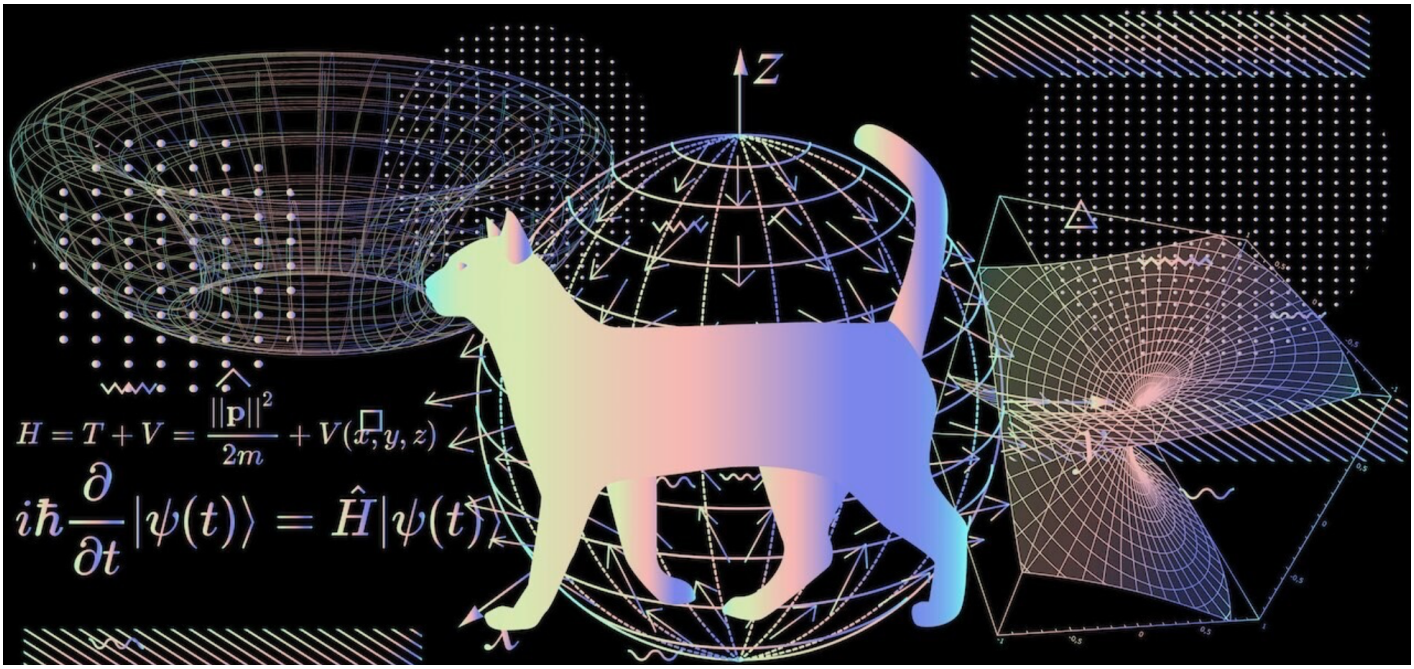
Lecture Notes on Quantum Physics II

Javier Rubio



This work is licensed under a Creative Commons "Attribution-NonCommercial-NoDerivatives 4.0 International" license.

CONTENTS

PREFACE

These notes are intended for a one-semester *Quantum Physics II* course at the Universidad Complutense de Madrid, presuming familiarity with *Quantum Physics I*, and in particular with the hydrogen atom and the harmonic oscillator. Although far from being a book, they stem from many excellent books and lecture notes, including those in the subsequent bibliography. The choice of content, the order in which the different topics are presented, and in many cases, their reformulation, are, however, my own. Feedback on the content is very welcome, especially if you identify typos or places where the text lacks clarity.[1]

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. You are permitted to copy and share this document in its entirety for noncommercial purposes; however, creating and distributing modified versions, using it for commercial purposes or uploading it in money-based platforms like Wuolah is not allowed.

The content is organized as follows. After a quick recap of wave mechanics in Chapter 1, Chapter 2 introduces the concept of spin through the so-called Stern-Gerlach experiment. Choosing this *spin first approach* has contextual and mathematical advantages. On the one hand, the inherently quantum mechanical nature of spin will force us to leave behind any reliance on classical physics' descriptions. On the other hand, the spin one-half system is one of the simplest and most instructive quantum setups, making it the perfect candidate for illustrating the foundational principles of quantum mechanics. This will be done in Chapter 3, where we will present the abstract mathematical formalism, ideas and postulates of quantum mechanics. Throughout this journey, we will discover that a plethora of unrelated physical systems and phenomena, ranging from simple molecules to neutrino and kaon oscillations, can be formally brought down to the aforementioned spin one-half setup, as explicitly illustrated in Chapter 4. A detailed treatment of composite systems, statistical mixtures and identical particles is then presented in Chapter 5. The peculiar attributes of spin will prompt us to study the algebraic theory of angular momentum in Chapter 6. Finally, Chapters 7 and 8 provide a good collection of tools and approximation methods to deal with realistic quantum mechanical systems, covering topics such bound states and relativistic quantum corrections, which we discuss in detail in Chapter 9.

Quantum mechanics, at this level, presents a dual nature: it is conceptually and computationally challenging. While this course will certainly demand a considerable effort from your side, it will also offer a delightful experience, confronting you for the very first time with the practical applications of quantum mechanics. In order to facilitate the understanding of the different topics, I have prepared a handful of supplementary materials:

[1] Please send comments and corrections to javier.rubio@ucm.es.

- **Examples:** I have made an effort to supplement the theoretical developments with detailed examples illustrating the most important results. These should, be considered, however, as examinable material.

- **Solved and proposed exercises**:

  Practicing exercises in quantum mechanics is as essential as swimming practice for learning to swim or regular training for building muscles. Just as you can't expect to master swimming by jumping into the water once, you can't fully grasp quantum mechanics by simply reading these lecture notes or passively attending my lectures. To aid your training, I've compiled a collection of over 150 exercises.

  At the end of each chapter, you will find a selection of *worked-out exercises*. These exercises come with full, detailed solutions and are intended to reinforce the chapter's main concepts, illustrate key techniques, and provide worked examples to support your learning and problem-solving skills.

  Immediately following the solved examples, you will encounter the *Proposed Exercises*. These problems are left for you to solve independently and range in difficulty from straightforward applications to more challenging questions that require a deeper understanding. These are categorized as follows:

  –  Exercises marked with this symbol should be understood as essential problems you should master to excel in the midterm test and the final examination. Some of these will be addressed during the exercise sessions.

  –  Exercises marked with this symbol typically involve deriving specific identities provided without proof during the lessons or exploring more formal aspects of the course not strictly necessary in a first encounter with the subject. On general grounds, they will not be covered during the exercise sessions.

  –  Exercises marked with this symbol often require thoughtful consideration of the problem's formulation before proceeding with calculations. Though usually too involved for an exam, some of them offer valuable insights and will be discussed during the exercise sessions.

References

There is a plethora of excellent books and online lecture notes on Quantum Mechanics, each of them emphasising different aspects of the theory, from the mathematical formalism to the physical applications. Which one to use depends only on your taste. Said that, here are some of the books and online materials I found useful while preparing these notes:

Theory Books:

1. D.J. Griffiths, *Introduction to Quantum Mechanics*, Pearson, 2005.

2. S. Gasiorowicz, *Quantum Physics*, Wiley International, 2003.

3. C. Cohen-Tannoudji, B. Diu and F. Lalo ë, *Quantum Mechanics I and II*, Wiley, 1977.

4. J.J. Sakurai, *Modern Quantum Mechanics*, Addison-Wesley, 1991.
5. S. Weinberg, *Lectures on Quantum Mechanics*, Cambridge University Press, 2013.
6. M. Le Bellac, *Quantum Physics*, Cambridge University Press, 2011.
7. K. Konishi and G. Paffuti, *Quantum Mechanics: A new introduction*, Oxford, 2009.
8. E. Merzbacher, *Quantum Mechanics*, Wiley, 1998.
9. G. Auletta, M. Fortunato and G. Parisi, *Quantum Mechanics*, Cambridge University Press, 2013.
10. B. Bransden and C. Joachain, *Quantum Mechanics*, Prentice hall, 2000.
11. F. Mandl, *Quantum Mechanics*, Wiley 1992.
12. L. Susskind and A. Friedman *Quantum Mechanics: The Theoretical Minimum*, Penguin Books, 2014.
13. F. Schwabl, *Advanced Quantum Mechanics*, Springer, 2004.
14. K.S. Lam, *Non-relativistic Quantum Theory*, World Scientific Publishing, 2009.

Mathematical purists and rigour enthusiasts should realize from the very beginning that they are enrolled in the wrong course with the wrong lecturer, being kindly invited to consult the quite advanced references:

14. J. von Neumann, Mathematical Foundations of Quantum Mechanics, 2018.
15. B.Hall, *Quantum Theory for Mathematicians*, Springer, 2013.
16. V. Moretti, *Fundamental Mathematical Structures of Quantum Theory*, Springer, 2019.

Exercise Books:

1. E. d'Emilio, *Problems in Quantum Mechanics*, Springer, 2016.
2. F. Constantinescu and E. Magyari, *Problems in Quantum Mechanics*, Pergamon Press, 1976.
3. S. Flügge, *Practical Quantum Mechanics*, Springer-Verlag, 1971.
4. H. Mavromatis, *Exercises in Quantum Mechanics*, Springer-Verlag, 1992.

Youtube Courses:

1. MIT 8.04 Quantum Physics I by Barton Zwiebach.
2. MIT 8.05 Quantum Physics II by Barton Zwiebach.
3. The Theoretical Minimum: Quantum Mechanics by Leonard Susskind.

CONTENTS

Other useful online resourses:

1. NIST Digital Library of Mathematical Functions
2. Gerald Teschl, *Mathematical Methods in Quantum Mechanics With Applications to Schrödinger Operators*, American Mathematical Society, Providence, RI, 2009.
3. The Particle Data Group.
4. The Quantum Mechanics Visualisation Project

A certainly incomplete timeline of Quantum Mechanics

Year	Event	Year	Event
1672	Newton's corpuscular theory of light	1932	Von Neumann formalizes quantum mechanics using Hilbert spaces
1678	Huygens' wave theory of light	1935	EPR paradox
1801	Young's double-slit experiment	1935	Schrödinger's cat
1887	Hertz observes the photoelectric effect	1948	Feynman's path integral formulation
1895	Roentgen discovers X-rays	1952	Bohm's pilot-wave theory
1896	Becquerel discovers radioactivity	1957	Everett's many-worlds interpretation
1897	Thomson discovers the electron	1961	Wigner's friend thought experiment
1900	Planck quantizes energy to explain black-body radiation	1964	Bell's theorem
1905	Einstein proposes photons to explain photoelectric effect	1967	Kochen–Specker theorem on contextuality
1909	Geiger and Marsden observe alpha particle scattering	1981	Zurek introduces decoherence
1911	Rutherford proposes the nuclear atom	1982	Aspect's experiment on entanglement
1913	Bohr develops a quantized model of the atom	1994	Shor's algorithm
1919	Rutherford identifies the proton	1998	First quantum logic gate
1922	Stern and Gerlach observe the deflection of atoms in a spatially varying magnetic field	2019	Google claims quantum supremacy
1923	Compton effect demonstrates photon momentum		
1924	de Broglie proposes wave-particle duality for matter		
1925	Heisenberg introduces matrix mechanics		
1925	Pauli exclusion principle		
1925	Goudsmit and Uhlenbeck propose electron spin		
1926	Schrödinger formulates wave mechanics		
1926	Born introduces the probabilistic interpretation of the wavefunction		
1927	Heisenberg's uncertainty principle		
1927	Davisson–Germer experiment confirms electron diffraction		
1927	Solvay Conference on quantum theory		
1928	Dirac unifies quantum mechanics with relativity		

CHAPTER 1

AN EXECUTIVE SUMMARY OF WAVE MECHANICS

I insist upon the view that 'all is waves'.

ERWIN SCHRÖDINGER IN A LETTER TO JOHN LIGHTON SYNGE, 1959.

In this introductory chapter, we provide an overall picture of wave quantum mechanics, discuss its more important and surprising features, and introduce a series of concepts that will be important during the whole course. We start our discussion by introducing the Hamilton-Jacobi equation, which Schrödinger himself used as the starting point for deriving his seminal time-independent quantum wave equation (E.Schrödinger, An Undulatory Theory of the Mechanics of Atoms and Molecules, Physical Review (1926) Vol. 28, No. 6 pp. 1049-1070).

1.1 The Hamilton-Jacobi equation

Classical mechanics can be formulated through four fundamental approaches: the Newtonian, the Lagrangian, the Hamiltonian, and the Hamilton-Jacobi formalisms. The last one, in particular, offers a refined framework for analyzing Lagrangian systems, encompassing both geometrical optics and classical mechanics, and establishing a profound link between trajectories and wave-like behavior.

A transformation $(q_i, p_i) \to (Q_i, P_i)$ of the spatial coordinates q_i and canonical momenta p_i $(i = 1, 2, \ldots, n)$ of an n-dimensional system is said to be canonical if it preserves the form of the Hamilton equations, namely

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}\,, \qquad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}\,, \quad \longrightarrow \quad \frac{dQ_i}{dt} = \frac{\partial K}{\partial P_i}\,, \qquad \frac{dP_i}{dt} = -\frac{\partial K}{\partial Q_i}\,. \tag{1.1}$$

The new Hamiltonian K is related to the original Hamiltonian H as

$$K\left(Q_i, P_i, t\right) = H\left(q_i, p_i, t\right) + \frac{\partial F}{\partial t}, \tag{1.2}$$

with F the generating function of the canonical transformation, which can take the alternative forms $F\left(q_i, Q_i, t\right)$, $F\left(q_i, P_i, t\right)$, $F\left(p_i, Q_i, t\right)$ or $F\left(p_i, P_i, t\right)$. The so-called *Hamilton's principle function* $S \equiv F\left(q_i, P_i, t\right)$ is the specific generating function for which the transformed Hamiltonian K vanishes, leading to the equation,

$$H\left(q_i, p_i, t\right) + \frac{\partial S}{\partial t} = 0\,. \tag{1.3}$$

In this case, the transformation equations relating coordinates and momenta take the form

$$p_i = \frac{\partial S}{\partial q_i}\,, \qquad Q_i = \frac{\partial S}{\partial P_i}\,. \tag{1.4}$$

Substituting the first of these expressions into (1.3), we obtain the seminal *Hamilton-Jacobi equation*,

$$H\left(q_i, \frac{\partial S}{\partial q_i}, t\right) + \frac{\partial S}{\partial t} = 0\,. \tag{1.5}$$

The solution of this *partial* differential equation provides the function S, and with it the equations of motion of the system when Q_i and P_i are identified with initial values of the coordinates and momenta (the variables Q and P remain constant with time, cf. (1.1)).[1] In particular, the set of equations

$$Q_i = \frac{\partial S\left(q_i, P_i, t\right)}{\partial P_i} \tag{1.6}$$

can be inverted to express the coordinates q_i in terms of the transformed variables

$$q_i = q\left(Q_i, P_i, t\right)\,. \tag{1.7}$$

The momenta p_i are then obtained using

$$p_i = \left.\frac{\partial S\left(q_i, P_i, t\right)}{\partial q_i}\right|_{q_i = q(Q_i, P_i, t)}\,. \tag{1.8}$$

Since the variables P_i remain constant, the Hamilton's principal function S depends only on q_i and t, i.e.

$$dS = \sum_{i=1}^{n} \frac{\partial S}{\partial q_i} dq_i + \frac{\partial S}{\partial t} dt\,. \tag{1.9}$$

Making use of Eq. (1.3) and the first expression in (1.4), we have therefore

$$\frac{dS}{dt} = \sum_{i=1}^{n} p_i \frac{dq_i}{dt} - H = L\,, \tag{1.10}$$

with $L\left(q_i, dq_i/dt, t\right)$ the Lagrangian. Integrating this equation, identifies S as the action or indefinite time integral of the Lagrangian,

$$S = \int L\, dt + \text{constant}\,. \tag{1.11}$$

Note that while the Newtonian, Lagrangian, and Hamiltonian formalisms ultimately give rise to systems of coupled ordinary differential equations, often with a large number of variables, the Hamilton-Jacobi equation elegantly consolidates this complexity into a single partial differential equation for the action S, a quantity that plays only a minor explicit role in the other formalisms. Moreover, while the equations of motion in these traditional approaches depend on the gradient of the potential, the Hamilton-Jacobi equation uniquely incorporates $V(x)$ in its undifferentiated form.

[1] We are not suggesting that solving this differential equation is straightforward in general, but rather that we know the number of integration constants required in the final solution.

Example 1: The free particle in one-dimension

The Hamiltonian of a free particle in one-dimension, $H = \frac{p^2}{2m}$, is independent of the coordinate x, meaning it does not explicitly depend on the generalized coordinate $q = x$,

$$H = H(p) = H\left(\frac{\partial S}{\partial q}\right) = \frac{1}{2m}\left(\frac{\partial S}{\partial q}\right)^2 .$$

The corresponding Hamilton-Jacobi equation,

$$H\left(q, \frac{\partial S}{\partial q}; t\right) + \frac{\partial S}{\partial t} = 0 ,$$

takes then the form

$$\frac{\partial S}{\partial t} = -\frac{1}{2m}\left(\frac{\partial S}{\partial x}\right)^2 .$$

Since the right-hand side of this expression is independent of time, we can express S as

$$S = W(x) - Et . \tag{1.12}$$

with $W(x)$ the so-called *Hamilton's characteristic function*. We then have

$$-E = -\frac{1}{2m}\left(\frac{\partial S}{\partial x}\right)^2 \quad \longrightarrow \quad \frac{\partial S}{\partial x} = \sqrt{2mE} \equiv \alpha \quad \longrightarrow \quad W(x) = \alpha x + C ,$$

and consequently

$$S(x,t) = \alpha x + C - \frac{\alpha^2}{2m}t .$$

Taking now into account that $K = H + \frac{\partial S}{\partial t} = 0$, we get

$$\frac{p^2}{2m} - \frac{\alpha^2}{2m} = 0 \quad \longrightarrow \quad p = \alpha = \text{constant} ,$$

and therefore

$$S\left(x, p_0, t\right) = p_0 x - \frac{p_0^2}{2m}t .$$

where we have used the freedom to add an arbitrary constant to S to set $C = 0$. Since

$$P = \frac{\partial S}{\partial x} = p_0 , \qquad Q = \frac{\partial S}{\partial p_0} = x - \frac{p_0}{m}t ,$$

the solution of the original problem is

$$x = Q + \frac{p_0}{m}t , \qquad p = p_0 ,$$

with the new canonical variables (Q, P) corresponding to the initial position x_0 and the initial momentum p_0.

For reasons that will become evident soon, let us consider the Hamilton-Jacobi equation (1.5) for a non-relativistic particle of mass m moving along the x-axis under the influence of a real-valued potential $V(x) \in \mathbb{R}$,

$$\frac{1}{2m}\left(\frac{\partial S}{\partial x}\right)^2 + V + \frac{\partial S}{\partial t} = 0\,, \tag{1.13}$$

and introduce a quantity

$$\varphi(x,t) \equiv \exp\left(-\frac{S(x,t)}{S_0}\right)\,, \tag{1.14}$$

with S_0 a constant with dimensions of action. Inverting this expression, $S = -S_0 \ln\varphi$, and differentiating the result with respect to x and t, we obtain

$$\frac{\partial S}{\partial x} = -\frac{S_0}{\varphi}\frac{\partial \varphi}{\partial x}\,, \qquad \frac{\partial S}{\partial t} = -\frac{S_0}{\varphi}\frac{\partial \varphi}{\partial t}\,. \tag{1.15}$$

Differentiating now the first of these equations with respect to x,

$$\frac{\partial^2 \varphi}{\partial x^2} = -\frac{1}{S_0}\frac{\partial \varphi}{\partial x}\frac{\partial S}{\partial x} - \frac{1}{S_0}\varphi\frac{\partial^2 S}{\partial x^2}\,, \tag{1.16}$$

and combining it with the result of differentiating the first transformation in (1.4),

$$\frac{\partial^2 S}{\partial x^2} = \frac{\partial p_x}{\partial x} = m\frac{d}{dt}\left(\frac{\partial x}{\partial x}\right) = 0\,, \tag{1.17}$$

results in

$$\left(\frac{\partial S}{\partial x}\right)^2 = \frac{S_0^2}{\varphi}\frac{\partial^2 \varphi}{\partial x^2}\,. \tag{1.18}$$

This allows us to express Eq. (1.13) as a wave equation

$$S_0\frac{\partial \varphi}{\partial t}(x,t) = \left(\frac{S_0^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right)\varphi(x,t)\,. \tag{1.19}$$

Note that, although φ is dimensionless by definition in (1.14) , this expression remains valid when φ is scaled by any dimensional constant necessary for its proper physical interpretation in a specific application.

1.2 The Schrödinger equation

In quantum mechanics, the evolution of a non-relativistic particle of mass m moving along the x-axis[2] under the influence of a real-valued potential $V(x) \in \mathbb{R}$ is described by a partial differential equation, the so-called *Schrödinger equation*

$$i\hbar\frac{\partial \varphi}{\partial t}(x,t) = \left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right)\varphi(x,t)\,, \quad \longleftrightarrow \quad i\hbar\frac{\partial \varphi}{\partial t}(x,t) = H\varphi(x,t)\,. \tag{1.20}$$

[2]The restriction to one dimension is just done for the sake of clarity; generalization to three dimensions is straightforward.

Here $\hbar = 1.054 \times 10^{-34}\text{J} \cdot \text{s} = 6.5822 \times 10^{-16}\text{eV} \cdot \text{s}$ is the reduced Planck constant (pronounced "h-bar"), $\varphi(x,t)$ is the *wave function* of the system, and the differential operator in parentheses is the Hamiltonian of the theory.

The similarity between the quantum Schrödinger equation (1.20) and the classical Hamilton-Jacobi equation (1.19) is apparent. A simple inspection of Eq. (1.20) reveals several interesting facts:

- The Schrödinger equation is linear in the wave function, since the operator H does not depend itself on it and

$$L(a\,\varphi) = a\,L\varphi\,, \qquad L\,(\varphi_1 + \varphi_2) = L\varphi_1 + L\varphi_2\,,$$

with

$$L \equiv i\hbar\frac{\partial}{\partial t} - H\,, \tag{1.21}$$

and a a constant that, as we will discuss below, does need not be real. Therefore, any linear superposition of solutions with arbitrary coefficients is also a solution. This makes quantum mechanics much simpler than many classical theories. Consider for instance the corresponding 1-dimensional Newton's equation of motion for a non-relativistic particle of mass m moving in a time-independent potential $V(x)$,

$$m\frac{d^2x(t)}{dt^2} = -V'[x(t)]\,. \tag{1.22}$$

Since the function $V'[x(t)] \equiv \partial V(x)/\partial x|_{x=x(t)}$ is generically non-linear in its argument,

$$V'(ax(t)) \neq aV'(x(t))\,, \qquad V'(x_1(t) + x_2(t)) \neq V'(x_1(t)) + V(x_2(t))\,, \tag{1.23}$$

neither the scaling of a solution by a constant nor the superposition of solutions are guaranteed to be a solution.

- The Hamiltonian is a Hermitian operator, meaning that for any single-valued[3] functions f and g vanishing at infinity we have

$$\int dx\, f^*(Hg) = \int dx\,(Hf)^* g\,, \tag{1.24}$$

This is a trivial condition for the potential term V and also true for the operator $\partial^2/\partial x^2$ (or the corresponding Laplacian in 3 dimensions), as can be easily seen be integrating the identity

$$\left(\frac{\partial^2 f}{\partial x^2}\right)^* g - f^*\left(\frac{\partial^2 g}{\partial x^2}\right) = \frac{\partial}{\partial x}\cdot\left[\left(\frac{\partial f}{\partial x}\right)^* g - f^*\left(\frac{\partial g}{\partial x}\right)\right]. \tag{1.25}$$

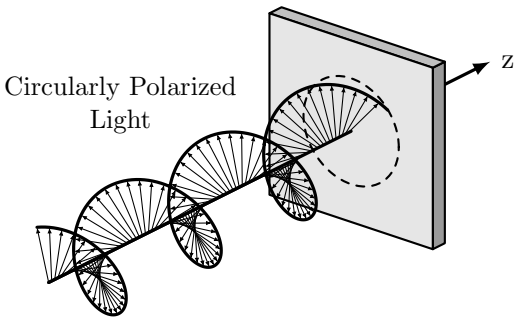
- The wave function is necessarily complex since otherwise the right-hand side of Eq. (1.20) would be real, while the left-hand side would be imaginary due to the i factor. Note that is a very unusual situation. Indeed, although we are certainly used to deal with complex numbers in physics, most of the time they are just convenient tools, not an

[3] This condition avoids multiple probability values, see below.

absolute necessity! Consider, for instance, the electric field associated to a circularly polarized wave moving along the z axis,

$$\vec{E}(z,t) = \mathrm{Re}\left(E_0(\vec{e}_x + i\vec{e}_y)e^{i(kz-\omega t)}\right) ,$$

with $\vec{e}_x$, $\vec{e}_y$ unitary vectors.



While this expression involves auxiliary complex numbers making it look nice and compact, the electric field is at the very end a real quantity,

$$\vec{E}(z,t) = E_0 \cos(kz-\omega t)\vec{e}_x - E_0 \sin(kz-\omega t)\vec{e}_y \,, \tag{1.26}$$

and so are the Maxwell equations! On the contrary, in quantum mechanics, both the Schrödinger equation and the dynamical variable $\varphi(x,t)$ are fundamentally complex, posing immediately a question on the physical interpretation of the wave function, since we just measure real quantities. This problem motivated Bohr to construct the most immediate real object one can derive from a complex number: its norm,

$$\varphi^*(x,t)\varphi(x,t) = |\varphi(x,t)|^2 \in \mathbb{R} \,. \tag{1.27}$$

This *positive-definite* quantity is then interpreted as a probability density for the particle to be near x at time t.[4] Imposing that the particle can be found somewhere at any time with unit probability leads to the wave function normalization condition

$$\int_{-\infty}^{\infty} dx\, |\varphi(x,t)|^2 = 1 \,. \tag{1.28}$$

For square integrable functions, this normalization can be easily achieved by a trivial rescaling of $\varphi(x)$, with the resulting wave function representing a physical state. On the other hand, if the integral of $|\varphi(x,t)|^2$ diverges, there is no possible rescaling leading to the fulfillment of the normalization condition (1.28), meaning that such a non-square integrable function does not represent a physical state. Note finally that the normalization condition allows us to quickly figure out the wave function units. In particular, in one dimension, since dx has the units of length and probabilities are unitless, we have $[\varphi] = 1/\sqrt{\text{length}}$.

[4] This probabilistic interpretation was opposed by many renowned physicists such as Einstein or Schrödinger itself and led to many debates along the years, most notably at the Solvay Conference in 1927. Although carefully explained in many books, we will not cover the interpretations of quantum mechanics in this course.

- The Schrödinger equation is a first-order differential equation in time. Therefore, in spite of the above probabilistic interpretation, the specification of the wave function all over space at an arbitrary initial time t_0 determines the wave function of the system at all times. The linearity of the equation guarantees also that the normalization condition (1.28) is necessarily conserved. Indeed, denoting the normalization as $N(t) \equiv \int_{-\infty}^{\infty} dx|\varphi(x,t)|^2$ and differentiating with respect to time, we get

$$\frac{dN}{dt} = \frac{d}{dt}\int_{-\infty}^{\infty} |\varphi(x,t)|^2 dx = \int_{-\infty}^{\infty} \frac{\partial}{\partial t}|\varphi(x,t)|^2\, dx \quad \rightarrow \quad \frac{d}{dt}(\varphi^*\varphi) = \varphi^*\frac{d\varphi}{dt} + \frac{d\varphi^*}{dt}\varphi = 0\,, \tag{1.29}$$

  where we have made use of the Leibniz's rule

$$\frac{d}{dt}\int_a^b f(x,t)\, dx = \int_a^b \frac{\partial f}{\partial t}(x,t)\, dx + f(b,t)\frac{db}{dt} - f(a,t)\frac{da}{dt}$$

  with constant integration limits for differentiating under the integral sign and taken into account the time-dependent Schrödinger equation (1.20) and its complex conjugate $-i\hbar\partial\,\varphi^*/\partial t = \varphi^* H$, where the Hermiticity of H is implicitly used. This result is important since, if the wave function at later times were not normalized, then the probability interpretation of quantum mechanics would not work.

- The Schrödinger equation for time-independent potentials is invariant under time-reversal, $T : \varphi(x,t) \mapsto \varphi^\star(x,-t)$, with the factor of i on the left-hand side playing once again a crucial role in this invariance. In this case, the equation admits separable solutions of the form

$$\varphi(x,t) = \sum_n a_n \exp\left(-\frac{iE_n t}{\hbar}\right)\varphi_n(x)\,, \tag{1.30}$$

  with the *stationary functions*[5] $\varphi_n(x)$ obeying the time-independent Schrödinger equation

$$-\frac{\hbar^2}{2m}\varphi_n''(x) + V(x)\varphi_n(x) = E_n\varphi_n(x)\,, \quad \longleftrightarrow \quad H\varphi_n(x) = E_n\varphi_n(x)\,, \tag{1.31}$$

  and E_n the energies of these solutions, which, as follows, from (1.24), are necessarily real,

$$E_n\int dx\,\varphi_m^*\varphi_n = \int dx\,\varphi_m^*(H\varphi_n) = \int dx\,(H\varphi_m)^*\varphi_n = E_m^*\int dx\,\varphi_m^*\varphi_n\,, \quad (n=m). \tag{1.32}$$

  As any other second-order linear differential equation, Eq. (1.31) displays two linearly independent solutions with symmetry and continuity properties depending on the specific potential $V(x)$. Regarding continuity, if the potential varies continuously with respect to x, the wave function $\varphi(x)$ will be continuous and at least two-times differentiable in that variable. Moreover, if $V(x)$ exhibits a finite discontinuity at a certain point, both $\varphi(x)$ and $\varphi'(x)$ will remain continuous at that specific point. However, if the potential features an infinite discontinuity at a particular location, $\varphi(x)$ will retain

[5] The states only change by a phase, hence the name. Note, also, that the associated probability $|\varphi(x,t)|^2 = |a_n|^2$ does not change with time.

its continuity at that point, but $\varphi'(x)$ will not. Regarding symmetry, potentials invariant under parity transformations or reflections about the origin, $V(-x) = V(x)$, will display wave functions that are either symmetric or antisymmetric under reflections.

- If we multiply (1.31) by $\varphi_n^*(x)$, integrate over x and perform a partial integration we obtain

$$E_n = \frac{\hbar^2}{2m}\int dx\ \left|\varphi_n'(x)\right|^2 + \int dx\, V(x)|\varphi_n(x)|^2 \geq \int dx\, V(x)|\varphi_n(x)|^2\,, \tag{1.33}$$

meaning that if the potential has a global minimum, one has

$$E_n \geq \int dx\, V(x)|\varphi_n(x)|^2 \geq V_{\min}\int dx\,|\varphi_n(x)|^2 = V_{\min} \quad \longrightarrow \quad E_n \geq V_{\min}\,. \tag{1.34}$$

1.3 Some exact solutions

For future reference, we summarize here some standard solutions of the Schrödinger equation (1.31) (or its 3D version), as usually done in *Quantum Physics I* courses. Alternative algebraic approaches to the harmonic oscillator and the hydrogen atom can be found in Appendices A and (B). The knowledge of these solutions will be taken for granted in this course.

