}{2}\beta & \sqrt{3}\sin^2\frac{1}{2}\beta\cos\frac{1}{2}\beta & -\sin^3\frac{1}{2}\beta \\ \sqrt{3}\sin\frac{1}{2}\beta\cos^2\frac{1}{2}\beta & \cos\frac{1}{2}\beta(3\cos^2\frac{1}{2}\beta - 2) & \sin\frac{1}{2}\beta(3\sin^2\frac{1}{2}\beta - 2) & \sqrt{3}\sin^2\frac{1}{2}\beta\cos\frac{1}{2}\beta \\ \sqrt{3}\sin^2\frac{1}{2}\beta\cos\frac{1}{2}\beta & -\sin\frac{1}{2}\beta(3\sin^2\frac{1}{2}\beta - 2) & \cos\frac{1}{2}\beta(3\cos^2\frac{1}{2}\beta - 2) & -\sqrt{3}\sin\frac{1}{2}\beta\cos^2\frac{1}{2}\beta \\ \sin^3\frac{1}{2}\beta & \sqrt{3}\sin^2\frac{1}{2}\beta\cos\frac{1}{2}\beta & \sqrt{3}\sin\frac{1}{2}\beta\cos^2\frac{1}{2}\beta & \cos^3\frac{1}{2}\beta \end{pmatrix}, \tag{A.26}$$

$$
\mathrm{d}^{(2)}(\beta) = \begin{pmatrix}
\frac{1}{4}(1+\cos\beta)^2 & -\frac{1}{2}\sin\beta(1+\cos\beta) & \sqrt{3/8}\sin^2\beta & -\frac{1}{2}\sin\beta(1-\cos\beta) & \frac{1}{4}(1-\cos\beta)^2 \\
\frac{1}{2}\sin\beta(1+\cos\beta) & \cos^2\beta - \frac{1}{2}(1-\cos\beta) & -\sqrt{3/2}\sin\beta\cos\beta & \frac{1}{2}(1+\cos\beta) - \cos^2\beta & -\frac{1}{2}\sin\beta(1-\cos\beta) \\
\sqrt{3/8}\sin^2\beta & \sqrt{3/2}\sin\beta\cos\beta & \frac{1}{2}(3\cos^2\beta - 1) & -\sqrt{3/2}\sin\beta\cos\beta & \sqrt{3/8}\sin^2\beta \\
\frac{1}{2}\sin\beta(1-\cos\beta) & \frac{1}{2}(1+\cos\beta) - \cos^2\beta & \sqrt{3/2}\sin\beta\cos\beta & \cos^2\beta - \frac{1}{2}(1-\cos\beta) & -\frac{1}{2}\sin\beta(1+\cos\beta) \\
\frac{1}{4}(1-\cos\beta)^2 & \frac{1}{2}\sin\beta(1-\cos\beta) & \sqrt{3/8}\sin^2\beta & \frac{1}{2}\sin\beta(1+\cos\beta) & \frac{1}{4}(1+\cos)^2
\end{pmatrix}. \tag{A.27}
$$

A.6 Clebsch–Gordan Coefficients

Tables of $\langle j_1 m_1 j_2 m_2 | j m\rangle$ are given for the case where j_2 takes the value 0, $\frac{1}{2}$, or 1. With the symmetry relations reproduced below from (17.24a)–(17.24c), the tables can be extended to all cases, where one of the three angular momenta involved in a Clebsch–Gordan coefficient is ≤ 1:

$$
\begin{aligned}
\langle j_1\ m_1\ j_2\ m_2 | j\ m\rangle &= (-1)^{j_1+j_2-j}\langle j_1 - m_1\, j_2 - m_2 | j - m\rangle \\
&= (-1)^{j_1+j_2-j}\langle j_2 m_2\, j_1 m_1 | j m\rangle \\
&= (-1)^{j_1-m_1}\sqrt{\frac{2j+1}{2j_2+1}}\langle j_1 m_1\, j - m | j_2 - m_2\rangle
\end{aligned} \tag{A.28}
$$