1.3.1 The free particle

Consider the Hamiltonian of a free particle of mass m,

$$H = \frac{p^2}{2m}\,. \tag{1.35}$$

The solution of the corresponding corresponding time-independent Schrödinger equation (1.31),

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\varphi(x) = E\varphi(x)\,, \tag{1.36}$$

are the momentum eigenstates

$$u_p(x) = \frac{1}{\sqrt{2\pi\hbar}}e^{\frac{i}{\hbar}px} \tag{1.37}$$

with

$$E_p = \frac{p^2}{2m}\,. \tag{1.38}$$

Taking into account that these momentum eigenstates are also energy eigenstates, the formal solution of the time-dependent Schrödinger equation can be written as

$$\varphi(x,t) = \int_{-\infty}^{\infty}\frac{dp}{\sqrt{2\pi\hbar}}\varphi(p,0)e^{-\frac{ip^2}{2m\hbar}t+\frac{i}{\hbar}px}\,, \tag{1.39}$$

with $\varphi(p,0)$ the initial wave function in the momentum representation. For a normalized Gaussian wavepacket

$$\varphi(p,0) = \left[\frac{2\sigma^2}{\pi\hbar^2}\right]^{\frac{1}{4}} e^{-\frac{\sigma^2}{\hbar^2}(p-p_0)^2} \tag{1.40}$$

describing a superposition of momentum eigenstates with momenta centered around p_0 and probability amplitudes becoming small as $|p-p_0| \gg \hbar/\sigma$, we have, for instance,

$$|\varphi(x,t)|^2 = \frac{\sigma}{\sqrt{2\pi\hbar^4|b(t)|^4}} e^{-\frac{\sigma^2}{2\hbar^4|b(t)|^4}(x-p_0t/m)^2}, \quad \hbar^2 b^2(t) = \sigma^2 + \frac{i\hbar t}{2m}. \tag{1.41}$$

This describes a Gaussian wavepacket moving with velocity p_0/m while broadening in time as

$$\sigma^2(t) = \sigma^2 + \left(\frac{\hbar t}{2m\sigma}\right)^2. \tag{1.42}$$

1.3.2 The infinite potential well

Consider a quantum particle confined within a one-dimensional potential well

$$V(x) = \begin{cases} 0 & \text{if } 0 < x < L\,. \\ \infty & \text{otherwise}\,. \end{cases} \tag{1.43}$$

Requiring the finite-energy solutions of the corresponding time-independent Schrödinger equation (1.31) for $0 < x < L$,

$$\varphi(x) = A\cos(kx) + B\sin(kx)\,, \qquad k \equiv \sqrt{\frac{2mE}{\hbar^2}}\,, \tag{1.44}$$

to satisfy the boundary conditions $\varphi(0) = \varphi(L) = 0$ determines the integration constant $A = 0$ and leads to the quantization of the wavenumber $k = \pi n/L$ and the associated energy levels

$$E_n = \frac{\hbar^2\pi^2 n^2}{2mL^2}\,, \qquad n = 1, 2, 3, \ldots\,, \tag{1.45}$$

The second integration constant B is fixed, up to a physically irrelevant phase, by the normalization condition

$$\int_0^a dx\,|\varphi(x)|^2 = 1\,. \tag{1.46}$$

The resulting normalized energy eigenstates

$$\varphi_n(x) = \sqrt{\frac{2}{L}}\sin\left(\frac{\pi n x}{L}\right) \tag{1.47}$$

are either symmetric or antisymmetric under reflections about $x = L/2$. Note that, in general, the n-th excited state has n nodes, as expected from the node theorem (Sturm-Lioville theorem).

1.3.3 The harmonic oscillator

Consider a quantum particle of mass m confined in a one-dimensional harmonic oscillator potential of frequency ω,

$$V(x) = \frac{1}{2} m\omega^2 x^2 \,. \tag{1.48}$$

To solve the corresponding time-independent Schrödinger equation, we define the dimensionless variable

$$u \equiv \sqrt{\frac{m\omega}{\hbar}}\, x \,. \tag{1.49}$$

In terms of this quantity, Eq. (1.31) becomes

$$\frac{d^2\varphi}{du^2} + \left(\epsilon - u^2\right) \varphi = 0 \,, \tag{1.50}$$

with

$$\epsilon \equiv \frac{2E}{\hbar\omega} \tag{1.51}$$

a dimensionless energy. Introducing in (1.50) an ansatz of the form

$$\varphi(u) = v(u) e^{-\frac{u^2}{2}} \,, \tag{1.52}$$

and performing some trivial simplifications, we obtain a Hermite equation

$$v'' - 2uv' + 2nv = 0 \,, \qquad\qquad 2n \equiv \epsilon - 1 \,, \tag{1.53}$$

that can be tackled by complex variable methods. To this end, let us consider an integral form solution

$$v(u) = \int_\gamma e^{ut}\, Y(t) dt \,, \tag{1.54}$$

with γ an as-yet unspecified contour in the complex and $Y(t)$ a function to be determined in what follows. Taking the first and second derivatives of $v(u)$,

$$v' = \int_\gamma e^{ut} t Y(t) dt \,, \qquad v'' = \int_\gamma e^{ut} t^2 Y(t) dt \,, \tag{1.55}$$

we can write

$$uv' = \int_\gamma tY(t) \left(ue^{ut} dt\right) = \int_\gamma tY(t) de^{ut} = tY(t) e^{ut}\Big|_A^B - \int_\gamma e^{ut} \frac{d}{dt}(tY(t)) \,, \tag{1.56}$$

with A and B the end points of the contour γ. Imposing this contour to satisfy

$$tY(t) e^{ut}\Big|_A^B = 0 \,, \tag{1.57}$$

the previous expression reduces to

$$uv' = -\int_\gamma e^{ut} \frac{d}{dt}(tY(t)) \,. \tag{1.58}$$

Inserting this result and the equations in (1.55) into the Hermite equation (1.53), we arrive to a first order ordinary differential equation for $Y(t)$,

$$\int_\gamma e^{ut}\left(2\frac{d}{dt}(tY(t)) + \left(t^2+2n\right)Y(t)\right)dt = 0 \quad \longrightarrow \quad \frac{d}{dt}(tY) + \left(\frac{t^2}{2}+n\right)Y = 0\,, \tag{1.59}$$

that can be integrated by separation variables. Defining $Z(t) = tY(t)$, we get

$$\frac{dZ}{dt} + \left(\frac{t}{2}+\frac{n}{t}\right)Z = 0 \quad \longrightarrow \quad \frac{dZ}{Z} = -\left(\frac{t}{2}+\frac{n}{t}\right)dt \quad \longrightarrow \quad Z(t) = \frac{1}{t^n}e^{-\frac{t^2}{4}}\,, \tag{1.60}$$

and therefore

$$Y(t) = \frac{1}{t^{n+1}}e^{-\frac{t^2}{4}}\,, \quad v(u) = \int_\gamma \frac{1}{t^{n+1}}e^{ut-\frac{t^2}{4}}dt\,, \quad \varphi(u) = e^{-\frac{u^2}{2}}\int_\gamma \frac{1}{t^{n+1}}e^{ut-\frac{t^2}{4}}dt\,. \tag{1.61}$$

The resulting wave function $\varphi(u)$ turns out to be normalizable only if n is a non-negative integer. By Eqs. (1.51) and (1.53), this implies the quantization of the energy spectrum,

$$E = \frac{\hbar\omega}{2}\varepsilon = \hbar\omega\left(n+\frac{1}{2}\right)\,. \tag{1.62}$$

The function $v(u)$ in (1.61) can be recast in a more familiar form by completing the squares in the exponentia,

$$v(u) = \oint_\gamma \frac{1}{t^{n+1}}e^{ut-\frac{t^2}{4}}dt = \oint_\gamma \frac{1}{t^{n+1}}e^{ut-\frac{t^2}{4}-u^2+u^2}dt = e^{u^2}\oint_\gamma \frac{1}{t^{n+1}}e^{-\left(u-\frac{t}{2}\right)^2}dt\,, \tag{1.63}$$

and introducing a displaced integration variable $z = u - \frac{t}{2}$. Dropping an irrelevant constant factor, we get

$$v(u) = e^{u^2}\oint_{\gamma'} \frac{e^{-z^2}}{(z-u)^{n+1}}dz\,, \tag{1.64}$$

with γ' a contour encircling the point $z = u$. Taking now into account the Cauchy's formula for the derivatives of homolomorfic functions,

$$\frac{d^k f(u)}{du^k} = \frac{k!}{2\pi i}\oint \frac{f(z)}{(z-u)^{k+1}}dz\,, \tag{1.65}$$

we can rewrite $v(u)$ as a Hermite polynomial of order n, [6]

$$v(u) = e^{u^2}\frac{d^n}{du^n}e^{-u^2} \quad \longrightarrow \quad H_n(u) = (-1)^n e^{u^2}\frac{d^n}{du^n}e^{-u^2} = \sum_{m=0}^{\left\lfloor\frac{n}{2}\right\rfloor}(-1)^m\frac{2^{n-2m}n!}{m!(n-2m)!}u^{n-2m}\,. \tag{1.66}$$

Note that this is not a general formula for the polynomials themselves-just an algorithms for calculating them. The corresponding wave function $\varphi(u)$ takes the form

$$\varphi(u) = \frac{1}{\sqrt{N}}\varphi_n(u) = \frac{1}{\sqrt{2^n n!\sqrt{\pi}}}e^{-\frac{u^2}{2}}H_n(u)\,, \tag{1.67}$$

[6] Hermite's polynomials $H_n(u)$ are defined with an additional factor of $(-1)^n$ to keep positive the coefficient next to the highest power of the argument.

with N a normalization factor determined by imposing

$$N = \int_{-\infty}^{\infty} \varphi_n^2 \, du = 1 \,. \tag{1.68}$$

Using (1.66) and integrating by part n times,

$$N = \int_{-\infty}^{\infty} e^{-u^2} H_n^2(u) du = \int_{-\infty}^{\infty} (-1)^n H_n(u) \left[\frac{d^n}{du^n} e^{-u^2} \right] du N = \int_{-\infty}^{\infty} e^{-u^2} \frac{d^n H_n}{du^n} du \,, \tag{1.69}$$

we get

$$\frac{d^n H_n}{du^n} = 2^n n! \quad \text{and} \quad N = 2^n n! \int_{-\infty}^{\infty} e^{-u^2} du = 2^n n! \sqrt{\pi} \,. \tag{1.70}$$

Note, once again, that, the n-th excited state has n nodes.

1.3.4 The hydrogen atom

Consider a hydrogen-like ion involving a single electron and a nucleus of charge Ze. Neglecting the spin degrees of freedom and other relativistic effects, the Hamiltonian of the 3-dimensional system can be approximated by

$$H = \frac{\vec{p}_n^2}{2m_n} + \frac{\vec{p}_e^2}{2m_e} - \frac{Ze^2}{|\vec{r}_e - \vec{r}_n|} \,, \tag{1.71}$$

with $e^2 = q_e^2/(4\pi\epsilon_0)$. The corresponding time-independent Schrödinger equation in the position representation takes the form

$$\left[-\frac{\hbar^2}{2m_n} \nabla_n^2 - \frac{\hbar^2}{2m_e} \nabla_e^2 - \frac{Ze^2}{|\vec{r}_e - \vec{r}_n|} \right] \varphi(\vec{r}_e, \vec{r}_n) = E \varphi(\vec{r}_e, \vec{r}_n) \,. \tag{1.72}$$

Introducing the center-of-mass coordinates

$$\vec{R} = \frac{m_e \vec{r}_e + m_n \vec{r}_n}{m_e + m_n} \,, \qquad \vec{r} = \vec{r}_e - \vec{r}_n \,, \tag{1.73}$$

we can rewrite the kinetic terms in this expression as

$$-\frac{\hbar^2}{2m_n} \nabla_n^2 - \frac{\hbar^2}{2m_e} \nabla_e^2 = -\frac{\hbar^2}{2} \left[\frac{1}{m_e + m_n} \nabla_{\vec{R}}^2 + \frac{1}{m} \nabla_{\vec{r}}^2 \right] \,, \tag{1.74}$$

with

$$m = \frac{m_e m_n}{m_e + m_n} \tag{1.75}$$

the reduced mass of the system, which for $m_n \approx 1836 \, m_e$ coincides approximately with the mass of the electron, $m \approx 0.999 \, m_e$. The full Schrödinger equation (1.72) becomes therefore

$$\left[-\frac{\hbar^2}{2(m_e + m_n)} \nabla_{\vec{R}}^2 - \frac{\hbar^2}{2m} \nabla_{\vec{r}}^2 - \frac{Ze^2}{r} \right] \varphi(\vec{R}, \vec{r}) = E \varphi(\vec{R}, \vec{r}) \,, \tag{1.76}$$

with $r = |\vec{r}|$. To solve this equation, we separate variables as $\varphi(\vec{R}, \vec{r}) = \Phi(\vec{R}) \phi(\vec{r})$, obtaining a sum of two independent terms

$$\frac{1}{\Phi(\vec{R})} \left[-\frac{\hbar^2}{2(m_e + m_n)} \nabla_{\vec{R}}^2 \Phi(\vec{R}) \right] + \frac{1}{\phi(\vec{r})} \left[-\frac{\hbar^2}{2m} \nabla_{\vec{r}}^2 - \frac{Ze^2}{r} \right] \phi(\vec{r}) = E \,, \tag{1.77}$$

which must individually equal a constant. This splits effectively the Schrödinger equation into two separate parts, one associated with the center-of-mass motion

$$-\frac{\hbar^2}{2(m_e+m_n)}\nabla^2_{\vec{R}}\Phi(\vec{R}) = E_{\rm CM}\Phi(\vec{R})\,, \tag{1.78}$$

and one with the relative motion of the two particles

$$\left[-\frac{\hbar^2}{2m}\nabla^2_{\vec{r}} - \frac{Ze^2}{r}\right]\phi(\vec{r}) = E_r\phi(\vec{r})\,, \tag{1.79}$$

with $E = E_{\rm CM} + E_r$ the total energy of the system. The solution for the center-of-mass piece is completely straightforward,

$$\Phi(\vec{R}) = \frac{1}{(2\pi\hbar)^{3/2}}e^{i\vec{K}\cdot\vec{R}}\,, \qquad E_{\rm CM} = \frac{\hbar^2\vec{K}^2}{2(m_n+m_e)}\,. \tag{1.80}$$

The relative motion part resembles that of a particle of mass m in a external Coulomb potential. Writing the gradient in spherical coordinates in terms the angular momentum operator,

$$\nabla^2_{\vec{r}} = \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} - \frac{1}{r^2}\frac{L^2}{\hbar^2}\,, \tag{1.81}$$

and separating the angular and radial variables,

$$\phi(\vec{r}) = R_{n,l}(r)Y_{l,m}(\theta,\phi)\,, \tag{1.82}$$

we obtain

$$\left[-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} - \frac{1}{r^2}\frac{L^2}{\hbar^2}\right) - \frac{Ze^2}{r} - E_{n,l}\right]R_{n,l}(r)Y_{l,m}(\theta,\phi) = 0\,. \tag{1.83}$$

Taking now into account that the spherical harmonics $Y_{l,m}$ are eigenstates of the angular momentum operator, $L^2Y_{l,m}(\theta,\phi) = \hbar^2\ell(\ell+1)Y_{l,m}(\theta,\phi)$, and dividing the resulting equation by $Y_{l,m}(\theta,\phi)$, we arrive at a radial Schrödinger equation

$$\left(-\frac{\hbar^2}{2m}\left[\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr}\right] + \frac{\hbar^2\ell(\ell+1)}{2mr^2} - \frac{Ze^2}{r}\right)R_{n,l}(r) = E_{n,l}R_{n,l}(r)\,, \tag{1.84}$$

which we choose to rewrite in terms of the dimensionless variable

$$\rho = \frac{r}{a_Z}\,, \qquad a_Z = \frac{\hbar^2}{me^2Z}\,, \tag{1.85}$$

to obtain

$$-R''_{n,l}(\rho) - \frac{2}{\rho}R'_{n,l}(\rho) + \left(\frac{\ell(\ell+1)}{\rho^2} - \frac{2}{\rho}\right)R_{n,l}(\rho) = \epsilon_{n,l}R_{n,l}(\rho)\,, \tag{1.86}$$

with

$$\epsilon_{n,l} = \frac{2ma_Z^2}{\hbar^2}E_{n,l}\,. \tag{1.87}$$

Note that (1.86) is divergent for $\rho \to 0$. Substituting the ansatz

$$R_{n,l}(\rho) = \rho^\ell e^{-\lambda_{n,l}\rho}F_{n,l}(\rho) \tag{1.88}$$

into it yields an expression

$$-\rho F''_{n,l}(\rho) - 2(\ell + 1 - \lambda_{n,l}\rho)F'_{n,l}(\rho) + 2(\lambda_{n,l}(\ell+1) - 1)F_{n,l}(\rho) = 0\,, \tag{1.89}$$

which can be recast as a generalized Laguerre equation by further introducing the variable $y = 2\rho\lambda_{n,l}$ and defining $F_{n,l}(\rho) = f_{n,l}(y)$,

$$y f''_{n,l}(y) + f'_{n,l}(y)(\alpha + 1 - y) + k f_{n,l}(y) = 0\,, \tag{1.90}$$

with

$$\alpha = 2\ell + 1\,, \qquad k = \frac{1}{\lambda_{n,l}} - \ell - 1\,. \tag{1.91}$$

The solutions of this seminal equation are Laguerre polynomials $L_k^{(\alpha)}(y)$ with $k = 0, 1, 2, \ldots$ or equivalently

$$\lambda_{n,l} = \frac{1}{n}\,, \qquad n \geq 1\,, \quad \ell < n\,. \tag{1.92}$$

Collecting all terms, the relative-motion wave function becomes $\phi_{nlm}(r,\theta,\phi) = R_{n,l}(r)Y_{l,m}(\theta,\phi)$ with

$$R_{n,l}(r) = N_{n,l} r^{\ell} L_{n-\ell-1}^{(2\ell+1)}(2r/na_Z)\, e^{-(r/na_Z)}\,, \tag{1.93}$$

and $N_{n,l}$ a normalization constant fixed by the requirement

$$1 = \int_0^\infty dr\, r^2 \int_0^\pi d\theta \sin(\theta) \int_0^{2\pi} d\phi\, |R_{n,l}(r)Y_{\ell m}(\theta,\phi)|^2 = \int_0^\infty dr\, r^2\, |R_{n,l}(r)|^2\,. \tag{1.94}$$

The energy eigenvalues associated with $\phi_{n\ell m}(r,\theta,\phi)$ are

$$E_n = -\frac{Z^2}{n^2} Ry\,, \tag{1.95}$$

with

$$Ry = \frac{m_e e^4}{2\hbar^2} = 13.6\,\text{eV} \tag{1.96}$$

the Rydberg constant.

1.4 Exercises

1. **The Gaussian integral** ✐

   Solve the integral

   $$I = \int_{-\infty}^{\infty} dx\, e^{-\alpha x^2} = \sqrt{\frac{\pi}{\alpha}}\,.$$

   Hint: Consider the two-dimensional integral $I^2 = \int_{-\infty}^{\infty} dx\, dy\, e^{-\alpha\left(x^2+y^2\right)}$.

2. **Expectation values for Gaussian states** ✐

   Use the previous result to compute the expectation value of x^2, x^4 and x^{2n} on a Gaussian state $\varphi(x) = e^{-\alpha x^2}$ with $\alpha > 0$ and n a natural number, i.e.

   $$\langle x^2\rangle = \int_{-\infty}^{\infty} dx\, x^2 e^{-2\alpha x^2}\,, \qquad \langle x^4\rangle = \int_{-\infty}^{\infty} dx\, x^4 e^{-2\alpha x^2}\,, \qquad \langle x^{2n}\rangle = \int_{-\infty}^{\infty} x^{2n} \mathrm{e}^{-\alpha x^2}\,.$$

   Hint: Differentiate the Gaussian integral in the previous exercise with respect to α *to introduce the terms* x^2*,* x^4 *and* x^{2n} *terms.*

3. **The residues' theorem** ✐

   Use the residues' theorem to show that

   $$\int_{-\infty}^{\infty} \frac{e^{i\omega x}}{x^2+\tau^2} dx = \frac{\pi}{\tau} e^{-\omega\tau}\,,$$

   with ω and τ real positive constants.

4. **Playing with matrices** ✐

   Consider two matrices

   $$A = \begin{pmatrix} a & 0 & 0 \\ 0 & -a & 0 \\ 0 & 0 & -a \end{pmatrix}\,, \qquad B = \begin{pmatrix} b & 0 & 0 \\ 0 & 0 & -ib \\ 0 & ib & 0 \end{pmatrix}\,,$$

   with a and b real constants.

   a) Are they Hermitian? Do they commute?

   b) Determine the corresponding eigenvalues and eigenvectors.

   c) Find a set of orthonormal eigenvectors that are simultaneous eigenvectors of both A and B.

   d) Construct a rotation matrix R such that $R^\dagger A R$ and $R^\dagger B R$ are diagonal. Verify that $R^\dagger R = R R^\dagger = \mathbb{I}$.

5. **The harmonic oscillator in the Hamilton-Jacobi approach** ✐

   Consider the Hamiltonian of a one-dimensional harmonic oscillator with mass m and spring constant k,

   $$H = \frac{p^2}{2m} + \frac{1}{2} k x^2\,.$$

   Use the Hamilton-Jacobi approach to determine the classical motion of the system. Compare your result with the standard solution obtained from Hamilton's equations.

6. **From Schrödinger to Hamilton–Jacobi**

A quantum particle moves in one dimension under the influence of a potential $V(x)$. Suppose its wavefunction can be written in the form

$$\varphi(x,t) = \exp\left(\frac{i}{\hbar}S(x,t)\right),$$

with $S(x,t)$ a real-valued function.

a) Show that substituting this ansatz into the time-dependent Schrödinger equation leads to an equation for $S(x,t)$ that reduces to the classical Hamilton–Jacobi equation in the limit $\hbar \to 0$. Not surprisingly classical mechanics is contained in Schrödinger wave mechanics!

b) Apply this result to the special case of a free particle, with $V(x) = 0$, and demonstrate that it gives the exact plane-wave solution. Why does the semiclassical approximation become exact in this situation?

7. **A particle in a one-dimensional potential well**

Determine the energies and wave functions for a particle confined to an interval $x \in [-L/2, +L/2]$, with $L > 0$ and boundary conditions $\varphi(L/2) = \varphi(-L/2) = 0$.

8. **Bound and scattering states in the exact solvable Pöschl–Teller potential.**

Consider a particle of mass m moving in the one-dimensional potential

$$V(x) = -\frac{\hbar^2}{m}\frac{1}{\cosh^2 x}.$$

a) Show that $\varphi(x) = (\tanh x + c)\exp(ikx)$ is a solution of the Schrödinger equation for a particular value of the constant c. Determine this value and find the corresponding energy. By analyzing the asymptotic behavior of $\varphi(x)$, compute the reflection and transmission coefficients.

b) Show that $\phi(x) = 1/\cosh x$ also satisfies the Schrödinger equation. Demonstrate that this solution corresponds to a bound state and calculate its energy. Provide an argument supporting the idea that this is the ground state of the system.

9. **The Floquet's and Bloch's theorems**

A potential $V(x)$ that satisfies the condition $V(x + na) = V(x)$ for $n = 0, \pm 1, \pm 2, \ldots$ is called a periodic potential with period a. In this case, the Schrödinger equation is invariant under transformations of the form $x \to x + na$, i.e., under translations by integer multiples of a.

Let $u_1(x)$ and $u_2(x)$ be two linearly independent solutions of the Schrödinger equation. Due to the aforementiones invariance, the functions $u_1(x + a)$ and $u_2(x + a)$ are also solutions. Therefore, there must exist constants $c_{11}, c_{12}, c_{21}, c_{22}$ such that

$$u_1(x + a) = c_{11}u_1(x) + c_{12}u_2(x),$$
$$u_2(x + a) = c_{21}u_1(x) + c_{22}u_2(x).$$

Prove the following seminal theorems:

a) *Floquet's Theorem:* Among all solutions of the Schrödinger equation, there exist two, $\varphi_1(x)$ and $\varphi_2(x)$, that satisfy $\varphi(x+a) = \lambda\varphi(x)$, where λ is a constant.

b) *Bloch's Theorem:* These solutions can be written in the form $\varphi(x) = e^{ikx}u_k(x)$, with $u_k(x)$ a periodic function of x with period a, i.e. $u_k(x+a) = u_k(x)$.

10. **The Kronig-Penney model**

The Kronig-Penney model describes a periodic potential formed by an infinite sequence of rectangular barriers. Each barrier has height V_0, width b, and is separated from the next by a distance $a - b$, such that a is the lattice period.

a) Determine the Bloch eigenfunctions of the Hamiltonian associated with this potential.

b) Show that for energies $E < V_0$, the energy spectrum consists of continuous bands of eigenvalues.

11. **What potential shapes this wave?**

Given the wave function

$$\varphi(x) = \left(\frac{x}{x_0}\right)^n e^{-x/x_0}\,,$$

where n and x_0 are constants, use the time-independent Schrödinger equation to determine the potential $V(x)$ for which $\varphi(x)$ is an eigenfunction and the corresponding energy eigenvalue E. Assume that $V(x) \to 0$ as $x \to \infty$.

12. **A particle in an impenetrable spherical shell**

Determine the discrete energy eigenvalues for a particle confined by an impenetrable spherical shell of radius R, inside which the potential is zero.

13. **Bound states in a finite spherical well**

In nuclear physics, a simple and effective model treats the nucleus as a finite spherical potential well: a particle inside the nucleus experiences a constant negative potential, while outside the potential vanishes. This is described by

$$V(r) = \begin{cases} -V_0, & \text{for } r < a\,, \\ 0, & \text{for } r > a\,, \end{cases}$$

with $V_0 > 0$ representing the depth of the potential well and a the nuclear radius. Determine the discrete energy eigenvalues associated with the bound states of a particle confined in this potential. Assuming the nucleus has a radius $a = 5 \times 10^{-13}\,\text{cm}$ and a well depth $V_0 = 10\,\text{MeV}$, calculate the minimum mass that a particle must have in order to be bound within the nucleus.

14. **Units in atomic physics**

Although we usually measure quantities in kilograms, meters and seconds or a combination of them, these conventional units are not the most suitable ones for atomic physics. The essential constants that come into play in this context are the reduced Planck constant $\hbar$, the speed of light c, the electron mass m_e and the squared electron electric charge $e^2 = q_e^2/(4\pi\epsilon_0)$. Combine $\hbar$, m_e and c to construct units of length, time and energy, stating the corresponding numerical values in the international system of units. Repeat the exercise for combinations of $\hbar$, m_e and e^2. Do the results ring the bell?

CHAPTER 2

SPIN AND THE STERN-GERLACH EXPERIMENT

No experiment is so dumb that it should not be tried.

WALTHER GERLACH'S REPLY TO BORN'S DISMISSIVE COMMENTS, QUOTING EDGAR MEYER

If you will simply admit that maybe [nature] does behave like this, you will find her a delightful, entrancing thing. Do not keep saying to yourself, if you can possible avoid it, 'But how can it be like that?' because you will get 'down the drain', into a blind alley from which nobody has escaped.

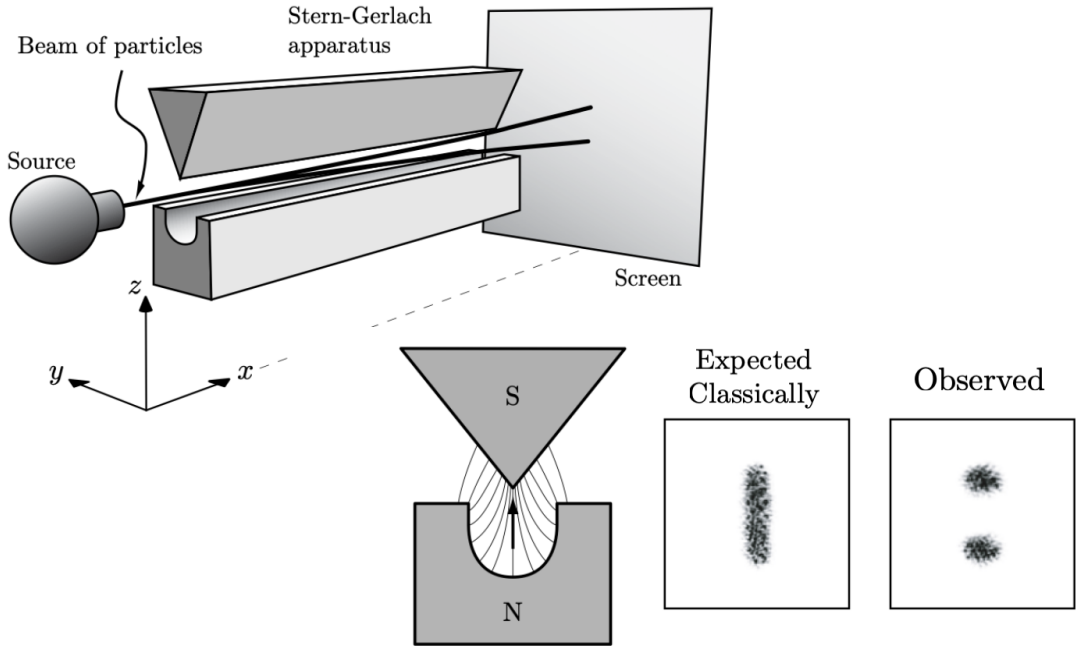
RICHARD FEYNMAN

In our development of quantum mechanics till now the behaviour of a particle was governed by a wave function depending on the position of the particle in a given coordinate system. There are, however, experimental observations that cannot be explained by a wave function depending only on coordinates. In this Chapter, we introduce the so-called Stern-Gerlach experiment and introduce the concept of spin, a quantum property with no classical analog.

2.1 The Stern-Gerlach experiment

The seminal Stern-Gerlach experiment was conducted in 1922 at the University of Frankfurt by Otto Stern and Walther Gerlach, both students of Ehrenfest. In its original version,[1] a collection of silver atoms is produced by evaporating silver into a high-temperature oven, from which they escape through a tiny hole. The produced atomic particles are then further collimated through a series of slits, producing a beam with a well-defined momentum p_x in the x-direction, which is subsequently passed through an inhomogeneous (non-uniform) magnetic field pointing mainly in the z direction and detected on a photographic plate. The results of the experiment were surprising to say the least: the silver atoms arrived on the screen at two spots, symmetrically located around the point of arrival in the absence of a magnetic field!

[1] The Stern-Gerlach experiment has been repeated with sodium, potassium, copper, gold, thalium and hydrogen atoms, with the same results.



To address this issue (and many other challenges in atomic physics that we will discover during this course), Wolfgang Pauli postulated the existence of a genuinely quantum mechanical degree of freedom that he cautiously called "a classical non-describable two valuedness" (*"eine klassisch nicht beschreibbare Art von Zweideutigkeit"*), an attribute that nowadays receives the much more favorable name of *spin*. Soon after, Ralph Kronig, Samuel Goudsmit and George Uhlenbeck approached this concept from a somehow less radical perspective, suggesting that this internal characteristic could be linked to an intrinsic rotation of particles. Let us adopt this naive point of view for the time being and analyze in detail the Stern-Gerlach experiment from a purely classical perspective.

2.2 Magnetic moment in classical physics

In classical physics, an atom is essentially a collection of charged particles moving in the field of a central force. Similarly to the Earth's motion around the Sun and its rotation on its own axis, we can imagine any of these charged particles having both orbital angular momentum $\vec{L}$ and intrinsic rotational angular momentum or spin $\vec{S}$.

According to the usual laws of electromagnetism, the current loop generated by a charge q in circular motion is associated with a magnetic moment $\vec{\mu}$ proportional to the area A of the loop. The latter quantity can be easily obtained by integrating the oriented area element swept out by the radius vector $\vec{r}$ over a period T. In particular, taking into account Kepler's area law, we have

$$\frac{d\vec{A}}{dt} = \frac{\vec{r}\times\vec{p}}{2m} = \frac{\vec{L}}{2m} \quad \longrightarrow \quad \vec{A} = \frac{T}{2m}\vec{L}\,, \tag{2.1}$$

with m the mass of the particle, $\vec{p}$ its linear momentum and $\vec{L}$ the associated angular momentum, a fixed vector perpendicular to the orbital plane. Multiplying the result by the current $I = q/T$ induced by the charge q when passing a given point $1/T$ times per period, we get the associated magnetic moment expression[2]

$$\vec{\mu}_L \equiv I\vec{A} = \frac{q}{2m}\vec{L}\,. \tag{2.2}$$

As should be clear from this expression, the magnetic moment $\vec{\mu}_L$ is a fixed vector perpendicular to the orbital plane with an orientation depending on the sign of q. In particular, $\vec{\mu}_L$ and $\vec{L}$ are counter-aligned for negatively charged particles. Note also that the proportionality constant is universal; in particular, it does not depend on the radius of the loop, nor on the velocity of the charge particle. The magnetic field moment has units of $A \cdot m^2$ in the International System of Units, with A standing for Amperes. It is, however, more convenient to express it in units of energy per magnetic flux density, $A \cdot m^2 = \text{Joules/Tesla}$.

A similar results holds for an intrinsic angular momentum $\vec{S}$,

$$\vec{\mu}_S = g\left(\frac{q\hbar}{2m}\right)\frac{\vec{S}}{\hbar}\,, \tag{2.3}$$

with g the so-called gyroscopic ratio, a factor effectively encoding our ignorance of the charge distribution. Experimentally, for electrons ($q = -e$, $g = 2$), protons ($q = e$, $g = 5.586$) and neutrons ($q = 0$, $g_n = -3.826$) we have

$$\vec{\mu}_{S,e} = -2\mu_B\frac{\vec{S}}{\hbar}\,, \qquad \vec{\mu}_{S,p} = 2.79\mu_N\frac{\vec{S}}{\hbar}\,, \qquad \vec{\mu}_{S,n} = -1.91\mu_N\frac{\vec{S}}{\hbar}\,, \tag{2.4}$$

with

$$\mu_B = \frac{e\hbar}{2m_e} = 9.27\times10^{-24}\frac{\text{Joules}}{\text{Tesla}}\,, \qquad \mu_N = \frac{e\hbar}{2m_p} = 5.05\times10^{-27}\frac{\text{Joules}}{\text{Tesla}}\,. \tag{2.5}$$

(Born magneton) (Nuclear magneton)

The value $g = 2$ for the electron will be derived in Chapter 2 using the so-called Pauli Hamiltonian. In reality, due to quantum electrodynamics' effects, this quantity turns out to be slightly larger from two, namely 2.00231930436. The deviation from 2 is known as the anomalous magnetic dipole moment. Note also that the existence of a non-vanishing neutron's magnetic moment reflects the fact that nucleons are not elementary particles. Indeed, protons and neutrons are composed of quarks, and the actual magnetic moments of the quarks can be used to compute the magnetic moments of the nucleons.