(a) $j_2 = 0$

$$
\langle j\ m\ 0\ 0 | j\ m\rangle = 1; \tag{A.29}
$$

(b) $j_2 = \frac{1}{2}$

$\langle j_1\ m_1\ \frac{1}{2}\ m_2 | j\ m\rangle$:

$j \setminus m_2$	$\frac{1}{2}$	$-\frac{1}{2}$
$j_1 + \frac{1}{2}$	$\sqrt{\frac{j_1+m+\frac{1}{2}}{2j_1+1}}$	$\sqrt{\frac{j_1-m+\frac{1}{2}}{2j_1+1}}$
$j_1 - \frac{1}{2}$	$-\sqrt{\frac{j_1-m+\frac{1}{2}}{2j_1+1}}$	$\sqrt{\frac{j_1+m+\frac{1}{2}}{2j_1+1}}$

(A.30)

(c) $j_2 = 1$

$\langle j_1\, m_1\, 1\, m_2 | j\, m\rangle$:

$j \setminus m_2$	1	0	-1
j_1+1	$\sqrt{\frac{(j_1+m)(j_1+m+1)}{(2j_1+1)(2j_1+2)}}$	$\sqrt{\frac{(j_1-m+1)(j_1+m+1)}{(2j_1+1)(j_1+1)}}$	$\sqrt{\frac{(j_1-m)(j_1-m+1)}{(2j_1+1)(2j_1+2)}}$
j_1	$-\sqrt{\frac{(j_1+m)(j_1-m+1)}{2j_1(j_1+1)}}$	$\frac{m}{\sqrt{j_1(j_1+1)}}$	$\sqrt{\frac{(j_1-m)(j_1+m+1)}{2j_1(j_1+1)}}$
j_1-1	$\sqrt{\frac{(j_1-m)(j_1-m+1)}{2j_1(2j_1+1)}}$	$-\sqrt{\frac{(j_1-m)(j_1+m)}{j_1(2j_1+1)}}$	$\sqrt{\frac{(j_1+m)(j_1+m+1)}{2j_1(2j_1+1)}}$

(A.31)

A.7 Normalized Irreducible Tensor Operators

The list of normalized irreducible tensor elements $\hat{\mathrm{T}}_{LM}(j)$, (19.23), with $L \le 2$ is

$$\hat{\mathrm{T}}_{00}(j) = a_0 \mathrm{I}, \tag{A.32}$$

$$\hat{\mathrm{T}}_{10}(j) = a_1 \mathrm{J}_z/\hbar, \qquad \hat{\mathrm{T}}_{1\pm1}(j) = \mp\sqrt{\frac{1}{2}} a_1 \mathrm{J}_\pm/\hbar, \tag{A.33}$$

$$\hat{\mathrm{T}}_{20}(j) = a_2\big(3\mathrm{J}_z^2 - \mathbf{J}^2\big)/\hbar^2, \qquad \hat{\mathrm{T}}_{2\pm1}(j) = \mp\frac{1}{2}\sqrt{6} a_2(\mathrm{J}_\pm \mathrm{J}_z + \mathrm{J}_z \mathrm{J}_\pm)/\hbar^2, \tag{A.34}$$

$$\hat{\mathrm{T}}_{2\pm2}(j) = \frac{1}{2}\sqrt{6} a_2 \mathrm{J}_\pm^2/\hbar^2,$$

with identity I and

$$\begin{aligned} a_0 &= \frac{1}{\sqrt{2j+1}}, \\ a_1 &= \sqrt{\frac{3}{j(j+1)(2j+1)}}, \\ a_2 &= \sqrt{\frac{20}{(2j-1)2j(2j+1)(2j+2)(2j+3)}}. \end{aligned} \tag{A.35}$$

All tensors $\hat{\mathrm{T}}_{LM}$ discussed in this book can be taken from this list.