Let us estimate which of the above magnetic moments is the dominant one in silver atoms. As you probably recall from your chemistry lessons, the most common silver atoms have 47 electrons, 47 protons, and 60 or 62 neutrons. Since the magnetic moment of a given particle is inversely proportional to its mass, the contribution of heavy protons and neutrons is effectively suppressed by a numerical factor $m_e/m_p \sim 10^{-3}$, with m_e and m_p the electron and proton masses.[3] This leaves us with the potential contribution of 47 electrons. Note, however,

[2] Be careful when comparing books. This formula is correct in the international system of units. In Gaussian units the magnetic moment is $\vec{\mu}_L = IA/c$.

[3] In spite of this suppression, nuclear magnetism is of great practical importance as it lies at the basis of nuclear magnetic resonance. We will return to these issues in the following chapters.

that, as follows from the electronic configuration of silver, $1s^2 2s^2 2p^6 3s^2 3p^6 4s^2 3d^{10} 4p^6 4d^{10} 5s$, most of these inner electrons are paired inside closed shells, forming a spherically symmetric cloud with no orbital or intrinsic angular momentum. This leaves us a single $5s$ electron as main contributor to the silver magnetic moment, and in particular to the intrinsic angular momentum (2.3), since the orbital angular momentum of an s state is zero. The energy of an interaction of this magnetic moment with the magnetic field $\vec{B}$ is given by

$$U = -\vec{\mu} \cdot \vec{B}\,, \tag{2.6}$$

which results on a *classical* force

$$\vec{F} = -\vec{\nabla} U = \vec{\nabla}(\mu_x B_x + \mu_y B_y + \mu_z B_z) \simeq \frac{\partial}{\partial z}(\vec{\mu} \cdot \vec{B}) = -\frac{eg}{2m_e} S_z \frac{\partial B_z}{\partial z}\,, \tag{2.7}$$

where we have taken into account that $q = -e$ for the electron un $5s$. Note that the magnetic field $\vec{B}$ cannot be strictly parallel to the z axis, since configurations with $\vec{B} = (0, 0, B)$ and $\partial \vec{B}/\partial z \neq 0$ are incompatible with the Maxwell equation $\vec{\nabla} \cdot \vec{B} = 0$. The expression (2.7) should be therefore understood as an effective description of the real setting valid only when the time scale associated with the (precession) effects induced in the S_x component is fast as compared to the device crossing time, so this spin component averages effectively to zero. Consequently, the deviation of the beam of silver particles in the Stern-Gerlach experiment provides a measurement of the intrinsic magnetic moment of the electron along the z-axis, which is the orientation of the magnetic field gradient. For a thermal environment such as the oven, the distribution of the spin directions is expected to be completely random, covering effectively all possible values between $S_z = -|\vec{S}|$ and $S_z = +|\vec{S}|$. However, the Stern-Gerlach experimental finds only two potential values for the electron spin projection. The magnitude of the deflections is compatible with the spin projections values

$$S_z = \pm \frac{\hbar}{2} \tag{2.8}$$

with

$$\hbar \equiv \frac{h}{2\pi} = 6.5822 \times 10^{-16} \text{eV} \cdot \text{s}. \tag{2.9}$$

the reduced Planck constant, a fundamental unit of angular momentum. The factor $1/2$ in this expression leads us to refer to this as a *spin* $1/2$ *system*, i.e. a system of two states, one with *spin up* $(+z)$ and another with *spin down* $(-z)$. When orienting the magnetic field along any other direction, e.g. the x-axis, a similar result was observed,[4] $S_x = \pm\hbar/2$. Note that this result is at odds not only with classical physics but also with Schrödinger wave mechanics. In fact, if the result of the experiment would be associated with a regular orbital angular momentum l, we should expect the beam to divide into a *odd* number of components, namely $(2l + 1)$.

[4] While the $\pm\hbar/2$ is universal, the relative intensity of the images in the photographic plate depends in practice on the direction of the magnet, unless the incident beam is completely unpolarized.

The day quantum mechanics got smoky

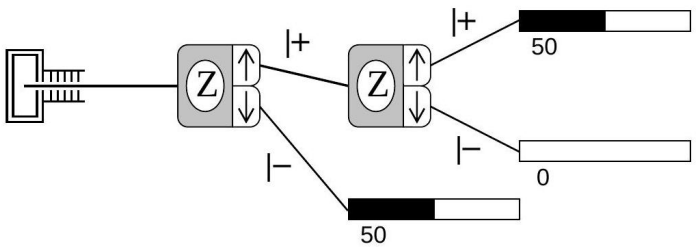
It's often said that science rewards careful planning, but sometimes a little bad habit can change everything. Otto Stern's fondness for cheap, foul-smelling cigars is a perfect example.

While working on the Stern–Gerlach experiment, Stern kept puffing away in the lab, filling the air with thick, sulfurous smoke. At first, the results of the experiment were barely visible. The silver atoms had landed on the detection plate, but the marks were so faint they were almost impossible to see. Then, almost by accident, Stern's cigar smoke came to the rescue. The sulfur in the smoke reacted with the tiny silver particles, darkening them and making the pattern stand out much more clearly. See here for further details.

2.3 Sequential Stern-Gerlach experiments

It is interesting to see what happens when performing several sequential Stern-Gerlach experiments. To simplify the figures, let us denote a Stern-Gerlach apparatus measuring the component of the spin $\vec{S}$ in a given direction x, y or z as a box with a X, Y or Z label.

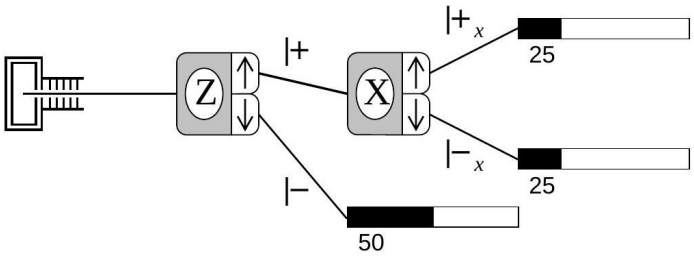
Experiment 1



This experiment involves two consecutive Stern-Gerlach devices both aligned along the z-axis. The first one acts essentially as a *polarizer or a state preparation device*, filtering out the $S_z = -\hbar/2$ component and letting only the $S_z = \hbar/2$ one proceed towards the next Z machine. By preparing the state in this manner, the details of the source can be ignored and we can focus on what happens in the second device, usually called the *analyzer*. Since all the ongoing particles have a definite spin projection $S_z = \hbar/2$, the second Z machine yields only the upper output, i.e. nothing comes out from the bottom port.

The quantum mechanical insight derived from this experience is that states with $S_z = \hbar/2$ do not have components or amplitudes along $S_z = -\hbar/2$. These states are referred to as orthogonal states.

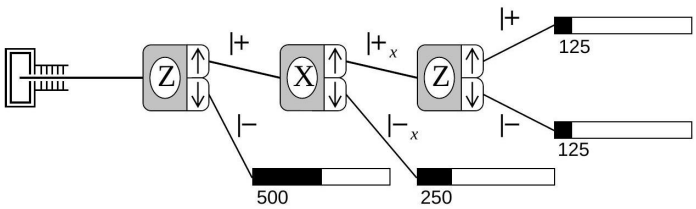
Experiment 2



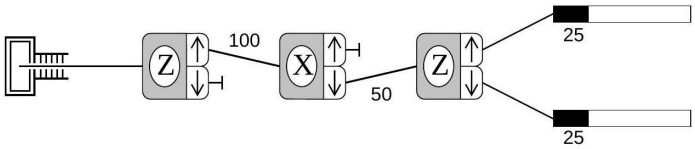
This experiment measures the spin projection along the z-axis and subsequently along the x-axis. As before, the relevant output of the first Z polarizer is just a $S_z = \hbar/2$ beam of particles directed towards the next X machine, which, as indicated in the figure, provides both possible outputs $S_x = \hbar/2$ and $S_x = -\hbar/2$, in clear analogy with a single Z-like device. Note that this is at odds with classical mechanics, where an object with angular momentum along the z axis has no component of angular momentum along the x axis. The results of this experiment remain unchanged if we rather use the lower Z polarizer port, i.e. atoms entering the X machine in a state $S_z = -\hbar/2$, result also in final states with $S_x = \hbar/2$ and $S_x = -\hbar/2$.

The quantum mechanical insight derived from this experience is that a state with a definite value of S_z has components along the states $S_x = \hbar/2$ and $S_x = -\hbar/2$. It also highlights the probabilistic nature of quantum mechanics. As can be experimentally demonstrated by recording the counts out of each port, the arrival sequence of atoms at any of the counters in the figure cannot be predicted. In particular, we can only say that there is a 50 % probability for an atom to exit the upper or lower port of *a particular* device. Note, however, that this is *not* the probability to pass through a whole system of Stern-Gerlach filters!

Experiment 3

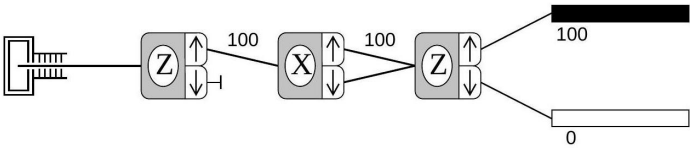


This experiment is an extension of the previous one where the spin projection is measured three times in a row. As before, the $S_z = \hbar/2$ beam exiting the first Z device goes into an X machine, which acts now also as a polarizer. Only the $S_x = \hbar/2$ component continues forward to an additional Z machine. Interestingly, while one could classically expect the final result to retain some memory of the first Z filter, half of the particles making it out to the third machine turn out to have $S_z = \hbar/2$ and the other half $S_z = -\hbar/2$. This means that, for all practical purposes, the last two analyzers in this experiment behave like the two analyzers in Experiment 2, but with $X \leftrightarrow Z$. The initial measurement resulting in $S_z = \hbar/2$ is effectively erased! It should be clear from the discussion of Experiment 2 that this conclusion remains unchanged if we rather use the lower X polarizer port, namely



The quantum mechanical insight to be derived from this experience is that there is a fundamental incompatibility when trying to know simultaneously the spin projection along two different directions; e.g. S_x and S_z are *incompatible observables*. Of course, this does not imply necessarily that all pairs of observables in quantum mechanics!

Experiment 4



This experiment is a slight variation of the previous one, where rather than disregarding or blocking one of the beams leaving the intermediate X device, we recombine them as an input of the following Z analyzer. [5] As displayed in the figure, all the atoms in this scenario exit the upper port of the last analyzer as if they remembered to have been initially measured to have $S_z = \hbar/2$! The combination of the two beams avoids the disturbance observed in Experiment 3, giving back the result of Experiment 1, as if the X device was not there! Note also that, in spite of allowing more ways or paths as compared to Experiment 3, the number of counts in the lower detector is zero.

The quantum mechanical insight derived from this experience is that the combination of two effects can lead to cancellation rather than enhancement. Does this ring the bell? Indeed, all the above is just the spin 1/2 version of the Young's double slit interference experiment you probably encountered in your *Quantum Physics I* course.

2.4 Not a spinning ball!

The things related to spin can be even more surprising. As Goudsmit and Uhlenbeck already recognised in their original paper, the interpretation of spin as an intrinsic rotation of particles we have adopted so far is doomed from the very beginning as it leads to serious problems with Special Relativity. [6] In particular, different experiments involving the scattering of electrons seemed to indicate that these should be extremely small ($r_e \lesssim 10^{-17}$ m across), behaving

[5] As discussed in the wonderful Feynman's Lectures on Physics, this is more easily said than done. In particular, the recombination must be done properly in order to avoid disturbing the beams.

[6] According to Goudsmit, this problem was pointed out by renowned physicist Hendrik Lorentz to at the time quite young Uhlenbeck, who overwhelmed implored Ehrenfest, his research supervisor at the University of Leiden, to withdraw their already submitted paper. Ehrenfest's reply was not precisely comforting: "It is too late. It will be published in two weeks. Both of you are young and can afford to do something stupid". The printed version does contain, however, a footnote acknowledging Lorentz's caveat.

essentially as a point-like particle and having to whirl at superluminal speed in order to match the intrinsic angular momentum $\hbar/2$ observed in the Stern-Gerlach experiments, namely

$$S = I\omega \sim mr^2\left(\frac{v}{r}\right) \sim mrv \sim \frac{\hbar}{2} \quad \longrightarrow \quad \frac{v}{c} \sim \frac{\hbar}{mrc} \sim \frac{10^{-34}\,\mathrm{J\cdot s}}{(10^{-30}\,\mathrm{kg})(10^{-17}\ \mathrm{m})(10^{8}\,\mathrm{m/s})} \sim 10^{4}\,.$$

with I the moment of inertia and ω the angular velocity. As Pauli strongly defended along his life (see e.g. the letter in Fig. 2.1), this problem gets resolved if one regards spin as an intrinsically internal property just like mass or charge, but without a classical analog.[7] In particular, in contrast to orbital angular momentum, *spin cannot be described by a differential operator*. This hypothesis was formally demonstrated by Paul Dirac in 1926, when he introduced his renowned relativistic equation, from which both spin and magnetic moment are automatically obtained without any further postulates. In some sense, relativity forces spin upon us.

[7] The term *spin* is indeed a misnomer that, unfortunately, has become entrenched within the terminology.

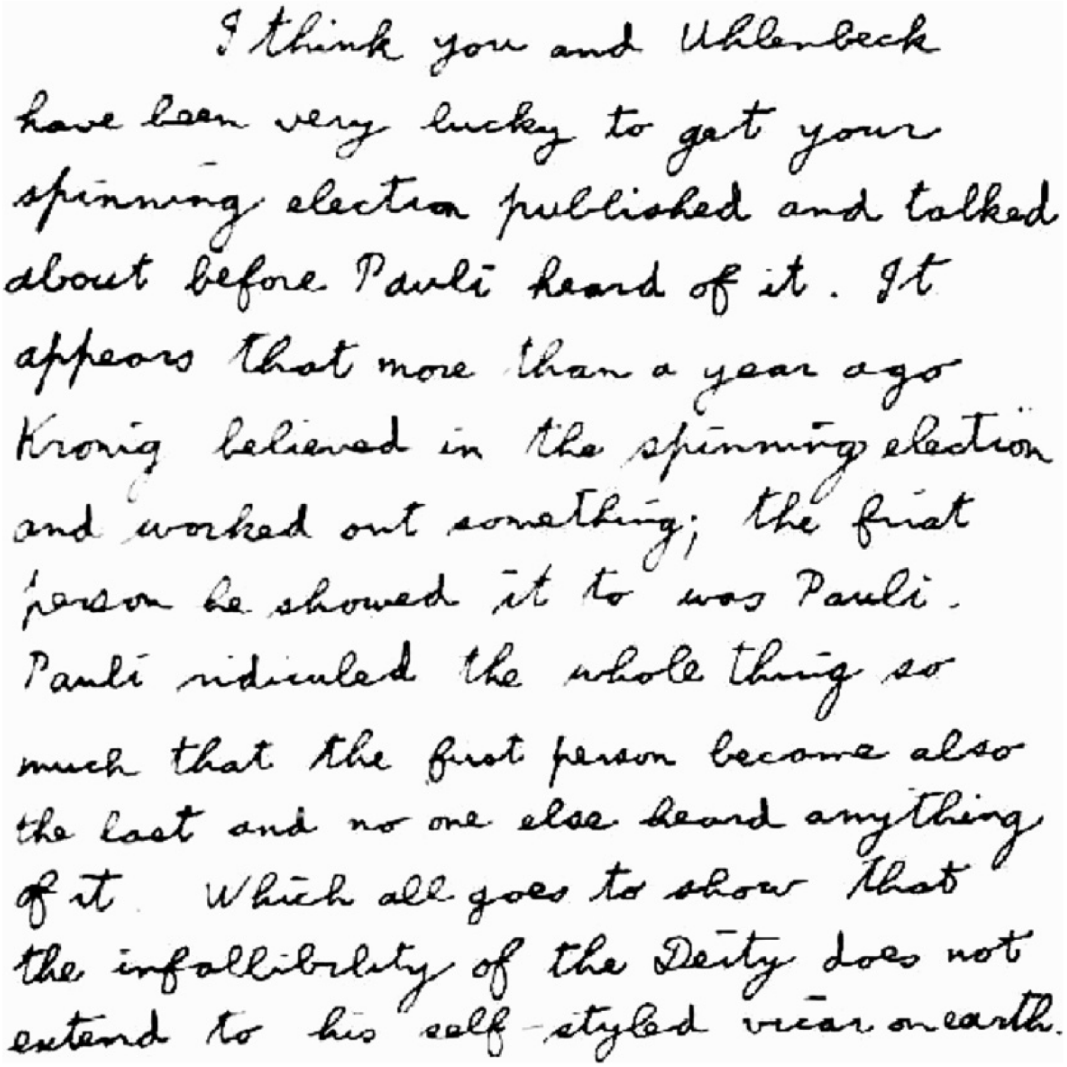

I think you and Uhlenbeck have been very lucky to get your spinning electron published and talked about before Pauli heard of it. It appears that more than a year ago Kronig believed in the spinning electron and worked out something; the first person he showed it to was Pauli. Pauli ridiculed the whole thing so much that the first person become also the last and no one else heard anything of it. Which all goes to show that the infallibility of the Deity does not extend to his self-styled vicar on earth.

Figure 2.1: Letter by L.H. Thomas to S. Goudsmit on March 25th 1926. Taken from The discovery of the electron spin by Goudsmit, 1971.

2.5 Exercises

1. **Maxwell's Equations and the Stern–Gerlach experiment**

   In a Stern–Gerlach-type experiment, a beam of silver atoms travels along the x-axis and passes through a region of non-uniform magnetic field. The field is predominantly oriented along the z-axis. Near the center of the magnet, the system is symmetric with respect to the $y-z$ plane and the field satisfies Maxwell's equations in vacuum. Let $\vec{B}(x, y, z)$ be the magnetic field near the point $(x, y, z) = (0, 0, 0)$.

   a) Use the symmetry of the field with respect to the $y-z$ plane to determine how $B_z(x, y, z)$ must behave under $x \to -x$. Deduce that $\partial B_z/\partial x|_{x=0} = 0$.

   b) Assuming the field is approximately invariant along the y-axis near the center of the magnet (neglect edge effects), argue that $\partial B_z/\partial y = 0$.

   c) Assume the field to have the form $\vec{B}(x, z) = B_0\,\hat{e}_z + \alpha z\,\hat{e}_z + \beta x\,\hat{e}_x$ near the center. Use the Maxwell's equation $\nabla \cdot \vec{B} = 0$ to find the relation between α and β.

2. **Modifying the Stern–Gerlach experiment**

   In the original Stern–Gerlach experiment, a beam of neutral silver atoms is passed through an inhomogeneous magnetic field directed along the z-axis, resulting in a splitting of the beam into two distinct spots on a detector screen.

   a) What would be the behavior of a beam of isolated electrons under the same magnetic field gradient? Would the beam still split into two components? What other effects, if any, would need to be considered?

   b) Would the beam still split if the magnetic field was constant in space? If not, describe the expected behavior of the particle beam.

   c) Suppose the silver atoms are ionized, losing their single $5s$ electron and becoming Ag^+ ions. Would a beam of these ions still split when passing through the original Stern–Gerlach setup?

3. **Deflection in the Stern–Gerlach experiment**

   In a Stern–Gerlach experiment, a beam of hydrogen atoms is emitted from an oven maintained at a temperature of $T = 500\,\mathrm{K}$. The atoms subsequently traverse a region characterized by a magnetic field gradient, which exerts a force on the magnetic moments associated with the electrons. The magnetic field gradient is $dB/dz = 0.2\,\mathrm{T/mm}$ and the length of the magnetic field region is $L = 20\,\mathrm{cm}$.

   a) Assuming that the atoms follow classical trajectories and that their electron spins are initially aligned along the $+z$-axis, determine the acceleration experienced by the atoms due to the magnetic force.

   b) Calculate the displacement of the beam in the z-direction upon exiting the magnetic field region.

4. **Magnetic fields and the 21 cm line**

   The 21 cm line observed in radio astronomy is extensively used to map the structure of our galaxy. This spectral line arises from the emission of a photon during a spin-flip transition

of the electron in a hydrogen atom. Specifically, the electron transitions from a spin state aligned with the proton's spin to an anti-aligned configuration. Determine the effective magnetic field experienced by the electron inside the hydrogen atom. The energy of the emitted photon is $E = \frac{hc}{\lambda}$, with h the Planck's constant, c the speed of light and λ the wavelength of the photon.

CHAPTER 3

PRINCIPLES OF QUANTUM MECHANICS

I knew of Heisenberg's theory, of course, but I felt discouraged, not to say repelled, by the methods of transcendental algebra, which appeared difficult to me, and by the lack of visualizability.

Erwin Schrödinger, 1926

The more I think about the physical portion of Schrödinger's theory, the more repulsive I find it. What Schrödinger writes about the visualizability of his theory is probably not quite right, in other words it's crap.

Werner Heisenberg, writing to Wolfgang Pauli, 1926

If you are receptive and humble, mathematics will lead you by the hand.

Paul A.M. Dirac, 1933.

Young man, in mathematics you don't understand things. You just get used to them.

John von Neumann.

As you learned in your *Quantum Physics I* course, Quantum mechanics is based on a set of postulates dictating the mathematical treatment of the theory and its interpretation. In this chapter, we will reformulate them using an algebraic formulation of quantum mechanics, the so-called transformation theory of Dirac. When formulating the postulates in this way, we will find a plethora of new terms and concepts, that will be immediately explained in detail. In particular, we will introduce the essential properties of Hilbert spaces. To illustrate the practical use of the postulates, we will consider the simple, but inherently quantum mechanical, Stern-Gerlach experiment presented in the previous chapter. This will allow us to break away from any classical concept or intuition, while paving the way to study more general systems.

3.1 States in a Hilbert space

Postulate I

A (pure) state of a quantum system is described by (the ray of) a vector in (the projective space of) a separable, complex Hilbert space. The state vector contains all the information about the system's quantum state.

As we have seen in the introduction, quantum mechanics is essentially a linear theory involving complex functions for which it is possible to compute a norm. It seems therefore rather natural to try to formulate it in a vector space with these properties.

A *complex Hilbert space* $\mathcal{H}$ is a vector space over the complex numbers $\mathbb{C}$ endowed with an *hermitian inner product* that maps any ordered pair of vectors to $\mathbb{C}$ and is *complete* with respect to the *norm*. This constitutes a generalization of finite-dimensional Euclidean vector spaces $\mathbb{R}^n$ to complex vector spaces with finite or infinite dimensions.

3.1.1 The superposition principle

Any element of $\mathcal{H}$ is called a *ket vector* or simply a *ket*. In Dirac notation, it is represented by the symbol $|\varphi\rangle$, with φ a distinctive label allowing to differentiate this specific vector from others and taking, depending on the situation, discrete or continuous values. Any linear combination

$$|\varphi\rangle = a|\varphi_1\rangle + b|\varphi_2\rangle\,, \tag{3.1}$$

with $|\varphi_1\rangle$ and $|\varphi_2\rangle$ in $\mathcal{H}$ and complex coefficients a and b, is also in $\mathcal{H}$. Physically, this is nothing else than the superposition principle leading to wave-like interference. This implies in particular that any vector $\lambda|\varphi\rangle$ with $\lambda \in \mathbb{C}^*$ describes the same physical state as $|\varphi\rangle$!

Rays in Hilbert space

Although not always sufficiently emphasized, a physical state does not strictly correspond to an element $|\varphi\rangle \in \mathcal{H}$ but rather to *a ray through the origin*.

A ray is an equivalence class of vectors that differ by multiplication by a nonzero complex scalar $\lambda \in \mathbb{C}^*$. Geometrically speaking, the equivalence $|\varphi\rangle \sim \lambda|\varphi\rangle$ for all $|\varphi\rangle \in \mathcal{H}/\{\mathbf{0}\}$ means that physical states live in the *projective Hilbert space* $\mathbb{P}\mathcal{H}$. Due to this property, one can always normalize the states to unity, leaving only an overall phase with no physical significance, i.e. $|\varphi\rangle$ and $e^{i\alpha}|\varphi\rangle$ describe the same state, with $\left|e^{i\alpha}\right| = 1$. However, relative phases are physically significant in superpositions. In particular, given two states $|\varphi_1\rangle, |\varphi_2\rangle$, we can identify $a|\varphi_1\rangle + b|\varphi_2\rangle$ with $e^{i\alpha}(a|\varphi_1\rangle + b|\varphi_2\rangle)$ but not with $a|\varphi_1\rangle + e^{i\alpha}b|\varphi_2\rangle$.

Superselection rules

In specific cases, there can exist *superselection rules* imposing restrictions on the types of states that can exist or be prepared in a quantum system. These rules emerge from symmetries and conservation laws and restrict the application of the superposition principle. For instance, the conservation of electric charge implies a superselection rule for charge particles, forbidding to express the state of a neutral particle, $|q=0\rangle$, as the superposition of two charged states $|+q\rangle$ and $|-q\rangle$.

Hilbert spaces follow the same rules of linear algebra as real vector spaces, but with the addition of complex scalar multiplication and complex conjugation. These rules include properties such as commutativity, associativity, distributivity and the existence of a unit element and a zero vector:

$$|\varphi_1\rangle + |\varphi_2\rangle = |\varphi_2\rangle + |\varphi_1\rangle\,, \qquad \text{(Commutativity of addition)} \tag{3.2}$$
$$(|\varphi_1\rangle + |\varphi_2\rangle) + |\varphi_3\rangle = |\varphi_1\rangle + (|\varphi_2\rangle + |\varphi_3\rangle)\,, \qquad \text{(Associativity of addition)} \tag{3.3}$$
$$a(|\varphi_1\rangle + |\varphi_2\rangle) = a|\varphi_1\rangle + a|\varphi_2\rangle\,, \qquad \text{(Distributivity of scalar multiplication)} \tag{3.4}$$
$$(a+b)|\varphi_1\rangle = a|\varphi_1\rangle + b|\varphi_1\rangle\,, \qquad \text{(Distributivity of scalar multiplication)} \tag{3.5}$$
$$(ab)|\varphi_1\rangle = a(b|\varphi_1\rangle)\,, \qquad \text{(Associativity of scalar multiplication)} \tag{3.6}$$
$$|\varphi\rangle - |\varphi\rangle = 0\,. \qquad \text{(Additive inverse)} \tag{3.7}$$

As usual, a set of vectors $|\varphi_1\rangle, |\varphi_2\rangle \ldots |\varphi_N\rangle$ are linearly independent if the condition

$$\sum_{i=1}^{N} c_i |\varphi_i\rangle = 0 \tag{3.8}$$

implies necessarily $c_1 = c_2 = \ldots = c_N = 0$. The dimensionality of the Hilbert space is determined by the largest possible number of linearly independent vectors we can find. If there exists not an upper number, we say the vector space has infinite dimension.

3.1.2 Dual space, inner product and norm

As any vector space, the dual space $\mathcal{H}^*$ of a Hilbert space $\mathcal{H}$ is defined as the space of linear functionals/maps from $\mathcal{H}$ to $\mathbb{C}$. In particular, the *ket* $|\varphi\rangle$ allows us to define a linear functional $\langle\chi|$ (called *bra*, dual vector, covector or one-form) that maps each element of $\mathcal{H}$ into a complex number equal to the scalar or inner product of $|\chi\rangle$ by $|\varphi\rangle$, that is,

$$\begin{aligned} \langle\cdot|\cdot\rangle : \mathcal{H}\times\mathcal{H} &\to \mathbb{C}\,, \\ (|\chi\rangle, |\varphi\rangle) &\mapsto \langle\chi|\varphi\rangle\,, \end{aligned} \tag{3.9}$$

with $\langle\cdot|\cdot\rangle$ the so-called *bra-ket* symbol ("bra(c)ket") popularized by Dirac[1]. As you probably demonstrated in your *Algebra* course, this provides a vector space isomorphism $\mathcal{H}^* \cong \mathcal{H}$ in finite dimensions. Non-trivially, the result also holds in infinite dimensions.

[1] This notation was first introduced by Hermann Grassmann, one hundred years earlier than Dirac.

Riesz theorem

For every ket vector $|\varphi\rangle \in \mathcal{H}$, there exists a unique bra vector $\langle\varphi| \in \mathcal{H}^*$, i.e.

$$|\varphi\rangle \in \mathcal{H} \longleftrightarrow \langle\varphi| \in \mathcal{H}^* \,.$$

In other words, a Hilbert space and its dual are naturally isomorphic.

The scalar product (3.9) satisfies the properties[2]

$$\langle\varphi|\varphi\rangle > 0 \quad \text{if} \quad |\varphi\rangle \neq 0\,, \qquad \langle\varphi|\varphi\rangle = 0 \quad \text{iff } |\varphi\rangle = 0\,, \qquad \text{(Positive-definiteness)} \tag{3.10}$$

$$\langle\chi|\varphi\rangle = \langle\varphi|\chi\rangle^*\,, \qquad \text{(Conjugate symmetry)} \tag{3.11}$$

$$\langle\chi|a\varphi_1 + b\varphi_2\rangle = a\langle\chi|\varphi_1\rangle + b\langle\chi|\varphi_2\rangle\,, \qquad \text{(Linearity in kets)} \tag{3.12}$$

$$\langle a\chi_1 + b\chi_2|\varphi\rangle = a^*\langle\chi_1|\varphi\rangle + b^*\langle\chi_2|\varphi\rangle\,, \qquad \text{(Antilinearity in bras)} \tag{3.13}$$

with * denoting complex conjugation and Eq. (3.13) following directly from Eqs. (3.11) and (3.12). By definition, two vectors are orthogonal to each other if their scalar product vanishes. Analogously, two subspaces $\mathcal{H}_1$ and $\mathcal{H}_2$ in $\mathcal{H}$ are orthogonal if each of the vectors in one is orthogonal to each of the vectors in the other, i.e. if $\mathcal{H}_1$ and $\mathcal{H}_2$ do not have common vectors.

The positivity of the scalar product $\langle\varphi|\varphi\rangle$ induces a norm[3]

$$\||\varphi\rangle\| \equiv \|\varphi\| = \sqrt{\langle\varphi|\varphi\rangle} \in \mathbb{R} \tag{3.14}$$

which generalizes the concept of modulus and allows us to define a distance

$$d(\varphi,\chi) = \|\varphi - \chi\| = \sqrt{\langle\varphi-\chi|\varphi-\chi\rangle} \tag{3.15}$$

between two arbitrary vectors φ, χ. $\mathcal{H}$ is therefore a *metric space*. The properties also ensure the important inequalities.

$$|\langle\chi|\varphi\rangle| \leq \|\chi\|\|\varphi\|\,, \qquad \text{(Cauchy-Schwarz inequality)} \tag{3.16}$$

$$\|\varphi - \chi\| \leq \|\varphi\| + \|\chi\|\,, \qquad \text{(Triangular inequality)} \tag{3.17}$$

with the equality in the last expression holding if and only if $|\varphi\rangle$ and $|\chi\rangle$ are proportional to each other. To prove the Cauchy-Schwartz inequality, let us define a state $|a\rangle = |\chi\rangle + a|\varphi\rangle$ with $a \in \mathbb{C}$. Taking into account the positive definiteness of the inner product,

$$\langle a|a\rangle = \langle\chi|\chi\rangle + a\langle\chi|\varphi\rangle + a^*\langle\varphi|\chi\rangle + |a|^2\langle\varphi|\varphi\rangle \geq 0\,,$$

and choosing $a = -\langle\varphi|\chi\rangle/\langle\varphi|\varphi\rangle$, we get

$$\langle\chi|\chi\rangle \geq \frac{|\langle\chi|\varphi\rangle|^2}{\langle\varphi|\varphi\rangle} \qquad \longrightarrow \qquad |\langle\chi|\varphi\rangle| \leq \|\chi\|\|\varphi\|\,.$$

[2] Note that in (3.13) the inner product is taken to be conjugate-linear in the first argument or bra. In the mathematical literature, the convention is usually the opposite.

[3] When convenient, we will rather use the notation $\||\varphi\rangle\|$.

To prove the triangular inequality, we consider the norm of the difference of two vectors $\|\varphi-\chi\|$ in (3.15). Expanding the inner product there as

$$\langle\varphi-\chi|\varphi-\chi\rangle=\langle\varphi|\varphi\rangle-\langle\varphi|\chi\rangle-\langle\chi|\varphi\rangle+\langle\chi|\chi\rangle\,,$$

we can write

$$\|\varphi-\chi\|^2=\|\varphi\|^2+\|\chi\|^2-2\mathrm{Re}(\langle\varphi|\chi\rangle)\leq\|\varphi\|^2+\|\chi\|^2-2|\langle\varphi|\chi\rangle|\,,$$

where $\mathrm{Re}(\langle\varphi|\chi\rangle)$ denotes the real part of $\langle\varphi|\chi\rangle$ and we have taken into account that $\mathrm{Re}(\langle\varphi|\chi\rangle)\leq|\langle\varphi|\chi\rangle|$. Using now the Cauchy-Schwarz inequality, $|\langle\varphi|\chi\rangle|\leq\|\varphi\|\|\chi\|$, we obtain

$$\|\varphi-\chi\|^2\leq\|\varphi\|^2+\|\chi\|^2+2\|\varphi\|\|\chi\|\quad\longrightarrow\quad\|\varphi-\chi\|^2\leq(\|\varphi\|+\|\chi\|)^2\quad\longrightarrow\quad\|\varphi-\chi\|\leq\|\varphi\|+\|\chi\|\,.$$

3.1.3 Separability

In the context of Hilbert spaces, the term *separable* refers to a specific property related to the cardinality or size of the set of vectors spanning the space. A Hilbert space $\mathcal{H}$ is said to be separable if admits a countable[4] (but not necessarily finite) orthonormal basis of vectors $\{|n\rangle\}\equiv\{|1\rangle,|2\rangle,\ldots,|n\rangle\}$ in $\mathcal{H}$ with $\langle n|m\rangle=\delta_{nm}$, such that any vector $|\varphi\rangle$ admits a decomposition

$$|\varphi\rangle=\sum_n c_n|n\rangle\,,\tag{3.18}$$

with $c_n=\langle n|\varphi\rangle$ the components of $|\varphi\rangle$ in that basis. In terms of this above basis, a ket state can be written as a column vector

$$|\varphi\rangle=\begin{pmatrix}c_1\\c_2\\\vdots\\c_n\end{pmatrix}\,,\tag{3.19}$$

a bra becomes a row vector

$$\langle\chi|=(d_1^*,d_2^*,\ldots,d_n^*)\,,\tag{3.20}$$

and the inner product becomes just the matrix product of a row vector and column vector

$$\langle\chi|\varphi\rangle=(d_1^*,d_2^*,\ldots,d_n^*)\begin{pmatrix}c_1\\c_2\\\vdots\\c_n\end{pmatrix}\,.\tag{3.21}$$

Finally, the squares norm becomes

$$\||\varphi\rangle\|^2=\langle\varphi|\varphi\rangle=\sum_n|c_n|^2\,.\tag{3.22}$$

[4] A set is countable if its elements can be put in a one-to-one correspondence with the natural numbers (1, 2, 3, ...). For example, the set of integers is countable. If a set is not countable, it is uncountable.

Example 2: Countable but not finite basis

The eigenfunctions of an infinite well of width L, namely $V(x) = 0$ for $0leqx \leq L$ and ∞ otherwise, form a *countable infinite basis*

$$\varphi_n(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right), \quad n = 1, 2, 3, 4, \ldots \tag{3.23}$$

3.1.4 Completeness

In mathematics, a Cauchy sequence is a sequence of elements $\{\varphi_n\} = \varphi_1, \varphi_2, \ldots$ in a metric or normed vector space that becomes arbitrarily close to each other as the sequence progresses; more formally

$$\forall \epsilon > 0, \quad \exists N \quad : \quad \|\varphi_m - \varphi_n\| \leq \epsilon \quad \forall\, n, m > N\,. \tag{3.24}$$

Intuitively, this means that the elements become arbitrarily close to each other as we progress through the sequence. No matter how small a positive value ϵ we choose, there always exists a point in the sequence beyond which the distance between any two elements is always smaller than ϵ. A seminal example of a space that is not complete is the set of rational numbers $\mathbb{Q}$, which is a subspace of the real numbers $\mathbb{R}$

Example 3: A simple Cauchy sequence

Let us have a look at the sequence $\{a_1, a_2, \ldots, a_n\}$ con $a_n = 1/n$. This sequence is Cauchy. Consider, for instance, a choice of $\epsilon = 1/10$. Then, for $N > 2/\epsilon = 20$, we have

$$\left\|\frac{1}{m} - \frac{1}{n}\right\| < \frac{1}{10}\,, \qquad \forall\, n, m > N\,.$$

To see this explicitly, we can consider 2 very separate elements in the sequence

$$\left\|\frac{1}{22 \times 10^6} - \frac{1}{22}\right\| = \frac{10^6 - 1}{22 \times 10^6} < \frac{10^6}{22 \times 10^6} = \frac{1}{22} < \frac{1}{10}\,.$$

A normed vector space $(\mathcal{H}, \langle\cdot|\cdot\rangle)$ is considered complete if every Cauchy sequence in $\mathcal{H}$ converges within the same Hilbert space $\mathcal{H}$, i.e. if for any sequence $\{\varphi_n\} = \varphi_1, \varphi_2, \ldots$ in $\mathcal{H}$ where the distance $d(\varphi_n, \varphi_m)$ tends to zero as $n, m \to \infty$, there exists an element χ in $\mathcal{H}$ such that $d(\varphi_n, \chi)$ approaches zero as n approaches infinity. Heuristically, this means there are no points 'missing' from $\mathcal{H}$.

Complete normed vector spaces are called Banach spaces. The condition of completeness is important in infinite-dimensional function spaces as it guarantees the convergence of specific eigenfunction expansions.

Example 4: A uniform spatial probability density

Consider a uniform spatial probability density throughout an infinite potential well of width L,

$$\varphi(x) = \frac{1}{\sqrt{L}}\,.$$

Albeit this wave function cannot be written as a finite linear combination of energy eigenvectors, it can be expressed as the sum of the convergent series

$$\varphi(x) = \sum_{n=1}^{\infty} \frac{2\sqrt{2}}{(2n-1)\pi}\varphi_{2n-1}(x) \quad \text{with} \quad \varphi_n(x) = \sqrt{\frac{2}{L}}\sin\left(\frac{n\pi x}{L}\right)\,.$$

To prove that the above sequence is a Cauchy sequence, we need to show that for any given positive real number ϵ, there exists an integer N such that for all m, $n \geq N$, the absolute difference between the m-th and n-th terms of the sequence is less than ϵ. In other words, we want to prove that

$$|\varphi_m(x) - \varphi_n(x)| < \epsilon \quad \text{for all} \quad m, n \geq N\,.$$

Taking into account the form of the m-th and n-th terms,

$$\varphi_m(x) = \frac{2\sqrt{2}}{(2m-1)\pi}\varphi_{2m-1}(x) = \frac{4}{(2m-1)\pi}\sin\left(\frac{(2m-1)\pi x}{L}\right)\,,$$
$$\varphi_n(x) = \frac{4}{(2n-1)\pi}\sin\left(\frac{(2n-1)\pi x}{L}\right)\,,$$

we have

$$\begin{aligned}|\varphi_m(x) - \varphi_n(x)| &= \left|\frac{4}{(2m-1)\pi}\sin\left(\frac{(2m-1)\pi x}{L}\right) - \frac{4}{(2n-1)\pi}\sin\left(\frac{(2n-1)\pi x}{L}\right)\right| \\ &= \frac{4}{\pi}\left|\frac{1}{2m-1}\sin\left(\frac{(2m-1)\pi x}{L}\right) - \frac{1}{2n-1}\sin\left(\frac{(2n-1)\pi x}{L}\right)\right|\,.\end{aligned}$$

Applying now the triangle inequality we get

$$|\varphi_m(x) - \varphi_n(x)| \leq \frac{4}{\pi}\left|\frac{1}{2m-1}\sin\left(\frac{(2m-1)\pi x}{L}\right)\right| + \frac{4}{\pi}\left|\frac{1}{2n-1}\sin\left(\frac{(2n-1)\pi x}{L}\right)\right|\,.$$

Since both terms on the right-hand side involve sine functions, we can use the fact that $|\sin(x)| \leq 1$ for all x to simplify further,

$$|\varphi_m(x) - \varphi_n(x)| \leq \frac{4}{\pi}\left(\frac{1}{2m-1} + \frac{1}{2n-1}\right)\,.$$

Now, we need to find an N such that for all m, $n \geq N$, the right-hand side of the above inequality is less than ϵ,

$$\frac{4}{\pi}\left(\frac{1}{2m-1} + \frac{1}{2n-1}\right) < \epsilon\,.$$

Choosing N such that

$$\frac{4}{\pi}\left(\frac{1}{2N-1}+\frac{1}{2N-1}\right)<\epsilon\,,$$

and simplifying, we get

$$N>\frac{8}{\pi\epsilon}+1\,.$$

Since ϵ is a positive real number, the expression on the right-hand side is a constant. Therefore, for any ϵ, you can choose an N such that for all m, $n\geq N$, the absolute difference between the m-th and n-th terms of the sequence is less than ϵ. This satisfies the definition of a Cauchy sequence, and thus, $\varphi(x)$ is a Cauchy sequence.

In this course, however, we will be mainly interested in inner-product spaces that have finite dimensions and are therefore always complete. If you are interested in the more complicated treatment of Hilbert spaces of infinite dimensions, you are suggested to enroll the UCM *Quantum Mechanics* course, where this topic is usually covered.

Example 5: Spin 1/2 bases and the Block sphere

We have seen in the previous chapter that when spin-1/2 atoms pass through a Z-type Stern-Gerlach device, they separate into two beams of particles with spin projections $\pm\hbar/2$ along the z-axis. This suggests that the spin states of the electron can be described using two basis states or kets

$$|+\rangle\longleftrightarrow S_z=\hbar/2\quad,\quad|-\rangle\longleftrightarrow S_z=-\hbar/2\,,\tag{3.25}$$

commonly known as *spin up* and *spin down* states This is the simplest non-trivial example of a Hilbert space. The signs in these expressions indicate only that the z-component of spin is positive or negative and could be replaced by any other distinctive label you like, such as arrows $|\uparrow\rangle$, $|\downarrow\rangle$, giving you some (inaccurate) feeling of "rotation around an axis" or even numbers $|1\rangle$, $|2\rangle$ indicating that you are dealing with a two-level system. Note also that we are not yet accounting for other possible degrees of freedom such as the position of the atom, its momentum or energy. We will come back to these things later.
As we learned in Experiment 1, the two states (3.25) are orthonormal,

$$\langle+|+\rangle=\langle-|-\rangle=1,\qquad\langle+|-\rangle=\langle-|+\rangle=0\,.\tag{3.26}$$

In particular, after passing through the second Z analyzer, all previously prepared $|+\rangle$ states emerge again as $|+\rangle$ states; no down spins appears.
An explicit representation of the basis vectors satisfying the above normalization condition could be for instance

$$|+\rangle=\begin{pmatrix}1\\0\end{pmatrix},\qquad|-\rangle=\begin{pmatrix}0\\1\end{pmatrix},\tag{3.27}$$

with the corresponding bras obtained by transposition and complex conjugation,

$$\langle +| = (1,0)\,, \qquad \langle -| = (0,1)\,. \tag{3.28}$$

A general ket or bra in the two-dimensional space spanned by the above basis vectors

$$|\varphi\rangle = a|+\rangle + b|-\rangle \in \mathcal{H} \quad \longleftrightarrow \quad \langle\varphi| = a^*\langle +| + b^*\langle -| \in \mathcal{H}^* \tag{3.29}$$

is completely defined by a set of two complex coefficients $a = |a|e^{i\alpha}$ and $b = |b|e^{i\beta}$, i.e. by four real parameters. The normalization condition $\langle\varphi|\varphi\rangle = 1$ implies, however, a relation

$$\begin{aligned} 1 = \langle\varphi|\varphi\rangle &= (a^*\langle +| + b^*\langle -|)(a|+\rangle + b|-\rangle) \\ &= |a|^2\langle +|+\rangle + a^*b\langle +|-\rangle + b^*a\langle -|+\rangle + |b|^2\langle -|-\rangle = |a|^2 + |b|^2, \end{aligned}$$

or

$$|a|^2 + |b|^2 = 1\,, \tag{3.30}$$

reducing the number of linearly independent real parameters to 3. Moreover, since the state is defined as a ray-vector in the Hilbert space, we can rewrite (3.29) as

$$|\varphi\rangle = e^{i\alpha}\left(|a||+\rangle + |b|e^{i(\beta-\alpha)}|-\rangle\right)\,, \tag{3.31}$$

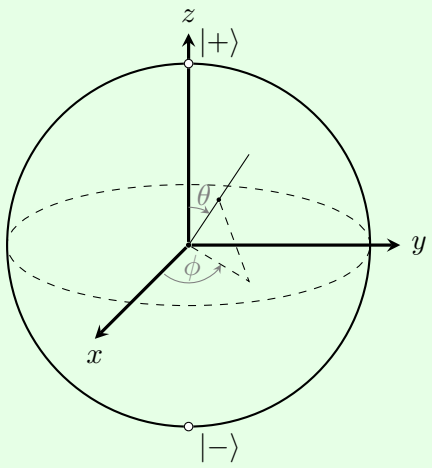
and simply neglect the irrelevant phase factor $e^{i\alpha}$. This leaves us with two real numbers that can be naturally represented by a point on a spherical surface called the *Block sphere*. This point is determined by the expectation value of the three Pauli matrices (see below) or alternatively by the polar angles θ and ϕ. In particular, choosing

$$|a| \equiv \cos(\theta/2)\,, \qquad |b| \equiv \sin(\theta/2)e^{i\phi}\,, \qquad \phi = \beta - \alpha\,, \tag{3.32}$$

with θ ranging from 0 to π and ϕ ranging from 0 to 2π, we can write the convenient expression

$$|\varphi\rangle = \cos\frac{\theta}{2}|+\rangle + \sin\frac{\theta}{2}e^{i\phi}|-\rangle\,. \tag{3.33}$$

The correspondence between quantum states in $\mathbb{C}^2$ and points on the unit-radius Block sphere in $\mathbb{R}^3$ is apparent in this figure.



The vector pointing towards the north pole represents the state $|+\rangle$ while that pointing towards the south pole represents the state $|-\rangle$. On the other hand, all vectors in the equatorial plane stand for symmetric superpositions of the two basis vectors, namely ($\theta = \pi/2$, $\sin(\pi/4) = \cos(\pi/4) = 1/\sqrt{2}$)

$$|\varphi\rangle = \frac{1}{\sqrt{2}}\left(|+\rangle + e^{i\phi}|-\rangle\right) . \tag{3.34}$$

We can make use of this expression to construct states $|+\rangle_x$ representing the spin projections $S_x = \pm\hbar/2$ along the x-axis. Since these must be orthogonal to each other, but not orthogonal to S_z in order to let half of the atoms go through in Experiment 2, we have ($\theta = \pi/2$, $\phi = 0, e^0 = 1$, $\phi = \pi, e^{i\pi} = -1$)

$$|+\rangle_x = \frac{1}{\sqrt{2}}\left(|+\rangle + |-\rangle\right) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} , \quad |-\rangle_x = \frac{1}{\sqrt{2}}\left(|+\rangle - |-\rangle\right) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} , \tag{3.35}$$

where in the last steps we have used the explicit representation (3.27). As anticipated, these states satisfy

$${}_x\langle -|+\rangle_x = {}_x\langle +|-\rangle_x^* = \frac{1}{2}\begin{pmatrix} 1 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 0 .$$

Note also that they are normalized

$${}_x\langle +|+\rangle_x = \frac{1}{2}\begin{pmatrix} 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1 , \qquad {}_x\langle -|-\rangle_x = \frac{1}{2}\begin{pmatrix} 1 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = 1 .$$

Similarly, one could construct states $|+\rangle_y$ representing the spin projections $S_y = \pm\hbar/2$ along the y-axis ($\theta = \pi/2$, $\phi = \pi/2, e^{i\pi/2} = i$, $\phi = \pi, e^{3i\pi/2} = -i$),

$$|+\rangle_y = \frac{1}{\sqrt{2}}\left(|+\rangle + i|-\rangle\right) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix} , \quad |-\rangle_y = \frac{1}{\sqrt{2}}\left(|+\rangle - i|-\rangle\right) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \end{pmatrix} . \tag{3.36}$$

Note that this superposition of $|\pm\rangle$ states involves complex numbers (there would be no way to find the y-type states without them). Special attention must be therefore taken when determining the associated bras and operating with them. For instance, we have

$$\langle +|_y = \frac{1}{\sqrt{2}}\left(\langle +| - i\langle -|\right) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -i \end{pmatrix} ,$$

and

$${}_y\langle +|-\rangle_y = {}_y\langle -|+\rangle_y^* = \frac{1}{2}\begin{pmatrix} 1 & -i \end{pmatrix}\begin{pmatrix} 1 \\ -i \end{pmatrix} = 0 .$$

3.2 Observables

Postulate II

Every measurable physical quantity in quantum mechanics is associated with a Hermitian or self-adjoint operator acting on the system's Hilbert space. This operator is said to be an *observable*.

Once we have defined the space of ket vectors, we can define the operators in that space. In Hilbert spaces, operators are linear transformations $A : \mathcal{H} \to \mathcal{H}$ mapping vectors in the Hilbert space to other vectors in the same Hilbert space, $A : |\varphi\rangle \mapsto A|\varphi\rangle$, in a way compatible with the vector space structure and therefore with the superposition principle, namely

$$A(a|\varphi_1\rangle + b|\varphi_2\rangle) = aA|\varphi_1\rangle + bA|\varphi_2\rangle\,, \tag{3.37}$$

with $|\varphi\rangle$, $|\varphi_1\rangle$, $|\varphi_2\rangle$ states in $\mathcal{H}$ and a and b complex coefficients. Given two linear operators A and B, we define their sum and product as

$$(A+B)|\varphi\rangle = A|\varphi\rangle + B|\varphi\rangle\,, \qquad (AB)|\varphi\rangle = A(B|\varphi\rangle)\,, \tag{3.38}$$

for all $|\varphi\rangle \in \mathcal{H}$. Note that the operator algebra is associative,

$$A(BC) = (AB)C\,, \tag{3.39}$$

but not necessarily commutative, meaning that, in general, the order of the operators is important $AB \neq BA$. The difference of two actuation's orders is given by the commutator

$$[A,B] \equiv AB - BA\,. \tag{3.40}$$

When this quantity vanishes, A and B are said commute; otherwise they are said to not commute. The commutator obeys a bunch of useful properties,

$$\begin{aligned}
&[A,B] = -[B,A]\,, && \text{(Antisymmetry)} && (3.41)\\
&[aA_1 + bA_2, B] = a\,[A_1,B] + b\,[A_2,B]\,, \quad \forall a,b \in \mathbb{C} && \text{(Linearity)} && (3.42)\\
&[A,BC] = [A,B]C + B[A,C]\,, && \text{(Leibniz identity)} && (3.43)\\
&[AB,C] = A[B,C] + [A,C]B\,, && \text{(Leibniz identity)} && (3.44)\\
&[A,[B,C]] + [B,[C,A]] + [C,[A,B]] = 0\,, && \text{(Jacobi identity)} && (3.45)\\
&[AB,CD] = [A,C]DB + C[A,D]B + A[B,C]D + AC[D,B]\,, && && (3.46)\\
&[f(A),A] = 0\,. && && (3.47)
\end{aligned}$$

whose proof can be lengthy but completely straightforward. Consider, for instance the Leibniz identity (3.44). Starting with the definition of the commutator (3.40) and taking into account the associative property (3.39), we get

$$\begin{aligned}
[AB,C] &= (AB)C - C(AB) = A(BC) - (CA)B = A(BC) - A(CB) + A(CB) - (CA)B\\
&= A(BC - CB) + (ACB - CAB) = A[B,C] + [A,C]B\,.
\end{aligned}$$

Similarly, for the Jacobi identity (3.45), we have

$$\begin{aligned}[A,[B,C]]+[B,[C,A]]+[C,[A,B]] &= [A,BC]-[A,CB]+[B,CA]-[B,AC]+[C,AB]-[C,BA]\\ &=[A,B]C+B[A,C]-[A,C]B-C[A,B]+[B,C]A+C[B,A]-[B,A]C\\ &\quad -A[B,C]+[C,A]B+A[C,B]-[C,B]A-B[C,A]\\ &=2([A,B]C-C[A,B]-[A,C]B+B[A,C]+[B,C]A-A[B,C])\\ &=2(ABC-BAC-CAB+CBA-ACB+CAB+BAC-BCA+BCA\\ &\quad -CBA-ABC+ACB)=0\,,\end{aligned}$$

and so on and so forth.

The inverse of a linear operator $A:\mathcal{H}\to\mathcal{H}$ is another linear operator A^{-1} satisfying the conditions

$$A^{-1}(A|\varphi\rangle)=|\varphi\rangle,\qquad\text{and}\qquad A\left(A^{-1}|\varphi\rangle\right)=|\varphi\rangle\,. \tag{3.48}$$

The inverse of a product of operators A and B is $(AB)^{-1}=B^{-1}A^{-1}$, provided of course that the inverse operators A^{-1} and B^{-1} exist.[5]

Bounded and unbounded operators

If there exists a real positive number C such that

$$\|A|\varphi\rangle\|<C\|\varphi\|\,,\qquad\forall\varphi\in\mathcal{H}\,, \tag{3.49}$$

the operator A is said to be *bounded*; otherwise it is said to be *unbounded*. The norm of the operator, $\|A\|$ is defined as the lower limit of the constant C satisfying the above equation, namely

$$||A||=\sup_{\varphi\neq0}\frac{\|A|\varphi\rangle\|}{\|\varphi\|}\,. \tag{3.50}$$

In other words the operator is bounded if $\|A\|<\infty$. In a finite-dimensional Hilbert space all operators are bounded. This is not the case, however, in an infinite dimensional Hilbert space. Common example of unbounded operators are for instance the position and momentum operators you are all familiar with and the Hamiltonian of the harmonic oscillator,

$$H=\frac{P^2}{2m}+\frac{1}{2}m\omega^2X^2\,, \tag{3.51}$$

since there exists states $H|\varphi_n\rangle=E_n|\varphi_n\rangle$, $\|\varphi\|=1$ with an arbitrarily large value of E_n.

The necessity of dealing with unbounded operators leads to situations in which their action on vectors in $\mathcal{H}$ does not generate a state in $\mathcal{H}$. In those cases, it is necessary to define the domain of an operator A,

$$|\varphi\rangle\in D(A)\subset\mathcal{H}\quad\text{ if }\quad A|\varphi\rangle\in\mathcal{H}\,. \tag{3.52}$$

[5] An inverse exists if the matrix representing the operator has a non-vanishing determinant, as discussed below.

3.2.1 Adjoint operators

Using the one-to-one correspondence between bras and conjugate kets, let us derive an analogous conjugation relation for linear operators. Consider a linear operator A and a ket $|\phi\rangle$ conjugate to the bra $\langle\varphi|A$. Since $|\phi\rangle$ depends antilinearly on $\langle\varphi|$, it must be linear in $|\varphi\rangle$, namely

$$|\phi\rangle = A^\dagger|\varphi\rangle\,, \tag{3.53}$$

with $A^\dagger$ a linear operator called the *Hermitian conjugate* or *adjoint operator* of A. Now, as $A^\dagger|\varphi\rangle$ is the ket conjugate to the bra $\langle\varphi|A$, the scalar product of this ket by an arbitrary bra $\langle\chi|$ is the complex conjugate of the scalar product of $|\chi\rangle$ by $\langle\varphi|A$,

$$\langle\chi|A^\dagger|\varphi\rangle = \langle\varphi|A|\chi\rangle^*\,. \tag{3.54}$$

Since this must hold for any $|\varphi\rangle$ and $|\chi\rangle$, the ket conjugate to $\langle\chi|A^\dagger$ must be necessarily equal to $A|\chi\rangle$,

$$(\langle\chi|A^\dagger)^\dagger = A|\chi\rangle\,, \qquad (A|\varphi\rangle)^\dagger = \langle\varphi|A^\dagger\,. \tag{3.55}$$

Consequently, the conjugate of the operator $A^\dagger$ is the operator A itself,

$$(A^\dagger)^\dagger = A\,, \tag{3.56}$$

Additionally, we can highlight the useful properties

$$(cA)^\dagger = c^*A^\dagger\,, \qquad (A+B)^\dagger = A^\dagger + B^\dagger\,, \qquad (AB)^\dagger = B^\dagger A^\dagger\,, \qquad [A,B]^\dagger = [B^\dagger, A^\dagger]\,, \tag{3.57}$$

with $c \in \mathbb{C}$. In other words, Hermitian conjugation is to operators what conjugation between bras and kets is to vectors and complex conjugation is to ordinary numbers.

Note finally that in matrix form, the Hermitian conjugate of A becomes the conjugate transpose $\left(A^T\right)^*$ of its associate matrix,

$$\langle\chi|A^\dagger|\varphi\rangle = \langle\varphi|A|\chi\rangle^* \quad \longrightarrow \quad (A^\dagger)_{mn} = A^*_{nm}\,,. \tag{3.58}$$

Example 6: The adjoint operation

To determine the bras' equation associated to a ket equation

$$|\varphi\rangle = aA|\chi\rangle + BC|\phi\rangle\,,$$

with $a \in \mathbb{C}$ and A, B and C linear operators in Hilbert space, we consider its product with an arbitrary bra $\langle\xi|$. This provides an equation between complex numbers,

$$\langle\xi|\varphi\rangle = a\langle\xi|A|\chi\rangle + \langle\xi|BC|\phi\rangle\,,$$

which can be subsequently conjugated to obtain

$$\langle\xi|\varphi\rangle^* = a^*\langle\xi|A|\chi\rangle^* + \langle\xi|BC|\phi\rangle^*\,.$$

Applying the properties of the inner scalar product on the left-hand side and the definition of the Hermitian conjugate in (3.54) on the right one, this expression can be rewritten as

$$\langle\varphi|\xi\rangle = a^*\langle\chi|A^\dagger|\xi\rangle + \langle\phi|(BC)^\dagger|\xi\rangle\,,$$

which upon canceling the arbitrary vector $|\xi\rangle$ provides the searched equation for bras',

$$\langle\varphi| = a^*\langle\chi|A^\dagger + \langle\phi|C^\dagger B^\dagger\,.$$

In practice, one typically skips all the above intermediate steps and directly writes the final expression by inspection.

3.2.2 Projection operators and resolution of the identity

The beauty of the Dirac notation lies on its ability to integrate vectors, linear operators and inner products into a single framework, allowing us to manipulate these quantities in a clear and intuitive way. Consider, for instance, the following expression and its interpretation within the Dirac framework,

$$(\underbrace{|n\rangle\langle m|}_{\text{operator}})\,\underbrace{|\varphi\rangle}_{\text{state}} = \underbrace{|n\rangle}_{\text{state}}\,(\underbrace{\langle m|\varphi\rangle}_{\text{number}})\,, \tag{3.59}$$

where we have taken into account that all operations with bras, kets and linear operators are associative, i.e. that we can order them the way we want provided we keep their order from left to right. As suggested by the notation, the outer product $|n\rangle\langle m|$ is not a number, but rather a linear operator. This can be easily verified by considering its action on a given superposition state in $(a|\varphi_1\rangle + b|\varphi_2\rangle) \in \mathcal{H}$,

$$|n\rangle\langle m|(a|\varphi_1\rangle + b|\varphi_2\rangle) = |n\rangle(a\langle m|\varphi_1\rangle + b\langle m|\varphi_2\rangle) = \underbrace{a\langle m|\varphi_1\rangle}_{\in\mathbb{C}}\,|n\rangle + \underbrace{b\langle m|\varphi_2\rangle}_{\in\mathbb{C}}\,|n\rangle \in \mathcal{H}\,. \tag{3.60}$$

Note also that Hermitian conjugation is naturally encoded by the notation. For instance, we have

$$(|n\rangle\langle m|)^\dagger|\varphi\rangle = (\langle\varphi|n\rangle\langle m|)^\dagger = |m\rangle\langle\varphi|n\rangle^* = |m\rangle\langle n|\varphi\rangle = (|m\rangle\langle n|)|\varphi\rangle, \quad \forall|\varphi\rangle. \tag{3.61}$$

and therefore

$$(|n\rangle\langle m|)^\dagger = |m\rangle\langle n|\,. \tag{3.62}$$

It immediately follows from (3.59) that an operator

$$P_n \equiv |n\rangle\langle n| \tag{3.63}$$

acts essentially as a filter, mapping any vector $|\varphi\rangle \in \mathcal{H}$ into a vector along $|n\rangle$, i.e. $P_n|\varphi\rangle = |n\rangle\langle n|\varphi\rangle \sim |n\rangle$. This operator is indeed a (self-adjoint) projection operator satisfying the idempotent property $P_n^2 = P_n$, as can be again easily verified in Dirac notation,

$$P_nP_n = (|n\rangle\langle n|)(|n\rangle\langle n|) = |n\rangle\langle n|n\rangle\langle n| = |n\rangle\langle n| = P_n\,. \tag{3.64}$$

If the vectors $\{|n\rangle\}$ form a basis of the Hilbert space, we can use the associated projection operators to obtain a very useful decomposition of the identity operator,

$$\sum_n |n\rangle\langle n| = \mathbb{I}\,, \tag{3.65}$$

as can be easily verified by acting on a general state $|\varphi\rangle = \sum_m c_m|m\rangle$,

$$\left[\sum_n |n\rangle\langle n|\right]|\varphi\rangle = \sum_m c_m \sum_n |n\rangle \underbrace{\langle n|m\rangle}_{\delta_{nm}} = \sum_m c_m|m\rangle = |\varphi\rangle\,. \tag{3.66}$$

This is the so-called *completeness condition* or *resolution of the identity* for the orthonormal set $\{|n\rangle\}$ and it is particularly useful when moving between bases.

Forbidden quantities

Quantities of the type $|\varphi\rangle|\chi\rangle$ make no sense for states belonging $|\varphi\rangle$ and $|\chi\rangle$ belonging to the same Hilbert space. They are neither bras or kets. They are allowed, however, if the states belong to different vector spaces (see Section 5.1).

3.2.3 Bases and representations

Depending on the specific Hilbert space under consideration and the nature of the problem, the relevant operators can take different forms, ranging from matrices to differential operators. If $\mathcal{H}$ is finite dimensional, an operator A in a given orthonormal basis $\{|n\rangle\}$ in $\mathcal{H}$ takes the form of a matrix, with elements A_{mn}. To obtain these entries, let as assume

$$A|\varphi\rangle = |\chi\rangle \tag{3.67}$$

and expand the vectors $|\varphi\rangle$ and $|\chi\rangle$ in terms of the orthonormal basis,

$$|\varphi\rangle = \mathbb{I}|\varphi\rangle = \sum_n |n\rangle\langle n|\varphi\rangle = \sum_n \underbrace{\langle n|\varphi\rangle}_{c_n}|n\rangle\,, \tag{3.68}$$

$$|\chi\rangle = \mathbb{I}|\chi\rangle = \sum_n |n\rangle\langle n|\chi\rangle = \sum_n \underbrace{\langle n|\chi\rangle}_{d_n}|n\rangle\,. \tag{3.69}$$

Making use of the linearity property (3.37), we can rewrite (3.67) as

$$\sum_n c_n A|n\rangle = \sum_n d_n|n\rangle\,, \tag{3.70}$$

which, multiplying on the left by $\langle m|$,

$$\sum_n c_n\langle m|A|n\rangle = \sum_n d_n\langle m|n\rangle = \delta_{mn}d_n = d_m\,, \tag{3.71}$$

gives us

$$d_m = \sum_n A_{mn} c_n \,, \tag{3.72}$$

with

$$A_{mn} = \langle m|A|n\rangle \tag{3.73}$$

the so-called *transition matrix element*. In this *basis-dependent* representation, the *basis-independent* equation (3.67) becomes simply the multiplication of a vector by a matrix,

$$\begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix} = \begin{pmatrix} A_{11} & \dots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \dots & A_{nn} \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} . \tag{3.74}$$

Similarly, the trace becomes the sum of the diagonal matrix elements,

$$\mathrm{Tr}\, A = \sum_n \langle n|A|n\rangle = \sum_n A_{nn} \,. \tag{3.75}$$

Note also that the matrix elements of the product of two operators A and B satisfies the usual rules of matrix multiplication, namely $(AB)_{mn} = \sum_i A_{mi} B_{in}$, as can be easily verified by using the identity resolution (3.65),

$$(AB)_{mn} = \langle m|AB|n\rangle = \langle m|A\Big(\sum_i |i\rangle\langle i|\Big)B|n\rangle = \sum_i \langle m|A|i\rangle\langle i|B|n\rangle = \sum_i A_{mi} B_{in} \,. \tag{3.76}$$

3.2.4 Hermitian operators and spectral decomposition

An operator A is called self-adjoint or Hermitian if it is equal to its adjoint $A^\dagger$,

$$A = A^\dagger \,, \tag{3.77}$$

so that

$$\langle \chi|A|\varphi\rangle = \langle \varphi|A|\chi\rangle^* \,. \tag{3.78}$$

Hermitian operators have three crucial properties that we will use constantly in the following:

- The Hermiticity of an operator ensures that their eigenvalues are real, allowing for meaningful interpretations of measurable quantities. Indeed, if A is a Hermitian operator in $\mathcal{H}$ with eigenvectors $|n\rangle \neq 0$ and eigenvalues $a_n \in \mathbb{C}$,

  $$A|n\rangle = a_n|n\rangle \,, \tag{3.79}$$

  the property (3.78) implies

  $$\langle n|A|n\rangle = \langle n|A|n\rangle^* \quad \longrightarrow \quad a_n\langle n|n\rangle = a_n^*\langle n|n\rangle \quad \longrightarrow \quad a_n = a_n^* \,. \tag{3.80}$$

- The eigenvectors $|n\rangle$ and $|m\rangle$ of Hermitian operators corresponding to distinct eigenvalues a_n and a_m ($a_n \neq a_m$) are orthogonal to each other, i.e. $\langle n|m\rangle = \delta_{nm}$, as can be easily verified,

$$0 = \langle n|A|m\rangle - \langle m|A|n\rangle^* = a_m\langle n|m\rangle - a_n\langle n|m\rangle = (a_m - a_n)\langle n|m\rangle\,. \tag{3.81}$$

  Note that if several eigenfunctions share the same eigenvalue (*degenerate states*), any linear combination of them will be still an eigenfunction with the same eigenvalue. This allows us to use the Gram-Schmidt orthogonalization you studied in your *Algebra* course to construct eigenfunctions within each degenerate subspace. Therefore, even in the presence of degeneracy, the eigenfunctions can be chosen to be orthogonal.