A.8 Coefficients of the Generalized Precession Equation

The coefficients

$$c_j(L_1 L_2 L) = -\frac{1}{\sqrt{2L+1}}(L||\hat{\mathrm{T}}_{L_1}(j)||L_2), \tag{A.36}$$

see (19.61), are obtained from the reduced tensor matrix elements (19.41). The only nonvanishing matrix elements for $L_1 = 1, 2$, corresponding to magnetic dipole and electric quadrupole interactions, respectively, are

$$(L||\hat{\mathrm{T}}_1(j)||L) = a_1\sqrt{L(L+1)(2L+1)}, \tag{A.37}$$

$$\begin{aligned}(L||\hat{\mathrm{T}}_2(j)||L-1) &= -(L-1||\hat{\mathrm{T}}_2(j)||L)\\ &= a_2\sqrt{3(2j+L+1)(2j-L+1)(L-1)L(L+1)}\end{aligned} \tag{A.38}$$

with a_1, a_2 given in (A.35). It follows for the coefficients of the generalized spin precession equation (19.60)

$$c_j(1LL) = -\sqrt{\frac{3L(L+1)}{j(j+1)(2j+1)}}, \tag{A.39}$$

$$\begin{aligned}c_j(2, L, L-1) &= \sqrt{\frac{60(2j+L+1)(2j-L+1)(L-1)L(L+1)}{(2L-1)(2j-1)2j(2j+1)(2j+2)(2j+3)}}\\ &= -c_j(2, L-1, L).\end{aligned} \tag{A.40}$$

A.9 Transforming away Part of an Interaction from the Liouville Equation

Let a Hamiltonian consists of two parts, $\mathrm{H} = \mathrm{H}_1 + \mathrm{H}_2$, with time independent H_1. We can remove H_1 from the Liouville equation by applying the partial time evolution operator $\mathrm{U}_1(t) = \mathrm{e}^{-\mathrm{i}\mathrm{H}_1 t/\hbar}$ to the density operator ρ, which gives $\tilde{\rho} = \mathrm{U}_1^\dagger \rho \mathrm{U}_1$. We apply this transformation to the Liouville equation of motion of the density operator $\dot{\rho} = (\mathrm{i}/\hbar)[\rho, \mathrm{H}]$ from (19.16). Using the fact (10.16) that U_1 obeys the Schrödinger equation $\dot{\mathrm{U}}_1 = -(\mathrm{i}/\hbar)\mathrm{H}_1\mathrm{U}_1$, and the fact from Sect. 6.1 that H_1 commutes with U_1, we find

$$\begin{aligned}\dot{\tilde{\rho}} &= \dot{\mathrm{U}}_1^\dagger \rho \mathrm{U}_1 + \mathrm{U}_1^\dagger \rho \dot{\mathrm{U}}_1 + \mathrm{U}_1^\dagger \dot{\rho} \mathrm{U}_1\\ &= \frac{\mathrm{i}}{\hbar}\big(\mathrm{H}_1\tilde{\rho} - \tilde{\rho}\mathrm{H}_1 + \mathrm{U}_1^\dagger[\rho, \mathrm{H}]\mathrm{U}_1\big)\\ &= \frac{\mathrm{i}}{\hbar}\big(-[\tilde{\rho}, \mathrm{H}_1] + \big[\tilde{\rho}, \mathrm{U}_1^\dagger \mathrm{H}_2 \mathrm{U}_1 + \mathrm{H}_1\big]\big)\\ &= \frac{\mathrm{i}}{\hbar}[\tilde{\rho}, \tilde{\mathrm{H}}_2],\end{aligned} \tag{A.41}$$

with $\tilde{\mathrm{H}}_2 = \mathrm{U}_1^\dagger \mathrm{H}_2 \mathrm{U}_1$. We see that H_1 has been transformed away from the Liouville equation.

Index

D. Dubbers, H.-J. Stöckmann, *Quantum Physics: The Bottom-Up Approach*, Graduate Texts in Physics,
DOI 10.1007/978-3-642-31060-7, © Springer-Verlag Berlin Heidelberg 2013