Gram-Schmidt decomposition

Given a complete basis $\{|\varphi_i\rangle\}$ of the Hilbert space $\mathcal{H}$, the sequence of states

$$|\chi_1\rangle = \frac{|\varphi_1\rangle}{\||\varphi_1\rangle\|}, \qquad |\chi_k\rangle = \frac{|\varphi_k\rangle - \sum_{j=1}^{k-1}\langle\chi_j|\varphi_k\rangle\,|\chi_j\rangle}{\||\varphi_k\rangle - \sum_{j=1}^{k-1}\langle\chi_j|\varphi_k\rangle\,|\chi_j\rangle\|} \qquad \forall k > 1$$

forms an orthonormal basis of $\mathcal{H}$. To show that $\langle\chi_i|\chi_j\rangle = \delta_{ij}$ we proceed by induction:

$$\begin{aligned}
\langle\chi_k|\chi_k\rangle &= \frac{\||\varphi_k\rangle - \sum_{j=1}^{k-1}\langle\chi_j|\varphi_k\rangle\,|\chi_j\rangle\|^2}{\||\varphi_k\rangle - \sum_{j=1}^{k-1}\langle\chi_j|\varphi_k\rangle\,|\chi_j\rangle\|^2} = 1\,,\\
\langle\chi_1|\chi_2\rangle &= \frac{1}{\||\varphi_1\rangle\|\,\||\varphi_2\rangle - |\chi_1\rangle\langle\chi_1|\varphi_2\rangle\|}\left(\langle\varphi_1|\varphi_2\rangle - \frac{\langle\varphi_1|\varphi_1\rangle}{\||\varphi_1\rangle\|^2}\langle\varphi_1|\varphi_2\rangle\right)\\
&= \frac{1}{\||\varphi_1\rangle\|\,\||\varphi_2\rangle - |\chi_1\rangle\langle\chi_1|\varphi_2\rangle\|}\left(\langle\varphi_1|\varphi_2\rangle - \langle\varphi_1|\varphi_2\rangle\right) = 0\,.
\end{aligned}$$

Assuming now that $\langle\chi_1|\chi_j\rangle = \delta_{1j}$ we show that $\langle\chi_1|\chi_{j+1}\rangle = \delta_{1,j+1}$:

$$\begin{aligned}
\langle\chi_1|\chi_{j+1}\rangle &= \frac{1}{\||\varphi_{j+1}\rangle - \sum_{k=1}^{j}|\chi_k\rangle\langle\chi_k|\varphi_{j+1}\rangle\|}\left\langle\chi_1\middle|\left(|\varphi_{j+1}\rangle - \sum_{k=1}^{j}|\chi_k\rangle\langle\chi_k|\varphi_{j+1}\rangle\right)\right.\\
&= \frac{1}{\||\varphi_{j+1}\rangle - \sum_{k=1}^{j}|\chi_k\rangle\langle\chi_k|\varphi_{j+1}\rangle\|}\Big(\langle\chi_1|\varphi_{j+1}\rangle - \sum_{k=1}^{j}\underbrace{\langle\chi_1|\chi_k\rangle}_{\delta_{1k}}\langle\chi_k|\varphi_{j+1}\rangle\Big)\\
&= \frac{1}{\||\varphi_{j+1}\rangle - \sum_{k=1}^{j}|\chi_k\rangle\langle\chi_k|\varphi_{j+1}\rangle\|}\left(\langle\chi_1|\varphi_{j+1}\rangle - \langle\chi_1|\varphi_{j+1}\rangle\right) = 0\,.
\end{aligned}$$

Finally, assuming that $\langle\chi_i|\chi_j\rangle = \delta_{ij}$ we show that $\langle\chi_{i+1}|\chi_j\rangle = 0$ for $i+1 < j$:

$$\begin{aligned}\langle\chi_{i+1}|\chi_j\rangle &= \frac{1}{\|\dots\|}\left(\left\langle\varphi_{i+1}\right| - \sum_{l=1}^{i}\langle\varphi_{i+1}|\chi_l\rangle\,\langle\chi_l|\right)|\chi_j\rangle \\ &= \frac{1}{\|\dots\|}(\langle\varphi_{i+1}|\chi_j\rangle - \sum_{l=1}^{i}\langle\varphi_{i+1}|\chi_l\rangle\,\underbrace{\langle\chi_l|\chi_j\rangle}_{=\delta_{lj}}) \\ &= \frac{1}{\|\dots\|}\left(\langle\varphi_{i+1}|\chi_j\rangle - \langle\varphi_{i+1}|\chi_j\rangle\right) = 0\,.\end{aligned}$$

Therefore $\{|\chi_i\rangle\}$ is an orthogonal basis of $\mathcal{H}$.

- Since the above eigenvectors form a complete, orthonormal set, $\langle m|n\rangle = \delta_{mn}$, we can use them as a basis in a finite dimensional Hilbert space. This allows to write a Hermitian operator A as a sum of projection operators with coefficients equal to its eigenvalues a_n,

$$A = \sum_n a_n|n\rangle\langle n|\,. \tag{3.82}$$

This relation is known as the *spectral decomposition* and allows us to define general functions of operators,

$$f(A) = \sum_n f\left(a_n\right)|n\rangle\langle n|\,, \tag{3.83}$$

provided that $f(a_n)$ is a well-defined function for all the eigenvalues a_n of A. For instance, a function

$$\ln A = \sum_n \ln\left(a_n\right)|n\rangle\langle n| \tag{3.84}$$

is only meaningful if all the eigenvalues of A are positive-definite.

Example 7: Spin 1/2 operators and Pauli matrices

In the Stern-Gerlach experiment, the spins in the $z-$ and x-axis are observables represented by the Hermitian operators S_z and S_x, both of them with eigenvalues $\pm\hbar/2$. Using the spectral decomposition (3.82) together with the previously constructed states (3.25), (3.35) and (3.36), we can easily build an explicit representation of these oper-

ators, namely

$$
\begin{aligned}
S_z &= \frac{\hbar}{2}|+\rangle\langle+| - \frac{\hbar}{2}|-\rangle\langle-| = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \equiv \frac{\hbar}{2}\sigma_3\,, \\
S_y &= \frac{\hbar}{2}|+\rangle_{yy}\langle+| - \frac{\hbar}{2}|-\rangle_{yy}\langle-| = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \equiv \frac{\hbar}{2}\sigma_2\,, \\
S_x &= \frac{\hbar}{2}|+\rangle_{xx}\langle+| - \frac{\hbar}{2}|-\rangle_{xx}\langle-| = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \equiv \frac{\hbar}{2}\sigma_1\,,
\end{aligned}
\tag{3.85}
$$

or more compactly

$$
S_i = \frac{\hbar}{2}\sigma_i\,, \tag{3.86}
$$

with the index identification $(x, y, z) = (1, 2, 3)$, and

$$
\sigma_1 \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\,, \qquad \sigma_2 \equiv \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\,, \qquad \sigma_3 \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\,, \tag{3.87}
$$

the so-called *Pauli matrices*. These are all Hermitian traceless matrices with determinant -1 and eigenvalues $+1$ and -1,

$$
\operatorname{tr}\sigma_i = 0\,, \qquad \det\sigma_i = -1\,. \tag{3.88}
$$

They also satisfy the following important and frequently used relations,

$$
\sigma_i\sigma_j = \delta_{ij}\mathbb{I} + i\epsilon_{ijk}\sigma_k\,, \qquad [\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k\,, \tag{3.89}
$$

with repeated indices summed over the values $1, 2, 3$ and ϵ_{ijk} the totally antisymmetric symbol, equal to $+1$ if $(i\,j\,k)$ is a cyclic permutation of $(1\,2\,3)$, -1 for a noncyclic permutation and zero otherwise. For example, one has $\epsilon_{231} = 1$, $\epsilon =_{213}= -1$ and $\epsilon_{112} = 0$.
The previous expressions can be used to derived an Hermitian operator

$$
S_{\vec{n}} \equiv \vec{n}\cdot\vec{S} \equiv n_x S_x + n_y S_y + n_z S_z = \frac{\hbar}{2}\vec{n}\cdot\vec{\sigma}\,, \tag{3.90}
$$

measuring spin in an arbitrary direction specified by a unit vector

$$
\vec{n} = (n_x, n_y, n_z) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)\,. \tag{3.91}
$$

More explicitly, we have

$$
\begin{aligned}
S_{\vec{n}} &= \frac{\hbar}{2}\left[n_x\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + n_y\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + n_z\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\right] \\
&= \frac{\hbar}{2}\begin{pmatrix} n_z & n_x - in_y \\ n_x + in_y & -n_z \end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix} \cos\theta & \sin\theta e^{-i\phi} \\ \sin\theta e^{i\phi} & -\cos\theta \end{pmatrix}\,.
\end{aligned}
$$

Note that, as expected, this general expression reduces to the particular cases in (3.85) for $(n_x, n_y, n_z) = (1,0,0), (0,1,0), (0,0,1)$, while sharing generically the same eigenvalues, independently of $\vec{n}$. This can be easily seen by direct computation

$$\det(S_{\vec{n}} - \lambda\mathbb{I}) = \lambda^2 - \frac{\hbar^2}{4}\left(\cos^2\theta + \sin^2\theta\right) = \lambda^2 - \frac{\hbar^2}{4} = 0\,, \quad \to \quad \lambda = \pm\frac{\hbar}{2}\,,$$

or more elegantly obtained by computing the square of the matrix and its trace

$$S_{\vec{n}}^2 = \left(\frac{\hbar}{2}\right)^2 (\vec{n}\cdot\vec{\sigma})^2 = \left(\frac{\hbar}{2}\right)^2 \mathbb{I}\,, \qquad \operatorname{tr} S_{\vec{n}} = n_i \operatorname{tr} S_i = n_i \frac{\hbar}{2}\operatorname{tr}(\sigma_i) = 0\,, \tag{3.92}$$

either by *brout force* or using the first property in (3.89) for $i = j$, namely

$$n_i\sigma_i n_j\sigma_j = n_i n_j \delta_{ij} + i \underbrace{(n_i n_j \epsilon_{ijk})}_{0 \text{ by symmetry}} \sigma_k\,, \quad \longrightarrow \quad (\vec{n}\cdot\vec{\sigma})^2 = \mathbb{I}\,. \tag{3.93}$$

Not surprisingly, especially if you understood the general discussion on the Bloch sphere in the first example of this chapter, the eigenstates $|\pm\rangle_n$ of $S_{\vec{n}}$ coincide in form with that in Eq. (3.33) and its orthogonal vector,

$$|+\rangle_n = \cos\frac{\theta}{2}|+\rangle + \sin\frac{\theta}{2}e^{i\phi}|-\rangle\,, \qquad |-\rangle_n = -\sin\frac{\theta}{2}e^{-i\phi}|+\rangle + \cos\frac{\theta}{2}|-\rangle\,. \tag{3.94}$$

as follows explicitly from the eigenvector equations $(S_{\vec{n}} - \pm(\frac{\hbar}{2})\mathbb{I})|\pm\rangle_n = 0$. For instance, for an ansatz $|+\rangle_n = c_1|+\rangle + c_2|-\rangle = (c_1\ c_2)^T$ with $|c_1|^2 + |c_2|^2 = 1$ we have

$$\frac{\hbar}{2}\begin{pmatrix} \cos\theta - 1 & \sin\theta\, e^{-i\phi} \\ \sin\theta\, e^{i\phi} & -\cos\theta - 1 \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = 0 \ \to \ c_2 = e^{i\phi}\frac{1-\cos\theta}{\sin\theta}c_1 = e^{i\phi}\frac{\sin\theta/2}{\cos\theta/2}c_1\,,$$

and

$$|c_1|^2 + |c_2|^2 = 1 \ \to \ |c_1|^2\left[1 + \left(\frac{1-\cos\theta}{\sin\theta}\right)^2\right] = \frac{|c_1|^2}{\cos^2\frac{\theta}{2}} = 1 \ \to \ \begin{cases} c_1 = \cos\frac{\theta}{2}\,, \\ c_2 = \sin\frac{\theta}{2}\,e^{i\phi} \end{cases}$$

where we have taken into account that the overall phase of the eigenstate is not observable to choose the simplest form of c_1.

A useful relation

A very useful relation for forthcoming developments is

$$\exp(i\,\vec{v}\cdot\vec{\sigma}) = \cos|\vec{v}|\mathbb{I} + i\frac{(\vec{v}\cdot\vec{\sigma})}{|\vec{v}|}\sin|\vec{v}|\,, \tag{3.95}$$

with $\vec{v}$ an arbitrary real vector with components v_j, $|\vec{v}| = \sqrt{v^j v_j}$ and $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ the spin operator, with $\sigma_x, \sigma_y, \sigma_z$ the Pauli matrices. The proof is straightforward. Expanding the above exponential in series and taking into account the first expression

in (3.88) for expressing $(\vec{v}\cdot\vec{\sigma})^2$ as

$$(\vec{v}\cdot\vec{\sigma})^2 = v_i v_j \sigma_i \sigma_j = v_i v_j (\delta_{ij}\mathbb{I} + i\epsilon_{ijk}\sigma_k) = v_i v_j \delta_{ij}\mathbb{I} = |\vec{v}|^2\mathbb{I}\,,$$

we get

$$\begin{aligned}\exp(i\,\vec{v}\cdot\vec{\sigma}) &= \sum_{k=0}^{\infty}\frac{i^k}{k!}(\vec{v}\cdot\vec{\sigma})^k = \sum_{k=0}^{\infty}\frac{i^{2k}}{(2k)!}(\vec{v}\cdot\vec{\sigma})^{2k} + \sum_{k=0}^{\infty}\frac{i^{2k+1}}{(2k+1)!}(\vec{v}\cdot\vec{\sigma})^{2k+1}\\ &= \sum_{k=0}^{\infty}\frac{(-1)^k}{(2k)!}(|\vec{v}|)^{2k}\mathbb{I} + i\frac{(\vec{v}\cdot\vec{\sigma})}{|\vec{v}|}\sum_{k=0}^{\infty}\frac{(-1)^k}{(2k+1)!}(|\vec{v}|)^{2k+1}\\ &= \cos|\vec{v}|\,\mathbb{I} + i\frac{(\vec{v}\cdot\vec{\sigma})}{|\vec{v}|}\sin|\vec{v}|\,. \end{aligned} \tag{3.96}$$

3.2.5 Position and momentum bases

Until now, we have assumed the bases of the Hilbert space $\mathcal{H}$ to be discrete, either finite or countably infinite, overlooking the fact that many systems cannot be adequately described within this framework. The simplest example is the free particle in one dimension, where the natural basis consists of the position eigenstates $\{|x\rangle\}$. Each state $|x\rangle$ represents the particle being exactly at position x, so that a measurement of the particle's position would yield x with complete certainty. Because the particle can be found at any point along the real line, the basis must include a distinct vector for every real number x. As a result, the set of basis states is uncountably infinite.

Working with infinite-dimensional vector spaces requires a more careful mathematical framework. Nonetheless, many key ideas from finite-dimensional linear algebra continue to apply by analogy. While the correspondence is not perfect, it is sufficiently strong to allow for practical calculations and to build an intuitive grasp of the behavior of quantum systems. In simple terms, the extension to the continuous requires the formal replacement of sums by integrals and Kronecker delta symbols by Dirac delta functions,

$$\sum_n \to \int ds\,, \qquad\qquad \delta_{nm} \to \delta(s-s')\,, \tag{3.97}$$

such that, in 1-dimension,

$$\begin{aligned}
|\varphi\rangle = \sum_i \varphi_i|i\rangle \qquad &\longrightarrow \qquad |\varphi\rangle = \int ds\,\varphi(s)|s\rangle\\
\langle m|n\rangle = \delta_{mn} \qquad &\longrightarrow \qquad \langle s|s'\rangle = \delta(s-s')\\
\langle\chi|\varphi\rangle = \sum_i \chi_m^*\varphi_m \qquad &\longrightarrow \qquad \langle\chi|\varphi\rangle = \int ds\,\chi^*(s)\varphi(s)\\
\varphi_m = \sum_n \delta_{mn}\varphi_n \qquad &\longrightarrow \qquad \varphi(s) = \int ds'\,\delta(s-s')\varphi(s')\\
\mathbb{I} = \sum_n |n\rangle\langle n| \qquad &\longrightarrow \qquad \mathbb{I} = \int ds\,|s\rangle\langle s|
\end{aligned}$$

Armed with this, we can establish the following links between matrix quantum mechanics (Heisenberg) and wave quantum mechanics (Schrödinger)

Position basis	Momentum basis
$X = \int dx\, x\, \|x\rangle\langle x\|$	$P = \int dp\, p\, \|p\rangle\langle p\|$
$X\|x\rangle = x\,\|x\rangle$	$P\|p\rangle = p\,\|p\rangle$
$\langle x\|x'\rangle = \delta\,(x - x')$	$\langle p\|p'\rangle = \delta\,(p - p')$
$\int dx\|x\rangle\langle x\| = \mathbb{I}$	$\int dp\,\|p\rangle\langle p\| = \mathbb{I}$
$\|\varphi\rangle = \int dx\,\varphi(x)\,\|x\rangle$	$\|\varphi\rangle = \int dp\,\tilde{\varphi}(p)\|p\rangle$

with $\varphi(x) = \langle x|\varphi\rangle$ and $\tilde{\varphi}(p) = \langle p|\varphi\rangle$ the *components* of the vector $|\varphi\rangle$ with respect to the bases $|x\rangle$ and $|p\rangle$ and the integrals integrals extending from $-\infty$ to ∞ unless otherwise stated.

Dirac's delta distribution

Don't panic. The notation $\delta(s)$ in the above expressions stands for the so-called Dirac's delta distribution, defined as

$$\int_{-\infty}^{\infty} ds\, f(s)\delta\left(s - s'\right) = f\left(s'\right) \tag{3.98}$$

in *Schwartz's distribution theory* for all sufficiently well-behaved trial functions $f(s)$ of argument s. This property is just the continuous analogue of the expression

$$\sum_i F\delta_{ij} = v_j\,, \tag{3.99}$$

with δ_{ij} the Kronecker delta you are completely familiar with. While δ_{ij} picks a vector component out of finitely many vector components, the Dirac distribution $\delta\,(s - s')$ picks a function value out of infinitely many function values, implying in particular that

$$\int_{-\infty}^{\infty} ds\,\delta\left(s - s'\right) = 1\,. \tag{3.100}$$

For a finite interval $[a, b]$, we have also

$$\int_a^b ds\, f(s)\delta\left(s - s'\right) = \begin{cases} f\left(s'\right)\,, & \text{if} \quad s \in [a, b]\,, \\ 0\,, & \text{if} \quad s \notin [a, b]\,. \end{cases} \tag{3.101}$$

Although the Dirac delta is not a function in the classical sense—since no function that is non-zero only on a set of measure zero can be integrable—it can be formally manipulated as such *within integrals*, with the following useful properties

$$\delta(s - s') = \delta(s' - s)\,, \tag{3.102}$$

$$\delta(a(s - s')) = \frac{1}{|a|}\,\delta(s - s')\,, \qquad a \in \mathbb{R},\ a \neq 0\,, \tag{3.103}$$

$$f(s)\,\delta(s - s') = f(s')\,\delta(s - s')\,, \tag{3.104}$$

$$\int_{-\infty}^{\infty} \delta(s-\eta)\,\delta(\eta-s')\,d\eta = \delta(s-s')\,, \tag{3.105}$$

$$\delta(g(s)) = \sum_i \frac{\delta(s-s_i)}{|g'(s_i)|}\,, \quad \text{where} \quad g(s_i)=0\,,\ g'(s_i)\neq 0\,. \tag{3.106}$$

In spite of the above, it might turn out convenient to represent the Dirac delta distribution as the $\epsilon \to 0^+$ limit of a smooth function $\delta(s-s',\epsilon)$ that becomes negligible away from its zero argument, but so strongly peaked there that its integral is unity, namely

$$\delta(s-s') = \lim_{\epsilon\to 0^+} \delta(s-s',\epsilon)\,, \qquad \lim_{\epsilon\to 0^+}\int_{\infty}^{\infty} ds\,\delta(s-s',\epsilon) = 1\,. \tag{3.107}$$

One could consider, for instance, a Gaussian representation

$$\delta_G(s-s',\epsilon) = \frac{1}{\epsilon\sqrt{\pi}} e^{-\frac{(s-s')^2}{\epsilon^2}}\,, \tag{3.108}$$

a Lorentzian representation

$$\delta_L(s-s',\epsilon) = \frac{1}{\pi}\frac{\epsilon}{\epsilon^2+(s-s')^2}\,, \tag{3.109}$$

or even give up continuity and define the delta function as a rectangle of width 2ϵ and height $1/2\epsilon$, ensuring that the total area is equal to 1,

$$\delta_R(s-s',\epsilon) = \begin{cases} \frac{1}{2\epsilon}\,, & \text{for} \quad |s-s'|<\epsilon\,, \\ 0\,, \text{for} & |s-s'|\geq\epsilon\,. \end{cases} \tag{3.110}$$

Another useful representation, equivalent to the Lorentzian one, appears in Fourier analysis,

$$\begin{aligned} \delta_F(s-s',\epsilon) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\, e^{ik(s-s')-\epsilon|k|} \\ &= \frac{1}{2\pi}\left(\frac{1}{\epsilon+i(s-s')}+\frac{1}{\epsilon-i(s-s')}\right) = \delta_L(s-s',\epsilon)\,. \end{aligned} \tag{3.111}$$

Finally, we highlight a sine-square representation

$$\delta_S(s-s',\epsilon) = \frac{1}{\pi}\frac{\epsilon\sin^2((s-s')/\epsilon)}{(s-s')^2}\,, \tag{3.112}$$

that will become useful when studying time-dependent perturbation theory in Chapter 8.

Note that the norm of position or momentum states is infinite,

$$\langle x|x\rangle = \delta(0) = \infty\,, \qquad \langle p|p\rangle = \delta(0) = \infty\,, \tag{3.113}$$

meaning that they are not square-integrable functions in $\mathbb{R}^3$, $L^2\left(\mathbb{R}^3\right)$, and therefore not suitable states for describing particles. This should not be a surprise, since the position of a particle cannot measured with infinite precision. Normalizable states can be however constructed using superpositions of a continuous set of states,

$$|\varphi\rangle = \int_{x-\Delta/2}^{x+\Delta/2} dx\, |x\rangle\langle x|\varphi\rangle = \int_{x-\Delta/2}^{x+\Delta/2} dx\, \varphi(x)|x\rangle\,, \tag{3.114}$$

with Δ a narrow but non-vanishing range around a position our detector cannot resolve. In this language, the wave function $\varphi(x)$ you have been using to describe physical states is identified with the components of an abstract vector $|\varphi\rangle$ in an infinite-dimensional space in which we have *chosen* to introduce a basis labeled by the values of the position. Answering one of the questions we formulated in the introduction, we notice that the same abstract vector could be alternatively described by its components along the direction associated to a definite value of the momentum ($\langle p|\varphi\rangle$) or other any other direction connected with alternative observables. From a mathematical perspective, wave mechanics is just one of the possible representations of quantum mechanics and there is nothing particularly exceptional about it.

To determine the momentum operator P, let us can consider the translation operator $T : \mathcal{H} \to \mathcal{H}$ leading to the *infinitesimal* displacement from x to $x + \delta\, x$, namely

$$\left.\begin{array}{l} T(x+\delta x) = T(x) + \frac{dT}{dx}\delta x \\ T(x+\delta x) = T(x)T(\delta x) = T(x)(\mathbb{I} - i\delta x K) = T(x) - i\delta x T(x) K \end{array}\right\} \quad \frac{dT}{dx} = -iTK\,, \tag{3.115}$$

where we have taken into account that the continuous set of translations can be composed or multiplied,

$$T\left(\vec{x}_1\right) T\left(\vec{x}_2\right) = T\left(\vec{x}_1 + \vec{x}_2\right)\,, \tag{3.116}$$

and introduced the so called generator of translations K through

$$T(\delta x) \equiv \mathbb{I} - i\delta x K + \mathcal{O}[(\delta x)^2]\,. \tag{3.117}$$

With the boundary condition $T(0) = 0$, the above first-order differential equation admits a solution

$$T(x) = e^{-iKx}\,. \tag{3.118}$$

The physical interpretation of K follows naturally from the commutation relation of the position operator X and $T(\delta X)$,

$$[X, T(\delta x)]|x\rangle = \delta x|x+\delta x\rangle = \delta x T(\delta x)|x\rangle \quad \longrightarrow \quad [X, T(\delta x)] = \delta x T(\delta x)\,, \tag{3.119}$$

where we have taken into account that

$$\begin{aligned} XT(\delta x)|x\rangle &= X|x+\delta x\rangle = (x+\delta x)|x+\delta x\rangle\,, \\ T(\delta x)X|x\rangle &= T(\delta x)x|x\rangle = x|x+\delta x\rangle\,. \end{aligned} \tag{3.120}$$

Combining this result with the definition (3.117), we have

$$\begin{aligned} [X, T(\delta x)] &= [X, \mathbb{I} - i\delta x K] = -i\delta x[X, K]\,, \\ \delta x T(\delta x) &= \delta x(\mathbb{I} - i\delta x K) \simeq \delta x\, \mathbb{I} + \mathcal{O}[(\delta x)^2], \end{aligned} \tag{3.121}$$

and therefore

$$[X, \hbar K] = i\hbar \mathbb{I}\,, \tag{3.122}$$

allowing us to identify the generator of translations in one dimension with the linear momentum $P = \hbar K$! The action of this operator on kets and bras in the position representation is given by

$$\left.\begin{aligned} &T(\delta x)|x\rangle = |x+\delta x\rangle \\ &T(\delta x)|x\rangle = (\mathbb{I} - \tfrac{i}{\hbar}P\delta x)|x\rangle = |x\rangle - \tfrac{i}{\hbar}\delta x P|x\rangle \end{aligned}\right\} \; P|x\rangle = i\hbar\frac{|x+\delta x\rangle - |x\rangle}{\delta x}\,, \tag{3.123}$$

and

$$\langle x|P^\dagger = \langle x|P = -i\hbar\frac{\langle x+\delta x| - \langle x|}{\delta x}\,, \tag{3.124}$$

since P is self-adjoint,

$$\left.\begin{aligned} &T^{-1}(\delta x) = \mathbb{I} + \tfrac{i}{\hbar}\delta x P \\ &T^\dagger(\delta x) = \mathbb{I} + \tfrac{i}{\hbar}\delta x P^\dagger \end{aligned}\right\} \; T^\dagger = T^{-1}\,. \tag{3.125}$$

These two expressions allow us to recover the form of the momentum operator you discovered in *Quantum Physics I*,

$$\begin{aligned} \langle x|P|\varphi\rangle &= \langle x|P\int dx'|x'\rangle\langle x'|\varphi\rangle = -i\hbar\frac{\langle x+\delta x| - \langle x|}{\delta x}\int dx'|x'\rangle\varphi(x') \\ &= -i\hbar\frac{1}{\delta x}\int dx'[\delta(x+\delta x - x') - \delta(x-x')]\varphi\left(x'\right) = -i\hbar\frac{\varphi(x+\delta x)-\varphi(x)}{\delta x} \\ &= -i\hbar\frac{d\varphi(x)}{dx}\,. \end{aligned} \tag{3.126}$$

namely

$$P = -i\hbar\frac{d}{dx}\,. \tag{3.127}$$

Hermitian versus self-adjoint

In finite-dimensional Hilbert spaces, the different operators can be represented by matrices, which can generically act on every vector in the Hilbert space. In infinite-dimensional Hilbert spaces, however, an operator A may not be able to act on the entire Hilbert space. This leads to some technical distinctions between the concepts of Hermitian and self-adjoint operators, with self-adjointness being somehow a stronger condition that requires not only that $A = A^\dagger$, but also that A and $A^\dagger$ share the same domain, $D(A^\dagger) = D(A)$. Usually the question of domains $D(A) \subset D(A^\dagger)$ appears in problems with non-trivial boundary conditions. Consider for instance the momentum operator of a particle confined to a one-dimensional interval $0 \leq x \leq L$. The corresponding domain is given by the set of twice-differentiable wave functions that vanish at the boundaries of the box, namely

$$D(P) = \left\{\phi(x) \in \mathcal{L}^2([0,L]) \mid \phi(0) = \phi(L) = 0\right\}\,.$$

Since the difference between the terms $\langle\varphi|P\phi\rangle$ and $\langle P\varphi|\phi\rangle$ vanishes due to the condi-

tion $\phi(0) = \phi(L) = 0$,

$$\langle\varphi|P|\phi\rangle - \langle P\varphi|\phi\rangle = \int_0^L dx \left[\langle x|\varphi\rangle \left(-i\hbar\frac{d}{dx}\langle x|\phi\rangle\right) - \left(i\hbar\frac{d}{dx}\langle x|\varphi\rangle\right)\langle x|\phi\rangle\right]$$
$$= -i\hbar\int_0^L dx\,\frac{d}{dx}\left[\langle x|\varphi\rangle\langle x|\phi\rangle\right] = -i\hbar\left[\bar\varphi(L)\phi(L) - \bar\varphi(0)\phi(0)\right] = 0\,,$$

the operator P is Hermitian within D. However, the operator is not self-adjoint (without extensions) because the domain of its adjoint, $P^\dagger$, is larger than that of P. Indeed, while having the same differential form as P, the adjoint operator $P^\dagger$ acts on the space of all square-integrable functions, without any restriction on their values at the boundaries,

$$D(P^\dagger) = \left\{\varphi(x) \in \mathcal{L}^2([0,L]) \mid \text{no boundary conditions on } \varphi(0) \text{ and } \varphi(L)\right\}\,.$$

To make P self-adjoint, you would need to find boundary conditions that make $D(P) = D(P^\dagger)$. This leads to a family of self-adjoint extensions, where the boundary conditions (like $\phi(L) = e^{i\theta}\phi(0)$) are chosen to ensure this equality.

Using Eq. (3.127) and the states in the position and momentum representation, we can determine the form of the overlaps $\langle x|p\rangle$,

$$-i\hbar\frac{d}{dx}\langle x|p\rangle = p\langle x|p\rangle \quad\longrightarrow\quad \langle x|p\rangle = Ne^{\frac{\mathrm{i}}{\hbar}px}\,, \tag{3.128}$$

with N a normalization constant that can be determined as

$$\left.\begin{aligned}\langle p|p'\rangle &= \delta(p'-p) = \tfrac{1}{\hbar}\delta((p'-p)/\hbar) = \tfrac{1}{2\pi\hbar}\textstyle\int dx\, e^{\frac{\mathrm{i}}{\hbar}(p'-p)x}\\ \langle p|p'\rangle &= \textstyle\int dx\,\langle p|x\rangle\langle x|p'\rangle = |N|^2\int dx\, e^{\frac{\mathrm{i}}{\hbar}(p'-p)x}\end{aligned}\right\}\quad N = \frac{1}{(2\pi\hbar)^{1/2}}\,, \tag{3.129}$$

where in the first expression we have used the Dirac delta property (3.103) to pull out a $1/\hbar$ prefactor and the Fourier representation (3.111) in the $\epsilon \to 0^+$ limit. Combining Eqs. (3.128) and (3.129) we have therefore

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}}e^{ipx/\hbar}\,, \qquad \langle p|x\rangle = \frac{1}{\sqrt{2\pi\hbar}}e^{-ipx/\hbar}\,. \tag{3.130}$$

These overlaps allows us to change between position and momentum wave functions

$$\tilde\varphi(p) = \langle p|\varphi\rangle = \int dx\langle p|x\rangle\langle x|\varphi\rangle = \int dx\langle p|x\rangle\varphi(x) = \frac{1}{\sqrt{2\pi\hbar}}\int dx\, e^{-ipx/\hbar}\varphi(x)\,,$$

$$\tilde\varphi(p) = \frac{1}{\sqrt{2\pi\hbar}}\int dx\, e^{-ipx/\hbar}\varphi(x)\,, \qquad \varphi(x) = \frac{1}{\sqrt{2\pi\hbar}}\int dp\, e^{ipx/\hbar}\tilde\varphi(p)\,. \tag{3.131}$$

We have therefore $\tilde\varphi(p)$ is the Fourier transform of $\varphi(x)$. Analogously, $\varphi(x)$ is the inverse Fourier transform of $\tilde\varphi(p)$, with the invariance under the transformations

$$\varphi(x) \leftrightarrow \tilde\varphi(p)\,, \qquad x \leftrightarrow p\,, \qquad i \leftrightarrow -i\,, \tag{3.132}$$

leading to the *Parcival theorem*

$$\int_{-\infty}^{\infty} dx\, \varphi^*(x)\varphi(x) = \int_{-\infty}^{\infty} dp\, \tilde{\varphi}^*(p)\tilde{\varphi}(p)\,. \tag{3.133}$$

Note that the transformations (3.131) are well-defined even in the case of a well-localized particle with $\varphi(x) = \langle x|x_0\rangle = \delta(x - x_0)$, for which

$$\tilde{\varphi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int dx\, e^{-ipx/\hbar}\delta(x - x_0) = \frac{1}{\sqrt{2\pi\hbar}} e^{-\frac{i}{\hbar}px_0}\,. \tag{3.134}$$

In this case, however, the associated momentum distribution is constant

$$|\tilde{\varphi}(p)|^2 = \frac{1}{2\pi\hbar}\,, \tag{3.135}$$

and all momenta values equally probable. This is just a manifestation of the uncertainty principle, to be described in the next section.

3.3 Measurements

Postulate III

When a physical quantity (observable) is measured in a quantum system, the system's state collapses into one of the eigenstates of the corresponding operator. The probability of obtaining a particular outcome (eigenvalue) is determined by the square of the inner product between the system's state vector and the eigenstate. After the measurement, the system's state is given by the eigenstate corresponding to the observed measurement outcome, so, if the measurement is immediately repeated, the same outcome is obtained with probability one.

Prior to this statement, our journey has primarily focused on conventional linear algebra and notation. The pivotal moment that elevates quantum mechanics to the realm of physics lies in establishing the correlation between squared overlaps and probabilities.

According to the above postulate, if a system is prepared in a general superposition state $|\varphi\rangle$, the probability of obtaining an eigenvalue corresponding to the particular eigenstate $|m\rangle$ is given by

$$\text{Prob}\left(|\varphi\rangle \to |m\rangle\right) = \frac{|\langle m|\varphi\rangle|^2}{\langle\varphi|\varphi\rangle\,\langle m|m\rangle} = \frac{\langle\varphi|P_m|\varphi\rangle}{\langle\varphi|\varphi\rangle\,\langle m|m\rangle}\,, \tag{3.136}$$

where $P_m = |m\rangle\langle m|$ stands for the projector operator on the state $|m\rangle$, and we have allowed for arbitrary normalizations for the states $|\varphi\rangle$ and $|m\rangle$. Note that the potential outcomes of this formula have all the fundamental properties one should expect for probabilities. First, they are all positive definite

$$\text{Prob}\left(|\varphi\rangle \to |m\rangle\right) \geq 0\,, \tag{3.137}$$

and remain the same if we multiply $|\varphi\rangle$ and/or $|\varphi\rangle$ by complex numbers. Second, the sum to all possible outcomes equals to one, namely

$$\sum_m \text{Prob}\left(|\varphi\rangle \to |m\rangle\right) = \sum_m \frac{|c_m|^2 \langle m|m\rangle}{\langle\varphi|\varphi\rangle} = 1\,, \tag{3.138}$$

where in the last step we have used

$$\langle\varphi|\varphi\rangle = \sum_{n,m} c_n c_m^* \langle m|n\rangle = \sum_m |c_m|^2 \langle m|m\rangle\ . \tag{3.139}$$

Of course, if our initial and final states are properly normalized, $\langle\varphi|\varphi\rangle = \langle m|m\rangle = 1$, the above probability depends only on the probability amplitude

$$\text{Prob}\left(|\varphi\rangle \to |m\rangle\right) = |\langle m|\varphi\rangle|^2 = \langle\varphi|P_m|\varphi\rangle \qquad \text{(Born's rule)}\,, \tag{3.140}$$

being still invariant if $|\varphi\rangle$ and/or $|\varphi\rangle$ are multiplied by a phase.

Example 8: Revisiting the Stern-Gerlach experiment

Stern-Gerlach Experiment 1

$|\varphi\rangle = |+\rangle$

$$\begin{aligned} &\text{Prob}(|+\rangle \to |+\rangle) = |\langle+|+\rangle|^2 = 1, \quad \to \quad \text{Final state} \quad |\varphi'\rangle = |+\rangle\ , \\ &\text{Prob}(|+\rangle \to |-\rangle) = |\langle-|+\rangle|^2 = 0 \quad \to \quad |-\rangle \text{ never happens}\,. \end{aligned}$$

Stern-Gerlach Experiment 2

$|\varphi\rangle = |+\rangle$

$$\begin{aligned} &\text{Prob}(|+\rangle \to |+\rangle_x) = |_x\langle+|+\rangle|^2 = \frac{1}{2}\,, \quad \to \quad \text{Final state} \quad |\varphi'\rangle = |+\rangle_x\,, \\ &\text{Prob}(|+\rangle \to |-\rangle_x) = |_x\langle-|+\rangle|^2 = \frac{1}{2}\,, \quad \to \quad \text{Final state} \quad |\varphi'\rangle = |-\rangle_x\,. \end{aligned}$$

Stern-Gerlach Experiment 3

$|\varphi\rangle = |+\rangle_x$ (after filtering one half of the atoms)

$$\begin{aligned} &\text{Prob}(|+\rangle_x \to |+\rangle) = |\langle+|+\rangle_x|^2 = \frac{1}{2}, \quad \to \quad \text{Final state} \quad |\varphi'\rangle = |+\rangle\ , \\ &\text{Prob}(|+\rangle_x \to |-\rangle) = |\langle-|+\rangle_x|^2 = \frac{1}{2}, \quad \to \quad \text{Final state} \quad |\varphi'\rangle = |-\rangle\,. \end{aligned}$$

Projection along $\vec{n}$ (cf. (3.94))

$|\varphi\rangle = |+\rangle$

$$\text{Prob}(|+\rangle \to |+\rangle_n) = |\langle +|+\rangle_n|^2 = \left|\cos\frac{\theta}{2}\langle +|+\rangle + \sin\frac{\theta}{2}e^{i\phi}\langle +|-\rangle\right|^2 = \cos^2\frac{\theta}{2}\,,$$
$$\text{Prob}(|+\rangle \to |-\rangle_n) = |\langle +|+\rangle_n|^2 = \left|-\sin\frac{\theta}{2}e^{i\phi}\langle +|+\rangle + \cos\frac{\theta}{2}\langle +|-\rangle\right|^2 = \sin^2\frac{\theta}{2}\,.$$

Note that if there exist several states with the same eigenvalue (degenerate states), one must sum over them when computing probabilities. Also, immediately after the measurement, the state is the superposition state associated to all degenerate eigenvalues, namely

$$|\varphi\rangle \to \frac{P_m|\varphi\rangle}{\langle\varphi\,|P_m|\,\varphi\rangle^{1/2}}\,, \tag{3.141}$$

with P_m the projection operator on the corresponding subspace

$$P_m = \sum_{r=1}^{g(m)} |m,r\rangle\langle m,r|, \tag{3.142}$$

and $g(m)$ the degeneracy of the eigenvalue m. Note that the denominator in (3.141) ensures the proper normalization of the resulting state, since, as any projector operator, $P_m^2 = P_m$.

Example 9: Probabilities and degenerate eigenvalues

Consider a state

$$|\varphi\rangle = \frac{1}{\sqrt{6}}(2|1\rangle + |2\rangle + |3\rangle)$$

in a quantum mechanical system with three orthonormal states $|1\rangle, |2\rangle, |3\rangle$ on which a observable A has 2 degenerate eigenvalues $a_1 = a_2 = 0$ and a non-degenerate eigenvalue $a_3 = 1$. The probabilities of measuring the eigenvalues 1 and 0 are then given by

$$\text{Prob}(|\varphi\rangle \to |3\rangle) = \left|\frac{1}{\sqrt{6}}\right|^2 = \frac{1}{6}\,, \quad \to \quad |\varphi'\rangle = |3\rangle\,,$$
$$\text{Prob}(|\varphi\rangle \to |1\rangle, |2\rangle) = \left|\frac{2}{\sqrt{6}}\right|^2 + \left|\frac{1}{\sqrt{6}}\right|^2 = \frac{5}{6}\,, \quad \to \quad |\varphi'\rangle = \frac{1}{\sqrt{5}}(2|1\rangle + |2\rangle)\,,$$

where we have indicated also the final states.

3.3.1 Compatible observables

Imagine that you want to measure two observables associated with two Hermitian operators A and B. As we have seen in Section 3.2.4, the eigenstates of either A and B can be used to

construct an orthonormal basis, namely

$$\begin{array}{l} A|a_i\rangle = a_i|a_i\rangle\,, \\ B|b_i\rangle = b_i|b_i\rangle\,, \end{array} \quad \longrightarrow \quad |\varphi\rangle = \sum_i \langle a_i|\varphi\rangle\,|a_i\rangle = \sum_i \langle b_i|\varphi\rangle\,|b_i\rangle\ . \tag{3.143}$$

If we start measuring the observable corresponding to A, and immediately after the one corresponding to B, we get

$$\begin{array}{lllll} & \underline{\text{Measure A}} & & \underline{\text{Measure B}} & \underline{\text{Final output}} \\ |\varphi\rangle \longrightarrow & a_i \text{ with prob. } |\langle a_i|\varphi\rangle|^2 & \longrightarrow & b_j \text{ with prob. } |\langle b_j|a_i\rangle|^2 & \longrightarrow \ (a_i, b_j) \\ & \text{System in state } |a_i\rangle & & \text{System in state } |b_i\rangle & |\langle a_i|\varphi\rangle|^2\,,\ |\langle b_j|a_i\rangle|^2\ . \end{array}$$

On the other hand, if we reverse the order of measurements, we rather obtain

$$\begin{array}{lllll} & \underline{\text{Measure B}} & & \underline{\text{Measure A}} & \underline{\text{Final output}} \\ |\varphi\rangle \longrightarrow & b_j \text{ with prob. } |\langle b_j|\varphi\rangle|^2 & \longrightarrow & a_i \text{ with prob. } |\langle a_i|b_j\rangle|^2 & \longrightarrow \ (a_i, b_j) \\ & \text{System in state } |b_j\rangle & & \text{System in state } |a_i\rangle & |\langle b_j|\varphi\rangle|^2\,,\ |\langle a_i|b_j\rangle|^2\ . \end{array}$$

The results obtained through these two procedures will generally differ, unless the second measurement leaves the state resulting from the former unaffected. This can only happen if there exists a *complete set of simultaneous eigenstates* of A and B, namely

$$A|u_j\rangle = a_j|u_j\rangle\,, \qquad B|u_j\rangle = b_j|u_j\rangle\,. \tag{3.144}$$

This condition is equivalent to

$$[A, B] = 0\,, \qquad \text{(Compatible observables)} \tag{3.145}$$

as follows easily from considering the matrix elements

$$\langle u_j|[A, B]|u_k\rangle = (a_k b_k - b_k a_k)\,\langle u_j|u_k\rangle = 0\,. \tag{3.146}$$

On the contrary, if $[A, B] \neq 0$, there will be at least one eigenstate $|a_j\rangle$ of A that is not an eigenstate of B, such that the above argument holds and the final outputs differ. In other words, it is impossible to measure simultaneously two observables represented by non-commuting operators because they cannot be diagonalized in the same basis, they are *incompatible*.

The largest set of mutually commuting observables is referred to as a *complete set of compatible observables (CSCO)*. When this set is identified, it becomes possible to consistently narrow down the quantum state of a system to a unique and well-defined state. This procedure operates as follows: Let us consider a complete set of mutually commuting observables $A, B, C, \ldots$. To start, we perform a measurement of the observable A, yielding outcome a_i. If a_i is not degenerate, the state collapses to $|a_i\rangle$. If a_i turns out to be degenerate, we proceed to measure the observable B, obtaining a result b_j. If b_j is non-degenerate, the state becomes $|a_i b_j\rangle$. If, on the contrary, b_j is also degenerate, we move on to measure the observable C, continuing this process. The sequence continues until we ultimately arrive at a state $|a_i b_j c_l \cdots\rangle$ completely specified by the eigenvalues of all the mutually compatible observables measured along the way.

3.3.2 Expectation values

Given the complete, orthonormal set of eigenstates $\{|n\rangle\}$ of some Hermitian operator A with eigenvalues a_n, $A|n\rangle = a_n|n\rangle$, the *expectation value* in a quantum state $|\varphi\rangle$

$$\langle A\rangle_\varphi = \langle\varphi|A|\varphi\rangle \tag{3.147}$$

is just the sum of all the eigenvalues associated to the states $|n\rangle$, weighted by the probability that $|\varphi\rangle$ agrees with $|n\rangle$,

$$\langle A\rangle_\varphi = \langle\varphi|A|\varphi\rangle = \sum_{m,n} c_m^* c_n \langle m|A|n\rangle = \sum_n a_n \left|c_n\right|^2 , \tag{3.148}$$

namely

$$\langle A\rangle_\varphi = \sum_n a_n \left|c_n\right|^2 . \tag{3.149}$$

Similarly the expectation values

$$\langle A^p\rangle_\varphi = \sum_n a_n^p |c_n|^2 \tag{3.150}$$

with integer p are the moments of the probability distribution associated with measuring the observable associated with A in the state $|\varphi\rangle$.

Example 10: Expectation value of Pauli matrices and the Bloch sphere

The expectation values of the Pauli matrices for a general state

$$|\varphi\rangle = \cos\frac{\theta}{2}|+\rangle + \sin\frac{\theta}{2}e^{\mathrm{i}\phi}|-\rangle ,$$

take the form

$$\begin{aligned}\langle\varphi\,|\sigma_x|\,\varphi\rangle &= \langle\varphi|\sigma_x|\varphi\rangle = \cos\theta ,\\ \langle\varphi\,|\sigma_y|\,\varphi\rangle &= \langle\varphi|\sigma_y|\varphi\rangle = \sin\theta\sin\phi,\\ \langle\varphi\,|\sigma_z|\,\varphi\rangle &= \langle\varphi|\sigma_z|\varphi\rangle = \cos\theta ,\end{aligned}$$

such that

$$\langle\sigma_x\rangle^2 + \langle\sigma_y\rangle^2 + \langle\sigma_z\rangle^2 = \cos^2\phi + \sin^2\phi\sin^2\theta + \cos^2\theta = 1.$$

This shows that the expectation values of the Pauli matrices form a unit vector in Cartesian coordinates, which corresponds to the Bloch vector for the given quantum state.

3.3.3 Generalized uncertainty principle

Let us define the uncertainty or standard deviation of A from its mean $\langle A\rangle_\varphi$ in a state $|\varphi\rangle$,

$$\Delta_\varphi A \equiv \sqrt{\langle\varphi|\,(A-\langle A\rangle_\varphi\mathbb{I})^2\,|\varphi\rangle}\,, \tag{3.151}$$

with the identity operator operator made here explicit for the sake of clarity, but omitted in what follows. Note that since

$$\langle\varphi|\,(A-\langle A\rangle_\varphi)^2\,|\varphi\rangle = \langle\varphi|\,(A-\langle A\rangle_\varphi)^\dagger\,(A-\langle A\rangle_\varphi)\,|\varphi\rangle = |\,(A-\langle A\rangle_\varphi)\,|\varphi\rangle|^2 \geq 0\,, \tag{3.152}$$

the uncertainty $\Delta_\varphi A = 0$ iff $|\varphi\rangle$ is an eigenstate of A, namely for $A|\varphi\rangle = \lambda|\varphi\rangle$ we have

$$\left.\begin{array}{l}(A-\langle A\rangle_\varphi)|\varphi\rangle = 0\\ \langle\varphi|A|\varphi\rangle = \lambda\langle\varphi|\varphi\rangle\end{array}\right\} A|\varphi\rangle = \langle A\rangle|\varphi\rangle = \lambda|\varphi\rangle\,. \tag{3.153}$$

This means that, unless we know that our state is in an eigenstate of A before carrying out a measurement, we will never be sure of the value we will obtain when measuring an observable quantity! Let us also emphasize that the uncertainty $\Delta_\varphi A$ is not related at all to the experimental limitations of a given experiment or the errors that could result from performing it. In particular, the resolution δA of a given device can be smaller than $\Delta_\varphi A$, but even in that case, the spread of the results will never be smaller than $\Delta_\varphi A$. In other words, the condition $\delta A \ll \Delta_\varphi A$ only allows to accurately determine the uncertainty $\Delta_\varphi A$.

Let us consider two Hermitian operators A and B and define the states

$$|\varphi_A\rangle \equiv A|\varphi\rangle - \langle A\rangle_\varphi|\varphi\rangle\,, \qquad |\varphi_B\rangle \equiv B|\varphi\rangle - \langle B\rangle_\varphi|\varphi\rangle\,, \tag{3.154}$$

such that (cf. Eq. (3.152))

$$(\Delta_\varphi A)^2 = \|\varphi_A\|^2\,, \qquad (\Delta_\varphi B)^2 = \|\varphi_B\|^2\,. \tag{3.155}$$

From the Cauchy-Schwarz inequality (3.16) it follows that

$$(\Delta_\varphi A)^2\,(\Delta_\varphi B)^2 = \|\varphi_A\|^2\|\varphi_B\|^2 \geq |\langle\varphi_A|\varphi_B\rangle|^2 \geq |\,\mathrm{Im}\,\langle\varphi_A|\varphi_B\rangle\,|^2\,, \tag{3.156}$$

which taking into account

$$\langle\varphi_A|\varphi_B\rangle = \langle\varphi|\,(A-\langle A\rangle_\varphi)\,(B-\langle B\rangle_\varphi)\,|\varphi\rangle = \langle\varphi|AB-\langle A\rangle_\varphi\langle B\rangle_\varphi|\varphi\rangle\,, \tag{3.157}$$

$$\langle\varphi_A|\varphi_B\rangle^* = \langle\varphi|AB-\langle A\rangle_\varphi\langle B\rangle_\varphi|\varphi\rangle^* = \langle\varphi|BA-\langle A\rangle_\varphi\langle B\rangle_\varphi|\varphi\rangle\,, \tag{3.158}$$

and therefore

$$\mathrm{Im}\,\langle\varphi_A|\varphi_B\rangle = \frac{1}{2i}\langle\varphi|[A,B]|\varphi\rangle\,, \tag{3.159}$$

can be written as

$$(\Delta_\varphi A)^2\,(\Delta_\varphi B)^2 \geq \left|\frac{1}{2i}\langle[A,B]\rangle_\varphi\right|^2\,. \tag{3.160}$$

This inequality constitutes the generalization of the *Heisenberg uncertainty relation* for position and momentum operators,

$$(\Delta_\varphi X)(\Delta_\varphi P) \geq \frac{\hbar}{2}\,, \tag{3.161}$$

following from the commutation relation $[X, P] = i\hbar$. In particular, Eq. (3.160) states the impossibility of finding states having simultaneously definite values of A or B if $[A, B] \neq 0$. Note, however, that the lower limit in (3.160) relies on the state $|\varphi\rangle$, so there could be cases in which the expectation value $\langle [A, B]\rangle_\varphi$ vanishes even if the observables A and B have a non-zero commutator $[A, B] \neq 0$. In those specific cases, it is a priori possible to measure simultaneously A and B with arbitrary accuracy. This happens, for instance, when considering the commutator expectation value $\langle \varphi\,|[L_x, L_y]|\,\varphi\rangle$ of two orbital angular momentum components in a spherically symmetric state $|\varphi\rangle$, which is generically vanishing even if $[L_x, L_y] = i\hbar L_z \neq 0$.

Example 11: Spin 1/2 algebra and uncertainty relations

The observables S_x and S_z are incompatible because

$$\begin{aligned}
[S_x, S_z] &= \frac{\hbar^2}{4}\left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right) \\
&= \frac{\hbar^2}{4}\left(\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right) = \frac{\hbar^2}{4}\begin{pmatrix} 0 & -2 \\ 2 & 0 \end{pmatrix} \\
&= -i\hbar\,\frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = -i\hbar S_y \neq 0\,.
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
[S_x, S_y] &= \frac{\hbar^2}{4}\left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} - \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right) \\
&= \frac{\hbar^2}{4}\left(\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} - \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}\right) = \frac{\hbar^2}{4}\begin{pmatrix} 2i & 0 \\ 0 & -2i \end{pmatrix} \\
&= i\hbar\frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = i\hbar S_z \neq 0\,.
\end{aligned}$$

It is indeed not difficult to convince oneself that the three hermitian operators S_i with $i = 1, 2, 3$ (x, y, z) satisfy the algebra

$$[S_i, S_j] = i\hbar\epsilon_{ijk}S_k\,, \tag{3.162}$$

with repeated indices summed over the values $1, 2, 3$ and ϵ_{ijk} the totally antisymmetric symbol with $\epsilon_{123} = +1$ or equivalently $\vec{S}\times\vec{S} = i\hbar\vec{S}$ if you don't feel comfy with indices yet. As for the angular momentum, this imply that the components of spin operator

$$\vec{S} = \begin{pmatrix} S_x \\ S_y \\ S_z \end{pmatrix} \tag{3.163}$$

along different axes cannot be simultaneously measured, meaning that they are incompatible observables. This is reflected in a generalized uncertainty relation

$$(\Delta S_x)^2(\Delta S_y)^2 \geq \frac{1}{4}\langle i[S_x, S_y]\rangle^2 = \frac{\hbar^2}{4}\langle S_z\rangle^2 . \tag{3.164}$$

Each spin component commutes, however, with the Casimir operator

$$S^2 = \frac{\hbar^2}{4}\left(\sigma_x^2 + \sigma_y^2 + \sigma_z^2\right) = \frac{3}{4}\hbar^2\mathbb{I}, \tag{3.165}$$

namely

$$\left[S^2, S_i\right] = 0, \tag{3.166}$$

meaning that it is always possible to measure simultaneously the absolute value of the spin and one of its components.
The operators S_x and S_y can be combined into the so-called raising and lowering operators

$$S_\pm = S_x \pm iS_y \tag{3.167}$$

satisfying the commutation relations

$$\left[S^2, S_\pm\right] = 0\,, \qquad [S_z, S_\pm] = \pm\hbar S_x\,, \qquad [S_+, S_-] = 2\hbar S_z\,, \tag{3.168}$$

The name given to these operators becomes obvious when considering the action of S^2 and S_z on the states $|\pm\rangle$

$$S_+\begin{pmatrix}0\\1\end{pmatrix} = \hbar\begin{pmatrix}0&1\\0&0\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} = \hbar\begin{pmatrix}1\\0\end{pmatrix}, \tag{3.169}$$

$$S_-\begin{pmatrix}1\\0\end{pmatrix} = \hbar\begin{pmatrix}0&0\\1&0\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} = \hbar\begin{pmatrix}0\\1\end{pmatrix}. \tag{3.170}$$

3.4 Temporal evolution

Postulate IV

The evolution of a quantum system is governed by a linear, unitary operator acting on the state vector. The evolution of the state vector over time is determined by the Schrödinger equation.

Symmetries are fundamental in physics. Imagine we take a normalized state $|\varphi, t_0\rangle$ at time t_0. Whatever the time evolution of this state turns out to be, it should end up in another

vector $|\varphi, t\rangle$ within the Hilbert space and with the *same length*. As we will show in what follows, this is precisely what a unitary operator does.

3.4.1 Unitary operators

An operator U is said to be unitary if it is the inverse of its own adjoint, namely

$$UU^\dagger = U^\dagger U = 1\,, \qquad \text{or} \qquad U^\dagger = U^{-1}\,. \tag{3.171}$$

Given such an operator, we can define a *similarity transformation* on states and operators,

$$|\varphi\rangle \mapsto |\varphi'\rangle = U|\varphi\rangle\,, \qquad \langle\varphi| \mapsto \langle\varphi'| = \langle\varphi|U^\dagger\,, \qquad A \mapsto A' = UAU^\dagger, \tag{3.172}$$

such that the scalar product and all physical properties remain unchanged,

$$\langle\chi'|\varphi'\rangle = \langle\chi|U^\dagger U|\varphi\rangle = \langle\chi|\varphi\rangle\,, \tag{3.173}$$

$$\langle\chi'|A'|\varphi'\rangle = \langle\chi\left|U^\dagger UAU^\dagger U\right|\varphi\rangle = \langle\chi|A|\varphi\rangle\,. \tag{3.174}$$

$$[A', B'] = [A, B]\,, \tag{3.175}$$

$$\vdots$$

including eigenvalues.

Unitary operators have some crucial properties:

- Contrary to Hermitian operators, the eigenvalues u_n of unitary operators are complex numbers of unit modulus. The eigenvectors $|n\rangle$ and $|m\rangle$ associated to distinct eigenvalues u_n and u_m are nonetheless still orthogonal to each other, i.e. $\langle m|n\rangle = \delta_{nm}$. Indeed, for

  $$U|n\rangle = u_n|n\rangle\,, \tag{3.176}$$

  we have

  $$\langle m|n\rangle = \langle m|U^\dagger U|n\rangle = u_m^* u_n \langle m|n\rangle \quad \longrightarrow \quad (1 - u_m^* u_n)\langle m|n\rangle = 0\,, \tag{3.177}$$

  and hence $|u_n| = 1$ if $m = n$ ($\langle n|n\rangle \neq 0$) and $\langle m|n\rangle = 0$ for $u_n \neq u_m$.

- Since the eigenvectors of a given unitary operator U form a complete, orthonormal set, $\langle m|n\rangle = \delta_{mn}$ we can use them as a basis.

- Using unitary operators, we can perform changes $|n\rangle \to U\,|n\rangle$ among orthonormal bases. The transformed states $U|n\rangle$ can be viewed as the columns of a matrix, which are both normalized and mutually orthogonal. Additionally, since the inverse of a unitary matrix is given by its conjugate transpose, denoted as $U^{-1} = U^\dagger$, the rows of this matrix are also normalized and orthogonal to each other. In general, hermitian matrices can be diagonalized using unitary matrices.

- A unitary operator U can be written as the exponential of a Hermitian operator A called the *generator* of U,
$$U = e^{iA}\,. \tag{3.178}$$
Indeed, as follows from the Taylor expansion of the exponential,
$$e^{iA} = \sum_{i=0}^{\infty} \frac{1}{n!}(iA)^n\,, \tag{3.179}$$
the properties $A = A^\dagger$ and $(A^n)^\dagger = (A^\dagger)^n$,
$$(e^{iA})^\dagger = \sum_{i=0}^{\infty} \frac{1}{n!}((iA)^n)^\dagger = \sum_{i=0}^{\infty} \frac{1}{n!}((-iA)^\dagger)^n = e^{-iA^\dagger}\,, \tag{3.180}$$
and the commutation of A with itself, we have
$$UU^\dagger = e^{iA}e^{-iA^\dagger} = e^{iA}e^{-iA} = \mathbb{I}\,. \tag{3.181}$$
Note also that taking into account the spectral decomposition (3.83) we can write (3.178) in terms of the eigenvalues and eigenvectors of A, $A|n\rangle = a_n|n\rangle$, namely
$$U = \sum_n e^{ia_n}|n\rangle\langle n|\,. \tag{3.182}$$

3.4.2 The time-evolution operator

As anticipated, the state vector $|\varphi, t\rangle$ at time t is assumed to follow from the state vector $|\varphi, t_0\rangle$ at time t_0 by applying a unitary operator $U(t, t_0)$, called the *evolution operator*[6]

$$|\varphi, t\rangle = U(t, t_0)|\varphi, t_0\rangle\,. \tag{3.183}$$

The evolution operator has some important properties:

- $U(t_0, t_0) = \mathbb{I}$, as follows from the defining equation (3.183).
- Time composition works like matrix multiplication: we go from t_0 to t', from t' to t'' etc,
$$\left.\begin{aligned} |\varphi, t''\rangle &= U(t'', t')\,|\varphi, t'\rangle = U(t'', t')\,U(t', t_0)\,|\varphi, t_0\rangle \\ |\varphi, t''\rangle &= U(t'', t_0)\,|\varphi, t_0\rangle \end{aligned}\right\} \; U(t'', t_0) = U(t'', t')U(t', t_0) \tag{3.184}$$

[6] Note that, if it exists, this operator is unique.

- The "inverse" (or "dagger", since U is unitary) can be removed by flipping the order of the arguments

$$U^{-1}(t,t_0) = U(t_0,t) = U^{\dagger}(t,t_0)\ , \tag{3.185}$$

as follows straightforwardly from setting $t'' = t_0$ in the final result of (3.184),

$$\mathbb{I} = U(t_0,t_0) = U\left(t_0,t'\right)U\left(t',t_0\right)\ . \tag{3.186}$$

Since we are looking for a differential equation (the Schrödinger equation), let us consider the derivative of Eq. (3.183). Taking into account that $|\varphi,t_0\rangle$ is independent of t, we get

$$\frac{\partial}{\partial t}|\varphi,t\rangle = \left(\frac{\partial U(t,t_0)}{\partial t}\right)|\varphi,t_0\rangle = \frac{\partial U(t,t_0)}{\partial t}U(t_0,t)\,|\varphi,t\rangle = \frac{\partial U(t,t_0)}{\partial t}U^{\dagger}(t,t_0)\,|\varphi,t\rangle\ , \tag{3.187}$$

where we have made use of the definition (3.183) and the property (3.185) to get an equation for $|\varphi,t\rangle$. Let us explore the properties of the operator

$$\Lambda \equiv \frac{\partial U(t,t_0)}{\partial t}U^{\dagger}(t,t_0) \tag{3.188}$$

connecting the left and right-hand side of this expression:

- *Anti-Hermiticity.*
  Differentiating

$$U(t,t_0)\,U^{\dagger}(t,t_0) = \mathbb{I} \tag{3.189}$$

with respect to time, we obtain,[7]

$$\frac{\partial U(t,t_0)}{\partial t}U^{\dagger}(t,t_0) + U(t,t_0)\frac{\partial U^{\dagger}(t,t_0)}{\partial t} = 0 \quad \longrightarrow \quad \Lambda = -\Lambda^{\dagger}\,, \tag{3.191}$$

meaning that the operator connecting the left and right-hand side of Eq. (3.187) is *anti-Hermitian*.

- t_0 *independency.*
  Introducing the identity operator into the definition (3.188), we get

$$\begin{aligned}\Lambda(t,t_0) &= \frac{\partial U(t,t_0)}{\partial t}U^{\dagger}(t,t_0)\\ &= \frac{\partial U(t,t_0)}{\partial t}U\left(t_0,t'\right)U^{\dagger}\left(t_0,t'\right)U^{\dagger}(t,t_0)\\ &= \frac{\partial}{\partial t}\left(U(t,t_0)\,U\left(t_0,t'\right)\right)U\left(t',t_0\right)U(t_0,t)\ ,\end{aligned} \tag{3.192}$$

[7] Note that a time derivative does not interfere with daggers, and therefore

$$\Lambda^{\dagger} = U(t,t_0)\frac{\partial U^{\dagger}(t,t_0)}{\partial t}\,. \tag{3.190}$$

where in the second step we have taken into account that $U(t_0, t')$ has no t dependence to include it into the derivative and made use of the property (3.185). By composition (cf. (3.184)), this this is equal to

$$\Lambda(t, t_0) = \frac{\partial}{\partial t} U(t, t') U(t', t) = \frac{\partial}{\partial t} U(t, t') U^\dagger(t, t') = \Lambda(t, t') \,, \tag{3.193}$$

meaning that Λ is independent of t_0.

- *Units of time*: Since the unitary operators U have no units and we are taking a single time derivative, $[\Lambda] = [1/T]$.

We have therefore an operator Λ that is anti-Hermitian, only depends on t and has units of inverse time. Multiplying (3.187) by $i\hbar$ and defining a *Hermitian* operator

$$H \equiv i\hbar\Lambda = i\hbar \frac{\partial U(t, t_0)}{\partial t} U^\dagger(t, t_0) \,, \tag{3.194}$$

that we name *Hamiltonian*, we get

$$i\hbar \frac{\partial}{\partial t} |\varphi, t\rangle = H(t) |\varphi, t\rangle \,. \tag{3.195}$$

This is nothing else that the generalization of the Schrödinger equation (1.20) for a wave function $\varphi(x, t)$, as can be easily seen by multiplying (3.195) by $\langle x|$ from the left and identifying

$$\varphi(x, t) \equiv \langle x|\varphi, t\rangle \,, \qquad H(t) = \frac{P^2}{2m} + V(x) \,. \tag{3.196}$$

This is quite nice: The time-dependent Schrödinger equation is just the result of unitary time evolution!

Example 12: Building a Hamiltonian for spin 1/2 particles

Let us consider the time-independent Schrödinger equation for a spin one-half particle,

$$H\chi = E\chi \,, \tag{3.197}$$

with

$$H = \frac{P^2}{2m} \mathbb{I}_{2\times 2} \tag{3.198}$$

a matrix Hamiltonian and $\chi = (\chi_1, \chi_2)^t$ the so-called *Pauli spinor*, a column vector encoding the two possible spin degrees of freedom. Taking into account the identity

$$(\vec{\sigma} \cdot \vec{a})(\vec{\sigma} \cdot \vec{b}) = \vec{a} \cdot \vec{b}\, \mathbb{I}_{2\times 2} + i\vec{\sigma} \cdot (\vec{a} \times \vec{b}) \tag{3.199}$$

with $\vec{a} = \vec{b} = \vec{P}$ and recognizing that $\vec{P} \times \vec{P} = 0$, we can write

$$(\vec{\sigma} \cdot \vec{P}) \cdot (\vec{\sigma} \cdot \vec{P}) = P^2\, \mathbb{I}_{2\times 2} \,, \tag{3.200}$$

and therefore

$$H = \frac{1}{2m}(\vec{\sigma}\cdot\vec{P})(\vec{\sigma}\cdot\vec{P})\,. \tag{3.201}$$

The corresponding Hamiltonian for a charged particle interacting with an electromagnetic field follows from the minimal coupling prescription

$$\vec{P} \to \vec{\pi} \equiv \vec{P} - q\vec{A}(\vec{x})\,, \tag{3.202}$$

with q the charge of particle and $\vec{A}$ the external vector potential. Including also a possible electromagnetic scalar potential $\phi(\vec{x})$, we obtain the so-called *Pauli Hamiltonian*

$$H_{\text{Pauli}} = \frac{1}{2m}(\vec{\sigma}\cdot\vec{\pi})(\vec{\sigma}\cdot\vec{\pi}) + q\,\phi(\vec{x})\,. \tag{3.203}$$

In this context, the term $\vec{\pi}\times\vec{\pi}$ in Eq. (3.199) gains significance ($\vec{a}=\vec{b}=\vec{\pi}$), since its contribution to the resulting Hamiltonian

$$H_{\text{Pauli}} = \frac{1}{2m}[(\vec{\pi}\cdot\vec{\pi})\mathbb{I} + i\vec{\sigma}\cdot(\vec{\pi}\times\vec{\pi})] + q\phi(\vec{x}) \tag{3.204}$$

is no longer zero. Indeed, we have

$$(\vec{\pi}\times\vec{\pi})_k = \epsilon_{ijk}\pi_i\pi_j = \frac{1}{2}\epsilon_{ijk}\left[\pi_i,\pi_j\right]\,, \tag{3.205}$$

with[a]

$$[\pi_i,\pi_j] = [p_i - qA_i, p_j - qA_j] = -\frac{\hbar}{i}q\left(\partial_i A_j - \partial_j A_i\right) = iq\hbar\left(\partial_i A_j - \partial_j A_i\right)\,. \tag{3.206}$$

or equivalently

$$(\vec{\pi}\times\vec{\pi})_k = \frac{1}{2}\epsilon_{ijk}(i\,q\hbar)\left(\partial_i A_j - \partial_j A_i\right) = (i\,q\hbar)\epsilon_{ijk}\partial_i A_j = (i\,q\hbar)(\nabla\times A)_k = (i\,q\hbar)B_k\,. \tag{3.207}$$

Back to the Pauli Hamiltonian (3.204) and leaving implicit identity matrices, this implies

$$H_{\text{Pauli}} = \frac{1}{2m}\left(\vec{P} - q\vec{A}\right)^2 - \frac{q\hbar}{2m}\vec{\sigma}\cdot\vec{B} + q\,\phi(\vec{x})\,. \tag{3.208}$$

The second term in this expanded Pauli Hamiltonian represents the coupling of the charged particle to the magnetic field, aligning precisely with the one anticipated in the previous Chapter. Indeed, comparing this piece with Eq. (2.3) and taking into account Eq. (3.86), we observe that the Pauli equation fixes the electron's magnetic moment to $g = 2$.

[a] The A_i components are functions of position and therefore commute among themselves.

The quantum mechanics dualism

At this point, it is interesting to highlight the mysterious dualism of quantum mechanics. A given quantum state can change in two distinct ways. On the one hand, we have unitary evolution, which is clearly deterministic. Given an initial condition $|\varphi(0)\rangle$, the Schrödinger equation (3.195) is able to predict the state $|\varphi, t\rangle$ at a later time t. On the other hand, we have measurements, which are inevitably probabilistic. In particular, a quantum theory cannot make definite predictions on the result of a measurement, but just assign probabilities for the different outcomes to occur. This contrasts with the classical descriptions of nature we are all used to, where the relationship between the state of a system and the result of a measurement is trivial.
The quest for a more elegant and transparent formulation of quantum mechanics is a valuable endeavor that could be eventually successful. However, it is important to acknowledge that, while not axiomatically proven, the postulates presented here accurately capture our current understanding of quantum physics, being supported by a plethora of experiments.

3.4.3 Time-ordering and Dyson series

In the above derivation, the Hamiltonian H is a derived quantity, obtained from the unitary evolution operator U. In practical situations, however, we will postulate/construct a Hamiltonian and calculate the unitary evolution for it. To do so, let us consider again Eq. (3.194). By simple manipulations, this can be written as a differential equation for $U(t, t_0)$,

$$\frac{dU(t,t_0)}{dt} = -\frac{i}{\hbar} H(t) U(t,t_0)\,, \tag{3.209}$$

with boundary condition $U(t_0, t_0) = \mathbb{I}$ and the partial derivatives replaced by total derivatives since there is not potential confusion.[8] This differential equation admits an integral representation

$$U(t,t_0) = \mathbb{I} - \frac{i}{\hbar}\int_{t_0}^{t} H(t')U(t',t_0)dt'\,, \tag{3.210}$$

[8] As long as non-commuting operators are involved, all we are going to discuss here applies also to other scenarios in mathematical physics where this type of differential equation appears.

which can be formally solved by iteration. In particular, by recursively substituting the left-hand side of Eq. (3.210) into the right-hand side, we obtain the so-called *Dyson series*[9]

$$\begin{aligned} U(t,t_0) &= \mathbb{I} - \frac{i}{\hbar}\int_{t_0}^{t} H(t')\left[\mathbb{I} - \frac{i}{\hbar}\int_{t_0}^{t'} H(t'')U(t'',t_0)dt''\right]dt' \\ &= \mathbb{I} - \frac{i}{\hbar}\int_{t_0}^{t} H(t')dt' + \left(-\frac{i}{\hbar}\right)^2 \int_{t_0}^{t} dt' \int_{t_0}^{t'} dt''\, H(t')H(t'') \\ &+ \dots \\ &+ \left(-\frac{i}{\hbar}\right)^n \int_{t_0}^{t} dt' \int_{t_0}^{t'} dt'' \dots \int_{t_0}^{t^{(n-1)}} dt^{(n)} H(t')H(t'')\dots H(t^{(n)}) + \dots , \end{aligned} \tag{3.211}$$

with $t_0 \leq t^{(n)} \leq \dots \leq t'' \leq t' \leq t$, or more compactly,

$$U(t,t_0) = \mathbb{I} + \sum_{n=1}^{\infty}\left(-\frac{i}{\hbar}\right)^n \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \cdots \int_{t_0}^{t_{n-1}} dt_n\, H(t_1)H(t_2)\cdots H(t_n)\,, \tag{3.212}$$

where we have replaced the increasingly cumbersome notation $t', t'', \dots$ the integrating variables by $t_1, t_2, \dots t_n$, with $t_0 \leq t_n \leq \dots \leq t_2 \leq t_1 \leq t$, most suitable for the subsequent developments. As we will discuss in Chapter 8, truncating this infinite sum is a usual procedure in the context of time-dependent perturbation theory, yielding approximate solutions for $U(t,t_0)$. It is nonetheless, possible to recast it in a much simpler exponential-like form. To achieve this goal, we must rewrite the integration regions in such a way that they do not depend of the integration variables, but instead go simply from t_0 to t in all the n integrations; we must uncouple the integrals. As a first step, let us define an instruction to reorder the operators in such a way that the corresponding time arguments decrease as we move from the left to the right, the so-called *time-ordered product of operators*

$$T\left[H(t_1)\,H(t_2)\cdots H(t_n)\right] = H(t_{i_1})\,H(t_{i_2})\cdots H(t_{i_n})\,, \tag{3.213}$$

with $t_{i_1} > t_{i_2} > \cdots > t_{i_n}$. Consider for instance the integral in the second line of Eq. (3.211). This can be written as

$$I_2 \equiv \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2\, H(t_1)\,H(t_2) = \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2\, T\left[H(t_1)\,H(t_2)\right]\,, \tag{3.214}$$

where, in order to insert the T symbol, we have taken into account that the integration region should be given by half of the square region $0 \leq t_1, t_2 \leq t$ where $t_2 \leq t_1$. The same integral can be also evaluated by interchanging the order of integration, without changing of course the value of the integral, i.e.

$$I_2 = \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2\, H(t_1)\,H(t_2) = \int_{t_0}^{t} dt_2 \int_{t_2}^{t} dt_1\, H(t_1)\,H(t_2)\,. \tag{3.215}$$

Since the times t_1 and t_2 in this expression are jusrt dummy integration variables, we can relabel them in the last integral as $t_1 \to t_2$ and $t_2 \to t_1$, getting

$$I_2 = \int_{t_0}^{t} dt_1 \int_{t_1}^{t} dt_2\, H(t_2)\,H(t_1) = \int_{t_0}^{t} dt_1 \int_{t_1}^{t} dt_2\, T\left[H(t_1)\,H(t_2)\right]\,, \tag{3.216}$$

[9] Note that the operators H in this expression with different time-arguments do not generically commute. The order must be preserved!

where the T symbol in the last expression follows from the condition $t_2 \geq t_1$, in which case $T\left[H\left(t_1\right) H\left(t_2\right)\right] = H\left(t_2\right) H\left(t_1\right)$. In other words, the integration region is now given by the area of the triangle above the diagonal line in the previous figure. We have therefore two different ways of writing I_2, namely

$$I_2 = \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2\, T\left[H\left(t_1\right) H\left(t_2\right)\right] , \qquad I_2 = \int_{t_0}^{t} dt_1 \int_{t_1}^{t} dt_2\, T\left[H\left(t_1\right) H\left(t_2\right)\right] , \tag{3.217}$$

with t_1 appearing respectively as a upper and lower index. Therefore, by adding these two integrals,

$$\begin{aligned} 2I_2 &= \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2\, T\left[H\left(t_1\right) H\left(t_2\right)\right] + \int_{t_0}^{t} dt_1 \int_{t_1}^{t} dt_2\, T\left[H\left(t_1\right) H\left(t_2\right)\right] \\ &= \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2\, T\left[H\left(t_1\right) H\left(t_2\right)\right] , \end{aligned} \tag{3.218}$$

we can eliminate the t_1 dependence, getting an expression whose integration region coincides with the full square area, i.e. $0 \leq t_1, t_2 \leq t$. Dividing this result by two, we end up with

$$I_2 = \frac{1}{2} \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2\; T\left[H\left(t_1\right) H\left(t_2\right)\right] . \tag{3.219}$$

This procedure can be easily extended to higher-orders. In particular, for the n-th term in Eq. (3.211), the original integration region in I_n is given by the hyper-volume bounded by the hypercube defined by $0 \leq t_1, t_2, \ldots, t_n \leq t$ such that $t_n \leq \ldots \leq t_3 \leq t_2 \leq t_1$, which is a fraction $1/n!$ of the hypercube. As before, by simple interchanging the order of integration in $n! - 1$ possible ways and appropriately relabeling the integration variables, we can obtain alternative expressions for I_n covering the remaining $n! - 1$ regions of the hypercube, such that we end with an integral over the entire hypercube $0 \leq t_1, t_2, \ldots, t_n \leq t$, namely

$$\begin{aligned} I_n &\equiv \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \cdots \int_{t_0}^{t_{n-1}} dt_n H\left(t_1\right) H\left(t_2\right) \cdots H\left(t_n\right) \\ &= \frac{1}{n!} \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 \cdots \int_{t_0}^{t} dt_n T\left[H\left(t_1\right) H\left(t_2\right) \cdots H\left(t_n\right)\right] . \end{aligned} \tag{3.220}$$

Using this result, Eq. (3.212) becomes equals to the Taylor expansion of a time-order exponential

$$U\left(t, t_0\right) = \mathbb{I} + \sum_{n=1}^{\infty} \frac{1}{n!} \left(-\frac{i}{\hbar}\right)^n \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 \cdots \int_{t_0}^{t} dt_n\; T\left[H\left(t_1\right) H\left(t_2\right) \cdots H\left(t_n\right)\right] , \tag{3.221}$$

which is usually compactly written as

$$\boxed{U(t, t_0) = T\left[\exp\left(-\frac{i}{\hbar} \int_{t_0}^{t} dt' H(t')\right)\right] .} \tag{3.222}$$

Given this general expression, we can considered two *specific cases*:

- If the Hamiltonians evaluated at different times commute, e.g. $[H(t'), H(t'')] = 0$, we can simply drop the T symbol in the previous expression. Eq. (3.222) simplifies then to

$$\begin{aligned} U(t,t_0) &= \mathbb{I} + \sum_{n=1}^{\infty} \frac{1}{n!} \left(-\frac{i}{\hbar}\right)^n \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 \cdots \int_{t_0}^{t} dt_n \, H(t_1) H(t_2) \cdots H(t_n) \\ &= \mathbb{I} + \sum_{n=1}^{\infty} \frac{1}{n!} \left(-\frac{i}{\hbar}\right)^n \left(\int_{t_0}^{t} H(t') \, dt'\right)^n = \exp\left(-\frac{i}{\hbar} \int_{t_0}^{t} dt' H(t')\right) , \end{aligned} \tag{3.223}$$

  where in the last step we have recalled the Taylor series of the exponential.

- For time-independent Hamiltonians, Eq. (3.222) becomes only depending on $(t - t_0)$,

$$U(t,t_0) = \exp\left(-\frac{i(t-t_0)}{\hbar} H\right) . \tag{3.224}$$

  This means that if the initial state of the system is an eigenstate of the Hamiltonian operator with eigenvalue E,

$$H|\varphi, t_0\rangle = E|\varphi, t_0\rangle , \tag{3.225}$$

  the action of the operator (3.224) becomes a trivial multiplication by a phase factor

$$|\varphi, t\rangle = e^{-\frac{iE(t-t_0)}{\hbar}} |\varphi, t_0\rangle . \tag{3.226}$$

  The state is then *stationary*; all its observable properties are constant in time.

Example 13: Spin 1/2 in a uniform magnetic field

Consider a spin 1/2 particle with aspin pointing along the direction specified by some angles (θ_0, ϕ_0) at time $t = 0$ (cf. Eq. (3.33)),

$$|\varphi, 0\rangle = \cos\frac{\theta_0}{2}|+\rangle + \sin\frac{\theta_0}{2} e^{i\phi_0}|-\rangle , \tag{3.227}$$

and placed in a uniform magnetic field oriented along the z axis $\vec{B} = B\vec{z}$. The associated Hamiltonian takes the form

$$H = -\vec{\mu} \cdot \vec{B} = -\gamma \vec{S} \cdot \vec{B} = \omega S_z , \tag{3.228}$$

where, for notational convenience, we have introduced the so-called Larmor frequency

$$\omega \equiv -\gamma B , \tag{3.229}$$

with $\gamma = gq/(2m)$ the Landé factor. Since the Hamiltonian is constant, we can make use of the expression (3.224) with $t = 0$,

$$U(t,0) = \exp\left(-\frac{iHt}{\hbar}\right) = \exp\left(-\frac{i\omega S_z t}{\hbar}\right) , \tag{3.230}$$

to determine the state of the system an any subsequent time $t > 0$, namely

$$|\varphi, t\rangle = e^{-iHt/\hbar}|\varphi, 0\rangle = e^{-\frac{i\omega t}{2}}\left(\cos\frac{\theta_0}{2}|+\rangle + \sin\frac{\theta_0}{2}e^{i(\phi_0+\omega t)}|-\rangle\right), \tag{3.231}$$

where we have taken into account that $S_z|\pm\rangle = \pm\hbar/2|\pm\rangle$ and factored out an overall and physically irrelevant phase. The evolved spin vector (3.231) corresponds therefore to a rotation around the z axis, with constant inclination $\theta(t) = \theta_0$ and angular velocity ω. Indeed, spin precesses in exactly the same way as magnetic dipoles do in classical electromagnetism. This is the basic principle behind the nuclear magnetic resonance technique, used chemistry and medicine. Numerically, a hydrogen atom in the earth magnetic field $B \simeq 5 \cdot 10^{-7}\,\mathrm{T}$ precesses with a frequency

$$\frac{\gamma B}{2\pi} \simeq 42.6\,\frac{\mathrm{MHz}}{\mathrm{T}} \times 5 \cdot 10^{-7}\,\mathrm{T} \simeq 2.13\,\mathrm{kHz}\,. \tag{3.232}$$

The Green's function in position space

Writing Eq. (3.209) in the position representation, we get a differential equation

$$\frac{d}{dt}G\left(x, t; x', t_0\right) = -\frac{i}{\hbar}H\left(x, -i\hbar\frac{d}{dx}\right)G\left(x, t; x', t_0\right), \tag{3.233}$$

for the matrix elements of the evolution operator, the so-called *propagator* or *Green's function in position space*,

$$G\left(x, t; x', t_0\right) \equiv \langle x|U(t, t_0)|x'\rangle\,, \tag{3.234}$$

with initial condition

$$G\left(x, t_0; x', t_0\right) = \delta\left(x - x'\right). \tag{3.235}$$

Given $G\left(x, t; x', t_0\right)$, the time-dependent solution of the Schrödinger equation for any given initial condition $\varphi(x, t_0)$ is obtained as (cf. Eq. (3.183))

$$\varphi(x, t) = \int_{\mathbb{R}} G\left(x, t; x', t_0\right)\varphi\left(x' t_0\right)dx'\,. \tag{3.236}$$

In addition to providing an explicit solution to the time-dependent Schrödinger equation, the propagator (3.234), and in particular the quantity $|G(x, t; x', t_0)|^2$, can be interpreted as a proxy for the probability density of finding the system at the point x at time t if it was initially localized around the point x' at time t_0. [a] Moreover, in the Feynman's path integral formulation of quantum mechanics, the propagator admits a natural interpretation as a sum over all possible paths connecting the initial point x' at time t_0 to the final point x at time t.

[a] Strictly speaking, since the position eigenstates $|x\rangle$ are not normalizable, $|G(x, t; x', t_0)|^2$ is only proportional to the probability density, rather than exactly equal to it.

3.4.4 The Schrödinger and Heisenberg pictures

Using the evolution operator U, one can express the time development of the expectation value of a given combination of operators $AB\cdots C$ in two equivalent representations,

$$\langle\varphi,t|AB\cdots C|\varphi,t\rangle = \langle\varphi,t_0|U^\dagger AB\cdots CU|\varphi,t_0\rangle = \langle\varphi,t_0|U^\dagger AUU^\dagger BU\cdots U^\dagger CU|\varphi,t_0\rangle. \tag{3.237}$$

The left-hand side of this expression is said to be in the *Schrödinger representation*, with the states evolving while the operators are fixed. The right-hand side is said to be in the Heisenberg representation, with the states fixed and the operators evolving according to

$$A_H(t) = U^\dagger(t)A_S U(t) \tag{3.238}$$

with the subindices S and H referring respectively to the Schrödinger and Heisenberg representations. Note that even if A_S has no time dependence, the operators $A_H(t)$ will depend on time unless A_S commutes with the Hamiltonian, as follows from

$$\frac{dA_H(t)}{dt} = \frac{\partial}{\partial t}\left[U^\dagger(t)A_S(t)U(t)\right] = \frac{i}{\hbar}U^\dagger(t)\left[H_S(t),A_S(t)\right]U(t) + U^\dagger(t)\frac{\partial A_S(t)}{\partial t}U(t)\,, \tag{3.239}$$

which taking into account that

$$U^\dagger(t)\left[H_S(t),A_S(t)\right]U(t) = \left[H_H(t),A_H(t)\right]\,, \qquad U^\dagger(t)\frac{\partial A_S(t)}{\partial t}U(t) = \frac{\partial A_H(t)}{\partial t}\,, \tag{3.240}$$

can be compactly written as

$$\frac{dA_H(t)}{dt} = -\frac{i}{\hbar}\left[A_H(t),H_H(t)\right] + \frac{\partial}{\partial t}A_H(t)\,. \tag{3.241}$$

This is the *Heisenberg equation of motion.*

3.4.5 Time-evolution of expectation values

Let us make use of the Heisenberg picture to compute the evolution of a *time-independent* Schrödinger operator A_S. Taking into account Eqs. (3.238) and (3.241), we have

$$\frac{d}{dt}\langle\varphi,t|A_S|\varphi,t\rangle = \frac{d}{dt}\langle\varphi,t_0|A_H|\varphi,t_0\rangle = \langle\varphi,t_0|\frac{dA_H}{dt}|\varphi,t_0\rangle = -\frac{i}{\hbar}\langle\varphi(t_0)|[A_H,H_H]|\varphi,t_0\rangle\,, \tag{3.242}$$

or more compactly

$$\frac{d\langle A_H\rangle_\varphi}{dt} = -\frac{i}{\hbar}\left\langle[A_H,H_H]\right\rangle_\varphi\,. \tag{3.243}$$

Note that since the expectation values of Schrödinger operators are the same as the expectation values of Heisenberg operators, this can be also written as

$$\frac{d}{dt}\left\langle A_S\right\rangle_\varphi = -\frac{i}{\hbar}\left\langle[A_S,H_S]\right\rangle_\varphi\,. \tag{3.244}$$

These equations are a natural generalization of the classical Hamiltonian equation for an observable function Q of x and p,

$$\frac{dQ}{dt} = \{Q, H\}\,, \tag{3.245}$$

with the formal replacement of the Poisson brackets

$$\{A, B\} \equiv \frac{\partial A}{\partial x}\frac{\partial B}{\partial p} - \frac{\partial A}{\partial p}\frac{\partial B}{\partial x}\,, \tag{3.246}$$

by commutators,

$$\{A, B\} \Longrightarrow -\frac{i}{\hbar}[A, B]\,. \tag{3.247}$$

The factor i in this prescription is important for hermiticity, since the commutator, by itself, is not Hermitian. As in classical mechanics, Eqs. (3.243) and (3.244) come also with a very important physical consequence: an operator A is conserved if it commutes with the Hamiltonian! This conclusion holds irrespectively of whether we use the Heisenberg picture or the Schrödinger picture.

Example 14: The quantum Newton's law

Consider the Hamiltonian of a moving particle in one dimension,

$$H = \frac{P^2}{2m} + V(X)\,, \tag{3.248}$$

with X and P position and momentum operators satisfying the commutation relation $[X, P] = i\hbar$. Since these operators are independent of time, the evolution of their expectation values is given by

$$\frac{d\langle P\rangle}{dt} = -\frac{i}{\hbar}\langle[P, H]\rangle\,, \qquad \frac{d\langle X\rangle}{dt} = -\frac{i}{\hbar}\langle[X, H]\rangle\,, \tag{3.249}$$

which taking into account that

$$[X, H] = \frac{1}{2m}\left[X, P^2\right] = \frac{i\hbar}{m}P\,, \qquad [P, H] = [P, V(X)] = -i\hbar\frac{\partial V(X)}{\partial X}\,, \tag{3.250}$$

can we rewritten as

$$\frac{d\langle P\rangle}{dt} = -\left\langle\frac{\partial V(X)}{\partial X}\right\rangle\,, \qquad \frac{d\langle X\rangle}{dt} = \frac{\langle P\rangle}{m}\,. \tag{3.251}$$

Combining these expressions, we obtain the quantum mechanical version of the Newton's law,

$$m\frac{d^2\langle X\rangle}{dt} = -\left\langle\frac{\partial V(X)}{\partial X}\right\rangle\,. \tag{3.252}$$

Note that this expression reduces to its classical counterpart,

$$m\frac{d^2x}{dt^2} = -\frac{\partial V(x)}{\partial x}\,, \tag{3.253}$$

only if those situations in which the average of the potential derivative can be replaced by a function of the average of X,

$$\left\langle \frac{\partial V(X)}{\partial X} \right\rangle \simeq \frac{\partial V(\langle X \rangle)}{\partial \langle X \rangle} \,. \tag{3.254}$$

This approximation is valid whenever the probability distribution is narrowly peaked around $\langle X \rangle$, or, equivalently, if the uncertainty ΔX is sufficiently small.

3.4.6 The time-energy uncertainty relation

Combining Eqs. (3.244) and (3.160) with B replaced by H, we get

$$\Delta_\varphi H \Delta_\varphi A \geq \frac{1}{2} \left| \langle [A, H] \rangle_\varphi \right| = \frac{\hbar}{2} \left| \frac{d\langle A \rangle_\varphi}{dt} \right| \,. \tag{3.255}$$

This expression can be alternatively rewritten as

$$\Delta_\varphi H \, \Delta_\varphi \tau \geq \frac{\hbar}{2} \,, \tag{3.256}$$

with

$$\Delta_\varphi \tau(A) = \frac{\Delta_\varphi A}{|d\langle A \rangle_\varphi / dt|} \tag{3.257}$$

the characteristic time in which the change in the expectation value of A becomes comparable to the uncertainty $\Delta_\varphi A$. In other words, it is the time after which we can deduce that a change has occurred by measuring A. For this reason, $\Delta_\varphi \tau(A)$ is sometimes called the lifetime of the state $|\varphi\rangle$ with respect to the observable A. Note, however, that in spite of the appearances, the above expression is not completely analogue to the uncertainty relation (3.161), since there is not a time operator T in quantum mechanics that could satisfy the commutation relation $[T, H] = i\hbar$ entering Eq. (3.160).

The inequality (3.256) is sometimes heuristically written as

$$\Delta E \, \Delta t \geq \frac{\hbar}{2} \,, \tag{3.258}$$

with ΔE the energy spread and Δt the characteristic evolution time. According to the above reasoning, we can interpret Δt as the lifetime of an excited state, existing only for a limited amount of time. A particular example could be an atom in an excited energy state eventually transitioning to a lower one by its interactions with the electromagnetic field. The energy of the emitted electron has an uncertainty $\Delta E = \hbar / 2\Delta t$, which is usually expressed in terms of the so-called *decay width* $\Gamma \equiv \Delta E / \hbar$. As explained in Section 3.3.3, ΔE is unrelated to the accuracy with which the energy can be measured. An example of this is the unstable Z boson mediating the weak interaction, whose mass or rest-frame energy is known more precisely than its width,

$$E_Z = m_Z c^2 = 91.1875 \pm 0.0021 \text{ GeV} \,, \qquad \Delta E_Z = \hbar\Gamma = 2.4952 \pm 0.0023 \text{ GeV} \,.$$

In other words, the peak's position can be identified much more precisely than its spread.

3.5 Some worked-out exercises

Physical states

A cavity supporting a single mode of oscillation of the electromagnetic field can have states with $n = 0, 1, 2, \ldots$ photons present, from which we can construct the basis of orthonormal vectors $\{|n\rangle; n = 0, 1, 2, \ldots\}$. Which of the following states could represent a physical state of the system?

$$|\varphi\rangle = \sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} |n\rangle \,, \qquad |\phi\rangle = \sum_{n=1}^{\infty} \frac{1}{n} |n\rangle \,, \qquad |\alpha\rangle = \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \,, \quad \alpha \in \mathbb{C} \,.$$

A physical state must be normalizable. The inner product $\langle\varphi|\varphi\rangle$ gives rise to a divergent series,

$$\langle\varphi|\varphi\rangle = \sum_{n=1}^{\infty} \frac{1}{n} \,,$$

not admitting therefore a probability interpretation; this vector does not represent a possible physical state of the system. On the contrary, the inner product $\langle\phi|\phi\rangle$ gives rise to a convergent series

$$\langle\phi|\phi\rangle = \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} \,,$$

allowing to normalize the state to unity and supporting its interpretations as a physical state of the system. Finally, for the final state $|\alpha\rangle$ we have

$$\langle\alpha| = \sum_{n=0}^{\infty} \frac{\alpha^{*n}}{\sqrt{n!}} \langle n| \,, \qquad \langle\alpha|\alpha\rangle = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{\alpha^{*m}}{\sqrt{m!}} \frac{\alpha^n}{\sqrt{n!}} \underbrace{\langle m|n\rangle}_{\delta_{nm}} = \sum_{n=0}^{\infty} \frac{|\alpha|^{2n}}{n!} = e^{|\alpha|^2} \,,$$

allowing to normalize the state,

$$|\alpha\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \,,$$

and supporting its interpretations as a physical state of the system. This type of states are called coherent states.

The Hadamard gate

One of the basic gates in quantum computing is the Hadamard gate

$$H = \frac{1}{\sqrt{2}} (\sigma_x + \sigma_z) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \,,$$

with σ_x and σ_z the Pauli matrices.

a) Show that $H^2 = \mathbb{I}$. Determine the eigenvalues and eigenvectors of H in the basis of eigenstates of σ_z.

b) Determine the form of H in the bases of σ_x and σ_y.

c) Provide a geometric interpretation of H in terms of the Bloch sphere.

a) By direct computation, we have

$$H^2 = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} = \mathbb{I}\,.$$

Alternatively, one can write

$$H = \frac{1}{\sqrt{2}}\Big[|0\rangle\langle 0| + |1\rangle\langle 0| + |0\rangle\langle 1| - |1\rangle\langle 1|\Big]\,,$$

with $|0\rangle \equiv |+\rangle$ and $|1\rangle \equiv |-\rangle$ the σ_z eigenstates (the $|0\rangle$ and $|1\rangle$ notation is chosen here for clarity), and use $\langle 0|0\rangle = \langle 1|1\rangle = 1$ and $\langle 0|1\rangle = \langle 1|0\rangle$ to obtain

$$\begin{aligned} H^2 &= \frac{1}{2}\Big[|0\rangle\langle 0| + |1\rangle\langle 0| + |0\rangle\langle 1| - |1\rangle\langle 1|\Big]\Big[|0\rangle\langle 0| + |1\rangle\langle 0| + |0\rangle\langle 1| - |1\rangle\langle 1|\Big] \\ &= \frac{1}{2}\Big[|0\rangle\langle 0|0\rangle\langle 0| + |0\rangle\langle 0|0\rangle\langle 1| + |1\rangle\langle 0|0\rangle\langle 0| + |1\rangle\langle 0|0\rangle\langle 1| \\ &\quad + |0\rangle\langle 1|1\rangle\langle 0| - |0\rangle\langle 1|1\rangle\langle 1| - |1\rangle\langle 1|1\rangle\langle 0| + |1\rangle\langle 1|1\rangle\langle 1|\Big] \\ &= \frac{1}{2}\Big[2|0\rangle\langle 0| + 2|1\rangle\langle 1|\Big] = \mathbb{I}\,. \end{aligned}$$

Since $H^2 = \mathbb{I}$, the eigenvalues must be either $+1$ or -1,

$$H^2_{\text{diag}} = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} = \begin{pmatrix} a^2 & 0 \\ 0 & b^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \longrightarrow \quad a = \pm 1\,,\ b = \pm 1\,.$$

However, if both eigenvalues were equal, H would be proportional to the identity matrix, which is not. Thus, one eigenvalue is $+1$ and the other -1. By direct calculation, we can find that the (normalized) eigenvectors are

$$(H - \lambda_\pm I)|\lambda_\pm\rangle = 0 \;\longrightarrow\; \begin{pmatrix} 1/\sqrt{2} \mp 1 & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \mp 1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \;\longrightarrow\; |\lambda_\pm\rangle = \frac{(1 \pm \sqrt{2})|0\rangle + |1\rangle}{\sqrt{2(2 \pm \sqrt{2})}}\,.$$

b) The eigenbasis of σ_x is formed by the two states $|\pm\rangle_x = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$. From the form of H in the σ_z basis,

$$H = \frac{1}{\sqrt{2}}(|0\rangle\langle 0| + |1\rangle\langle 0| + |0\rangle\langle 1| - |1\rangle\langle 1|)\,,$$

it is clear that we can express H as

$$H = |+\rangle_x \langle 0| + |-\rangle_x \langle 1|\,, \qquad \text{or} \qquad H = |0\rangle_x\langle +| + |1\rangle_x\langle -|\,,$$

where the last expression follows from the first, since $H^\dagger = H$. Finally, we can express the σ_z basis $|0\rangle$ and $|1\rangle$ in terms of the σ_x basis as

$$|0\rangle = \frac{1}{\sqrt{2}}(|+\rangle_x + |-\rangle_x)\,, \qquad |1\rangle = \frac{1}{\sqrt{2}}(|+\rangle_x - |-\rangle_x)\,.$$

Inserting these expressions in the equation for H we find

$$H = |0\rangle\,{}_x\langle +| + |1\rangle\,{}_x\langle -| = \frac{1}{\sqrt{2}}(|+\rangle_x\,{}_x\langle +|) + (|-\rangle_x\,{}_x\langle +|) + (|+\rangle_x\,{}_x\langle -|) - (|-\rangle_x\,{}_x\langle -|)\,,$$

or

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} .$$

Evidently, H has exactly the same representation in the σ_x basis. We could indeed anticipated this immediately once we noticed that H interchanges the σ_z and σ_x bases.
For σ_y, we can proceed differently and determine the action of H on the σ_y eigenstates $|\pm\rangle_y = \frac{1}{\sqrt{2}}(|0\rangle \pm i|1\rangle)$, namely

$$\begin{aligned} H\,|\pm\rangle_y &= \frac{1}{\sqrt{2}}(H|0\rangle \pm iH|1\rangle) = \frac{1}{2}(|0\rangle + |1\rangle \pm i|0\rangle \mp i|1\rangle) = \left(\frac{1 \pm i}{2}\right)|0\rangle + \left(\frac{1 \mp i}{2}\right)|1\rangle \\ &= \frac{1}{\sqrt{2}} e^{i\pm\frac{\pi}{4}} \left(|0\rangle + \left(\frac{1 \mp i}{1 \pm i}\right)|1\rangle\right) = \frac{1}{\sqrt{2}} e^{i\pm\frac{\pi}{4}} (|0\rangle \mp i|1\rangle) = e^{i\pm\frac{\pi}{4}}\,|\mp\rangle_y \,. \end{aligned}$$

Thus, the Hadamard operator

$$H = \begin{pmatrix} 0 & e^{-i\frac{\pi}{4}} \\ e^{i\frac{\pi}{4}} & 0 \end{pmatrix}$$

just swaps the two states in the σ_y basis.

c) Since $H^2 = \mathbb{I}$, the action of H must be a π rotation. Since

$$H = \frac{1}{\sqrt{2}}\,(\sigma_x + \sigma_z)$$

and the Pauli matrices represent rotations of angle π around the corresponding axis, H represents a π rotation around the axis $\frac{1}{\sqrt{2}}(x + z)$.

Operator representations

A given operator H on $\mathbb{C}^2$ is defined as

$$H|1\rangle = -\frac{1}{2}|1\rangle + \frac{\sqrt{3}}{2}|2\rangle\,, \qquad H|2\rangle = \frac{\sqrt{3}}{2}|1\rangle + \frac{1}{2}|2\rangle\,,$$

with $|1\rangle$ and $|2\rangle$ an orthonormal basis for $\mathbb{C}^2$.

a) Find the matrix representation of H with respect to the basis $\{|1\rangle, |2\rangle\}$.

b) Show that $H^2 = \mathbb{I}$, with $\mathbb{I}$ the identity operator on $\mathbb{C}^2$.

c) Express H in Dirac notation in this basis, i.e., write H in the form $\sum_{i,j=1}^{2} |i\rangle\langle j|H|i\rangle\langle j|$. Verify that the resulting expression yields the correct results when applied to $|1\rangle$ and $|2\rangle$.

d) Determine the eigenvalues and eigenvectors of H.

e) Write H in Dirac notation with respect to this new basis, i.e., write H in the form $\sum_{i,j=1}^{2} |e_i\rangle\langle e_j|$ with $|e_1\rangle$ and $|e_2\rangle$ the orthonormal eigenvectors in d). Confirm that the result matches the expression found in c).

a)

$$H = \begin{pmatrix} -1/2 & \sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{pmatrix} .$$

$H^\dagger$ is the conjugate transpose. Since the off-diagonal entries are the same and real, $H = H^\dagger$.

b)
$$H^2 = \begin{pmatrix} -1/2 & \sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{pmatrix} \begin{pmatrix} -1/2 & \sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbb{I}\,.$$

c)
$$H = -\frac{1}{2}|1\rangle\langle 1| + \frac{\sqrt{3}}{2}|1\rangle\langle 2| + \frac{\sqrt{3}}{2}|2\rangle\langle 1| + \frac{1}{2}|2\rangle\langle 2|\,,$$

so
$$H|1\rangle = -\frac{1}{2}|1\rangle\langle 1|1\rangle + \frac{\sqrt{3}}{2}|1\rangle\langle 2|1\rangle + \frac{\sqrt{3}}{2}|2\rangle\langle 1|1\rangle + \frac{1}{2}|2\rangle\langle 2|1\rangle = -\frac{1}{2}|1\rangle + \frac{\sqrt{3}}{2}|2\rangle\,,$$

and
$$H|2\rangle = -\frac{1}{2}|1\rangle\langle 1|2\rangle + \frac{\sqrt{3}}{2}|1\rangle\langle 2|2\rangle + \frac{\sqrt{3}}{2}|2\rangle\langle 1|2\rangle + \frac{1}{2}|2\rangle\langle 2|2\rangle = \frac{\sqrt{3}}{2}|1\rangle + \frac{1}{2}|2\rangle\,.$$

d) To calculate the eigenvalues of H, we let
$$\det \begin{pmatrix} -1/2-\lambda & \sqrt{3}/2 \\ \sqrt{3}/2 & 1/2-\lambda \end{pmatrix} = (-1/2-\lambda)(1/2-\lambda) - 3/4 = 0\,, \qquad \lambda = \pm 1\,.$$

The eigenvector for $\lambda = 1$ is given by
$$\begin{pmatrix} -3/2 & \sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This gives us that $a = b/\sqrt{3}$. The eigenvector is therefore some multiple of $\left(\frac{1}{\sqrt{3}}, 1\right)^t$. We normalise it to
$$|e_1\rangle = \frac{1}{2}|1\rangle + \frac{\sqrt{3}}{2}|2\rangle\,.$$

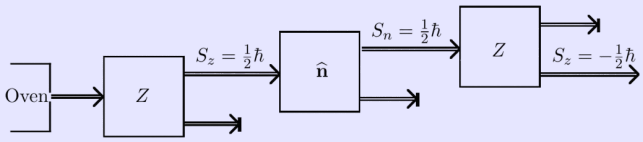
The eigenvector for $\lambda = -1$ is given by
$$\begin{pmatrix} 1/2 & \sqrt{3}/2 \\ \sqrt{3}/2 & 3/2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

So $a = -\sqrt{3}b$ and the eigenvector is some multiple of $(-\sqrt{3}, 1)^t$. We normalise it to
$$|e_2\rangle = \frac{-\sqrt{3}}{2}|1\rangle + \frac{1}{2}|2\rangle\,.$$

e)
$$\begin{aligned} H &= |e_1\rangle\langle e_1| - |e_2\rangle\langle e_2| \\ &= \left(\frac{1}{2}|1\rangle + \frac{\sqrt{3}}{2}|2\rangle\right)\left(\frac{1}{2}\langle 1| + \frac{\sqrt{3}}{2}\langle 2|\right) - \left(-\frac{\sqrt{3}}{2}|1\rangle + \frac{1}{2}|2\rangle\right)\left(-\frac{\sqrt{3}}{2}\langle 1| + \frac{1}{2}\langle 2|\right) \\ &= \frac{1}{4}|1\rangle\langle 1| + \frac{\sqrt{3}}{4}|1\rangle\langle 2| + \frac{\sqrt{3}}{4}|2\rangle\langle 1| + \frac{3}{4}|2\rangle\langle 2| - \frac{3}{4}|1\rangle\langle 1| + \frac{\sqrt{3}}{4}|1\rangle\langle 2| + \frac{\sqrt{3}}{4}|2\rangle\langle 1| - \frac{1}{4}|2\rangle\langle 2| \\ &= -\frac{1}{2}|1\rangle\langle 1| + \frac{\sqrt{3}}{2}|1\rangle\langle 2| + \frac{\sqrt{3}}{2}|2\rangle\langle 1| + \frac{1}{2}|2\rangle\langle 2|. \end{aligned}$$

Stern-Gerlach experiments and probabilities

Consider a beam of spin 1/2 particles passing through a series of three Stern-Gerlach devices, as illustrated in the figure.



The first device filters out particles with $S_z = -\hbar/2$. The remaining $S_z = \hbar/2$ component proceeds them towards a second device, whose magnetic field is oriented in a direction $\vec{n}$ in the xz plane, making an angle θ with the z axis. This device transmits particles with $S_{\vec{n}} = \vec{S} \cdot \vec{n} = \hbar/2$, filtering out those with $S_{\vec{n}} = -\hbar/2$. Finally, the last device transmits only particles with $S_z = -\hbar/2$.

a) Determine the fraction of particles going through the first Stern-Gerlach device that survive to the third measurement. How must the second-device magnetic field be oriented in order to maximize the number of particles transmitted by the final device? Which fraction of particles survive for this value of θ?

b) What fraction of the particles survive the last measurement if the second Stern-Gerlach device is removed?

a) The first device filters out particles with $S_z = -\hbar/2$. This gives a probability factor 1/2. Since the relative angle between the two magnetic fields in the first and second Stern-Gerlach devices is θ, the probability of a particle emerging from the second device with $S_{\vec{n}} = \hbar/2$ is simply $\cos^2(\theta/2)$. The relative angle of the magnetic fields between the second and the third Stern-Gerlach devices is $-\theta$. Therefore, the probability of a particle emerging from the second device with $S_z = \hbar/2$ is $\cos^2(-\theta/2) = \cos^2(\theta/2)$ and the probability of emerging with $S_z = -\hbar/2$ is given by $\sin^2(\theta/2)$. Consequently the probability of a particle emerging in the final device with $S_z = -\hbar/2$ beam is

$$P = \frac{1}{2}\cos^2\frac{\theta}{2}\sin^2\frac{\theta}{2} = \frac{1}{8}\sin^2\theta\,.$$

b) The maximum value of $\sin^2\theta$ happens for $\theta = \pi/2$, i.e. when the second magnetic field is oriented in the x direction. The maximum fraction of atoms making it through the three of Stern-Gerlach devices is then 1/8. Half of the initial particles are filtered out in the first device, half of the passing through particles are lost in the second device, and from those making it through, another half will be lost when passing through the final device.

c) If the intermediate Stern-Gerlach device is removed, all atoms entering the last device have $S_z = \frac{1}{2}\hbar$, therefore no atoms with $S_z = -\frac{1}{2}\hbar$ emerge.

Sequential measurements in a two-level system

Consider a quantum two level system with Hamiltonian

$$H = E_0 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

with E_0 a positive real constant. The system has only two additional observable

quantities

$$A = \hbar \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \qquad B = \hbar \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}.$$

a) Identify the complete set of commuting observables that includes the Hamiltonian.

b) Assuming the system to be initially ($t = 0$) in the state $|\varphi, 0\rangle = (1\,0)^t$, determine the state of the system at time $t > 0$.

c) If a measurement of the observable A is performed on $|\varphi, 0\rangle$ at time $t = T$, what values can be obtained? What are the corresponding probabilities?

d) Suppose that the measurement of the operator A at time T gave the larger of the two possible results. What are the possible outcomes of measuring the energy at $t = 3T$? And the corresponding probabilities?

a)

$$[H, A] = HA - AH = E_0\hbar \left[\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} - \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right]$$
$$= E_0\hbar \left[\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} - \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \right] = 0\,.$$

A commutes with the Hamiltonian.

$$[H, B] = HB - BH = E_0\hbar \left[\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix} - \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right]$$
$$= E_0\hbar \left[\begin{pmatrix} -i & 1 \\ 1 & i \end{pmatrix} - \begin{pmatrix} i & 1 \\ 1 & -i \end{pmatrix} \right] = 2i\hbar E_0 \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \neq 0\,.$$

A and H form a complete set of observables that commute.

b) To find the evolved version of the initial ket, we will need the eigenvalues and eigenvectors of the Hamiltonian. This are given by

$$\begin{vmatrix} -\lambda & E_0 \\ E_0 & -\lambda \end{vmatrix} = 0 \quad \longrightarrow \quad \lambda^2 - E_0^2 = 0 \quad \longrightarrow \quad E_\pm = \pm E_0\,.$$

The corresponding (normalized) eigenvectors follow straightforwardly from

$$H|+\rangle = E_0|+\rangle\,, \qquad H|-\rangle = -E_0|-\rangle\,, \quad \longrightarrow \quad |+\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad |-\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

In terms of $|\pm\rangle$, the initial state $|\varphi\rangle$ takes the form

$$|\varphi\rangle = \frac{1}{\sqrt{2}}|+\rangle + \frac{1}{\sqrt{2}}|-\rangle\,.$$

At time t, we have

$$|\varphi(t)\rangle = \frac{1}{\sqrt{2}}e^{-iE_0t/\hbar}|+\rangle + \frac{1}{\sqrt{2}}e^{iE_0t/\hbar}|-\rangle = \frac{1}{2}\begin{pmatrix} e^{-iE_0t/\hbar} + e^{iE_0t/\hbar} \\ e^{-iE_0t/\hbar} - e^{iE_0t/\hbar} \end{pmatrix} = \begin{pmatrix} \cos\left(\frac{E_0t}{\hbar}\right) \\ -i\sin\left(\frac{E_0t}{\hbar}\right) \end{pmatrix}.$$

c) Since A and H form a CSCO, they have simultaneous eigenvectors. Therefore, the A eigenvalues can be easily obtained by Acting with this operator on $|\pm\rangle$ states,

$$A|+\rangle = \hbar \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 1-1 \\ -1+1 \end{pmatrix} = 0 = 0|+\rangle$$

$$A|-\rangle = \hbar \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix} = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 2 \\ -2 \end{pmatrix} = 2\hbar|-\rangle\,.$$

The possible results of measuring A on the state are 0 and $2\hbar$. The probabilities of getting each result are given by

$$\text{Prob}(0) = |\langle+|\varphi,t\rangle|^2 = \left| \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} \cos\left(\frac{E_0 t}{\hbar}\right) \\ -i\sin\left(\frac{E_0 t}{\hbar}\right) \end{pmatrix} \right|^2 = \frac{1}{2}\left|e^{-iE_0 t/\hbar}\right|^2 = \frac{1}{2}\,.$$

and

$$\text{Prob}(2\hbar) = |\langle-|\varphi,t\rangle|^2 = \left| \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \begin{pmatrix} \cos\left(\frac{E_0 t}{\hbar}\right) \\ -i\sin\left(\frac{E_0 t}{\hbar}\right) \end{pmatrix} \right|^2 = \frac{1}{2}\left|e^{iE_0 t/\hbar}\right|^2 = \frac{1}{2}\,.$$

d) If the measurement at time T gave the larger A eigenvalue $2\hbar$), the state of the system collapses to the vector $|\varphi(T)\rangle = |-\rangle$ corresponding to this eigenvalue. Since this is already an eigenvector of the Hamiltonian, the time dependence can be directly added by matching the condition at time $t = T$, namely

$$|\varphi(t > T)\rangle = e^{iE_0(t-T)/\hbar}|-\rangle\,.$$

The only possible result at $t = 3T$ is $-E_0$,

$$\text{Prob}\,(E_0) = 0\,, \qquad \text{Prob}\,(-E_0) = 1\,.$$

More on probabilities

An Hermitian operator Q representing a physical observable for a quantum system with Hilbert space $\mathcal{H}$ and orthonormal basis $\{|1\rangle, |2\rangle\}$ is defined as $Q|1\rangle = 2|1\rangle - 2i|2\rangle$ and $Q|2\rangle = 2i|1\rangle - |2\rangle$. Which are the possible outcomes of measuring Q if the system is initially prepared in the state $|\varphi\rangle = \frac{1}{\sqrt{3}}|1\rangle + \frac{1+i}{\sqrt{3}}|2\rangle$. And the associated probabilities? Which are the states of the system after the measurement?

In matrix representation the Q operator takes the form

$$Q = \begin{pmatrix} 2 & 2i \\ -2i & -1 \end{pmatrix}$$

The corresponding eigenvalues and eigenvectors are $q_1 = -2, q_2 = 3$ and

$$|q_1\rangle = \frac{1}{\sqrt{5}}(|1\rangle + 2i|2\rangle)\,, \qquad |q_2\rangle = \frac{1}{\sqrt{5}}(2|1\rangle - i|2\rangle)\,.$$

Therefore the possible results of a measurement are -2 or 3. If the result $q_1 = -2$ is obtained, then the state of the system immediately after the measurement is performed is $|q_1\rangle$, and if the result $q_2 = 3$

is obtained, then the state of the system immediately after the measurement is performed is $|q_2\rangle$. The probability of obtaining $q_1 = -2$ is

$$|\langle q_1|\varphi\rangle|^2 = \frac{1}{3}|\langle q_1|1\rangle + (1+i)\langle q_1|2\rangle|^2 = \frac{1}{3}\left|\frac{1}{\sqrt{5}} - \frac{2i}{\sqrt{5}}(1+i)\right|^2 = \frac{1}{15}|1-2i+2|^2 = \frac{13}{15}\,.$$

The probability of obtaining the other result can be calculated either by simply subtracting the previous result from unity or by explicit computation, namely

$$|\langle q_2|\varphi\rangle|^2 = \frac{1}{3}\left|\langle q_2|1\rangle + (1+i)\,\langle q_2|2\rangle\right|^2 = \frac{1}{3}\left|\frac{2}{\sqrt{5}} + \frac{i}{\sqrt{5}}(1+i)\right|^2 = \frac{1}{15}|2+i-1|^2 = \frac{2}{15}\,.$$

Spin precession in an oscillating magnetic field

Consider an electron initially prepared in a spin-up state with respect to the x-axis, $|\varphi, 0\rangle = |x\rangle_x$, and placed at rest in a uniform but oscillating magnetic field $\vec{B} = B_0\cos(\omega t)\,\vec{e}_z$, with B_0 and ω constants and $\vec{e}_z$ a unit vector along the z-direction.

a) Write down the associated Hamiltonian and Schrödinger equation.

b) Determine the unitary operator $U(t)$ leading to time evolution.

c) Calculate the evolution of the state in terms of the time-dependent angles $\theta(t)$ and $\phi(t)$ defining the instantaneous direction of the spin.

d) Find the probability of getting $-\hbar/2$ if S_x is measured at time t.

e) Which is the minimum field amplitude required to force a complete flip in S_x?

a) The Hamiltonian is given by

$$H = -\gamma\vec{B}\cdot\vec{S} = -\gamma B_0\cos\omega t\;\mathrm{S}_z = -\frac{\gamma B_0\hbar}{2}\cos\omega t\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}.$$

Denoting the state at time t as

$$|\varphi, t\rangle = \begin{pmatrix}\alpha(t)\\ \beta(t)\end{pmatrix},$$

the corresponding Schrödinger equation reads

$$i\hbar\frac{\partial\varphi}{\partial t} = H\varphi = -\frac{\gamma B_0\hbar}{2}\cos\omega t\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = -\frac{\gamma B_0\hbar}{2}\cos\omega t\begin{pmatrix}\alpha\\ -\beta\end{pmatrix},$$

with initial conditions

$$|+\rangle_x = \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle) = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ 1\end{pmatrix} \quad\longrightarrow\quad \alpha(0) = \beta(0) = \frac{1}{\sqrt{2}}\,.$$

b) Since the considered magnetic field is classical, and therefore just a function and not an operator, the Hamiltonians at different times commute with each other

$$[H(t'), H(t'')] \propto [\cos(\omega t'), \cos(\omega t'')] = 0\,,$$

allowing us to write

$$U(t,0) = \exp\left(-\frac{i}{\hbar}\int_0^t H(t')dt'\right) = \exp\left(i\frac{\gamma B_0}{2\omega}\sin(\omega t)\sigma_z\right).$$

c) Taking into account the previous evolution operator and the fact that $(\sigma_z)^n|\pm\rangle = (\pm 1)^n|\pm\rangle$ and

$$\exp\left(i\Omega(t)\sigma_z\right)|\pm\rangle = \sum_{n=0}^{\infty}\frac{1}{n!}\left(i\Omega(t)\right)^n(\sigma_z)^n|\pm\rangle = \sum_{n=0}^{\infty}\frac{1}{n!}\left(\pm i\Omega(t)\right)^n|\pm\rangle = \exp\left(\pm i\Omega(t)\right)\,,$$

with $\Omega(t) \equiv \frac{\gamma B_0}{2\omega}\sin(\omega t)$, we get

$$|\varphi,t\rangle = U(t,0)|\varphi,0\rangle = U(t,0)|+\rangle_x = \frac{1}{\sqrt{2}}\begin{pmatrix} e^{i(\gamma B_0/2\omega)\sin\omega t} \\ e^{-i(\gamma B_0/2\omega)\sin\omega t}\end{pmatrix}.$$

Alternatively, we can solve the system of differential equation we obtained when deterimining the Schrödinger equation. Solving in particular the equation for α, and imposing the initial condition $\alpha(0) = \frac{1}{\sqrt{2}}$, one gets

$$\dot{\alpha} = i\left(\frac{\gamma B_0}{2}\right)\cos\omega t\,\alpha \quad\longrightarrow\quad \ln\alpha = \frac{i\gamma B_0}{2}\frac{\sin\omega t}{\omega} + \text{ constant} \quad\longrightarrow\quad \alpha(t) = \frac{1}{\sqrt{2}}e^{i(\gamma B_0/2\omega)\sin\omega t}\,.$$

Solving now for β, we have

$$\dot{\beta} = -i\left(\frac{\gamma B_0}{2}\right)\cos\omega t\beta \quad\longrightarrow\quad \beta(t) = \frac{1}{\sqrt{2}}e^{-i(\gamma B_0/2\omega)\sin\omega t}\,, \quad\longrightarrow\quad |\varphi,t\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} e^{i(\gamma B_0/2\omega)\sin\omega t} \\ e^{-i(\gamma B_0/2\omega)\sin\omega t}\end{pmatrix},$$

in agreement with the previous result.

d) The probability amplitude of getting $-\hbar/2$ if S_x is measured at time t is given by $P = |{}_+\langle -|\varphi,t\rangle|^2$ with

$${}_+\langle -|\varphi,t\rangle = \frac{1}{2}\left[e^{i(\gamma B_0/2\omega)\sin\omega t} - e^{-i(\gamma B_0/2\omega)\sin\omega t}\right] = i\sin\left[\frac{\gamma B_0}{2\omega}\sin\omega t\right].$$

Therefore, one has

$$P = |{}_+\langle -|\varphi,t\rangle|^2 = \sin^2\left[\frac{\gamma B_0}{2\omega}\sin\omega t\right].$$

e) The argument of $\sin^2$ must reach $\pi/2$ (so $P = 1$). This requires

$$\frac{\gamma B_0}{2\omega} = \frac{\pi}{2}\,, \qquad \text{or} \qquad B_0 = \frac{\pi\omega}{\gamma}\,.$$

An electron in a two-component magnetic field

A spin-1/2 particle is subject to a constant magnetic field $\vec{B}_3 = (0,0,B_3)$. At a given time t_0, a second magnetic field $\vec{B}_1 = (B_1,0,0)$ is introduced for a duration τ. During the interval from t_0 to τ, the spin will precess about the axis defined by the combined fields $\vec{B}_1 + \vec{B}_3$. Assuming the particle spin to be oriented along the positive z-axis at $t = t_0$, determine a) the angular frequency of the precession and b) the value of τ for which the expectation value of the x-component S_x of the spin reaches its maximum.

a) The angular frequency is $\omega = \gamma B = \gamma\sqrt{B_1^2 + B_3^2}$ with γ the gyromagnetic ratio and $B = |\vec{B}| = \left|\vec{B}_1 + \vec{B}_3\right| = \sqrt{B_1^2 + B_3^2}$

b) Setting $t_0 = 0$ for simplicity, the solution to the Schrödigner equation

$$i\hbar\frac{d}{dt}|\varphi\rangle = H|\varphi\rangle\,, \qquad |\varphi,0\rangle = |z_+\rangle = \begin{pmatrix}1\\0\end{pmatrix}, \qquad H = -\vec{\mu}_s\vec{B} = -\gamma\frac{\hbar}{2}\vec{\sigma}\cdot\vec{B} = -\gamma\frac{\hbar}{2}\begin{pmatrix} B_3 & B_1 \\ B_1 & -B_3\end{pmatrix},$$

in the interval $0 \le t \le \tau$ is given by

$$|\varphi,t\rangle = \exp\left[-\frac{i}{\hbar}Ht\right]|\varphi,0\rangle\,, \qquad e^{-iHt/\hbar} = \sum_{n=0}^{\infty}\frac{1}{n!}\left(-\frac{it}{\hbar}\right)^n H^n = \sum_{n=0}^{\infty}\frac{1}{n!}\left(\frac{it\gamma}{2}\right)^n(\vec{\sigma}\cdot\vec{B})^n\,.$$

Let us compute the different power of $\vec{\sigma}\cdot\vec{B}$,

$$\begin{aligned}
&\vec{\sigma}\cdot\vec{B} = B_1\sigma_1 + B_3\sigma_3\,,\\
&(\vec{\sigma}\cdot\vec{B})^2 = B_1^2\underbrace{\sigma_1^2}_{1} + B_3\underbrace{\sigma_3^2}_{1} + B_1B_3\underbrace{(\sigma_1+\sigma_3)}_{0} = \left(B_1^2+B_3^2\right)\mathbb{I}\,,\\
&(\vec{\sigma}\cdot\vec{B})^3 = \left(B_1^2+B_3^2\right)\vec{\sigma}\cdot\vec{B}\,,\\
&(\vec{\sigma}\cdot\vec{B})^{2k} = \left(B_1^2+B_3^2\right)^k\mathbb{I}\,,\\
&(\vec{\sigma}\cdot\vec{B})^{2k+1} = \left(B_1^2+B_3^2\right)^k\vec{\sigma}\cdot\vec{B}\,.
\end{aligned}$$

With this, we have

$$\begin{aligned}
e^{-iHt/\hbar} &= \sum_{k=0}^{\infty}\frac{(-1)^k}{(2k)!}\left(\frac{t\gamma}{2}\right)^{2k}\left(B^2+B_3^2\right)^k\mathbb{I} + \sum_{k=0}^{\infty}\frac{(-1)^k}{(2k+1)!}\left(\frac{ty}{2}\right)^{2k+1}\left(B_1^2+B_3^2\right)^k i\vec{\sigma}\cdot\vec{B}\\
&= \cos[\underbrace{\frac{\gamma}{2}\sqrt{B_1^2+B_3^2}}_{\frac{\omega}{2}}t] + \underbrace{\frac{1}{\sqrt{B_1^2+B_3^2}}}_{\gamma/\omega}\sin\left[\frac{1}{2}\sqrt{B_1^2+B_3^2}t\right]i\vec{\sigma}\cdot\vec{B}\\
&= \begin{pmatrix} \cos\frac{\omega t}{2} + i\frac{\gamma B_3}{\omega}\sin\frac{\omega t}{2} & i\gamma\frac{B_1}{\omega}\sin\frac{\omega t}{2} \\ i\gamma\frac{B_1}{\omega}\sin\frac{\omega t}{2} & \cos\omega t - i\frac{\gamma B_3}{2}\sin\frac{\omega t}{2}\end{pmatrix}
\end{aligned}$$

and therefore

$$|\varphi,t\rangle = \exp\left[-\frac{i}{\hbar}Ht\right]|\varphi,0\rangle = \begin{pmatrix}\cos\frac{\omega t}{2} + i\frac{\gamma B_3}{\omega}\sin\frac{\omega t}{2}\\ i\gamma\frac{B_1}{\omega}\sin\frac{\omega t}{2}\end{pmatrix}, \qquad 0\le t\le\tau\,.$$

The expectation value of S_x at $t=\tau$ is then

$$\begin{aligned}
\langle S_x\rangle_{t=\tau} &= \frac{\hbar}{2}\langle\varphi(\tau)|\sigma_x|\varphi(\tau)\rangle\\
&= \frac{\hbar}{2}\left(\cos\frac{\omega\tau}{2} - i\frac{\gamma_3}{\omega}\sin\frac{\omega\tau}{2}, -i\frac{\beta_1}{\omega}\sin\frac{\omega\tau}{2}\right)\begin{pmatrix}0&1\\1&0\end{pmatrix}\begin{pmatrix}\cos\frac{\omega\tau}{2} + \frac{i\gamma\beta_3}{\omega}\sin\frac{\omega\tau}{2}\\ i\frac{\beta_1}{\omega}\sin\frac{\omega\tau}{2}\end{pmatrix}\\
&= \frac{\hbar}{2}\left[\left(\cos\frac{\omega\pi}{2} - i\frac{B_3}{\omega}\sin\frac{\omega\tau}{2}\right)i\frac{\gamma^B}{\omega}\sin\frac{\omega\tau}{2} - i\frac{\gamma B_1}{\omega}\sin\frac{\omega\tau}{2}\left(\cos\frac{\omega t}{2} + \frac{\gamma\gamma\beta_3}{\omega}\sin\frac{\omega\tau}{2}\right)\right]\\
&= \frac{\hbar\gamma^2B_1B_3}{\omega^2}\sin^2\frac{\omega\tau}{2}\,.
\end{aligned}$$

The maximum of this quantity occurs at the maxima of $\sin^2\omega\tau$, namely

$$\sin^2\frac{\omega\tau}{2} = 1 \quad\longrightarrow\quad \frac{\omega\tau}{2} = \left(n+\frac{1}{2}\right)\pi\,, \quad n = 0,1,2,\ldots$$

or

$$\tau = \frac{(2n+1)\pi}{\omega} = \frac{(2n+1)\pi}{\gamma\sqrt{B_1^2+B_3^2}}, \quad n = 0,1,2,\ldots$$

3.6 Exercises

1. **The Tight-Binding Model**

Consider an electron moving in a one-dimensional chain of N atoms, equally spaced by distance a. The electron can only occupy discrete sites $|n\rangle$, where $n = 1, \ldots, N$, and sites are connected in a loop ($|N+1\rangle = |1\rangle$). The system is governed by the Hamiltonian:

$$H = E_0 \sum_{n=1}^{N} |n\rangle\langle n| - \alpha \sum_{n=1}^{N} \left(|n\rangle\langle n+1| + |n+1\rangle\langle n| \right),$$

where E_0 is the on-site energy and $\alpha > 0$ controls hopping between neighbouring sites.

a) Show that the state $|\varphi_k\rangle = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} e^{ikna} |n\rangle$ is an eigenstate of H, and find its energy $E(k)$.

b) Imposing $\varphi_{N+1} = \varphi_1$, find the allowed values of k. How many distinct values are there?

c) For $a = 1$, $E_0 = 2$, $\alpha = 1$, $N = 100$, plot $E(k)$. Where are the minimum and maximum energies?

d) Explain why adding the hopping term causes the electron's eigenstates to spread out over the entire lattice, even for very small α.

e) Show that for small k, $E(k) \approx (E_0 - 2\alpha) + \alpha a^2 k^2$. Compare this with $E = \frac{\hbar^2 k^2}{2m}$ for a free particle, and find the effective mass m^* of the electron in the lattice.

f) Determine the total range of energies in the spectrum. How does it depend on α? What does this range represent?

g) Suppose the Hamiltonian includes next-nearest neighbour hopping:

$$H' = H - \alpha' \sum_{n} \left(|n\rangle\langle n+2| + |n+2\rangle\langle n| \right).$$

Compute the new $E(k)$. How does the energy spectrum's shape change?

2. **Leibniz-type identities and the Baker-Campbell-Hausdorff formula**

Consider two operators A and B such that both commute with their commutator, that is, $[A, [A, B]] = [B, [A, B]] = 0$. Let $\lambda \in \mathbb{C}$.

a) Show that the commutator between A and B^n, for any positive integer n, satisfies $[A, B^n] = nB^{n-1}[A, B]$. Use this result to prove that $[A, e^B] = e^B [A, B]$.

b) Define the operator-valued function $F(\lambda) = e^{\lambda A} e^{\lambda B} e^{-\lambda(A+B)}$. Show that it satisfies the differential equation $dF(\lambda)/d\lambda = \lambda [A, B] F(\lambda)$, and deduce from this that $e^A e^B = e^{A+B+\frac{1}{2}[A,B]} = e^B e^A e^{[A,B]}$.

c) Suppose now that A and B are arbitrary operators. Show that $d\left(e^{\lambda A} B e^{-\lambda A}\right)/d\lambda = e^{\lambda A}[A, B]e^{-\lambda A}$, and use this to derive the expansion

$$e^A B e^{-A} = B + [A, B] + \frac{1}{2!}[A, [A, B]] + \frac{1}{3!}[A, [A, [A, B]]] + \ldots .$$

3. **Bogoliubov transformations**

Consider annihilation and creation operators a and $a^\dagger$ satisfying the canonical commutation relation $[a, a^\dagger] = 1$.

a) Demonstrate that the Bogoliubov transformation $b = a\cosh\eta + a^\dagger \sinh\eta$ with η a free parameter leads to a new pair of operators b and $b^\dagger$ that obey the same canonical commutation relation $[b, b^\dagger] = 1$.

b) Use the previous result to diagonalize the Hamiltonian $H = \hbar\omega\, a^\dagger a + \lambda\left(aa + a^\dagger a^\dagger\right)$, with ω and λ positive constants. Determine its eigenvalues. *Note that the diagonalization is only possible for values of λ below a certain critical threshold.*

c) Show that the unitary operator $U = \exp\left[\frac{\eta}{2}(aa - a^\dagger a^\dagger)\right]$ implements the Bogoliubov transformation via $b = U\, a\, U^{-1}$.

(d) Express the ground state of the Hamiltonian in terms of the number basis $|n\rangle$, defined by $a^\dagger a\,|n\rangle = n\,|n\rangle$.

4. **Hermitian operators and commutation relations**

Consider a Hermitian operator A satisfying the commutation relation $[A, B] = i\hbar\,\mathbb{I}$. Is B necessarily Hermitian? Is it possible to construct finite dimensional matrices A and B satisfying this commutation relation?

5. **A fermionic oscillator**

Consider an operator A satisfying the relations $AA^\dagger + A^\dagger A = \mathbb{I}$ and $A^2 = \left(A^\dagger\right)^2 = 0$.

a) Can A be Hermitian?

b) Show that the only possible eigenvalues of the operator $H = A^\dagger A$ are either 0 or 1.

c) If a state $|0\rangle$ satisfies $H|0\rangle = 0$ with $\langle 0|0\rangle = 1$, determine the states $|a\rangle \equiv A|0\rangle$ and $|b\rangle \equiv A^\dagger|0\rangle$.

6. **Pauli matrices as generators**

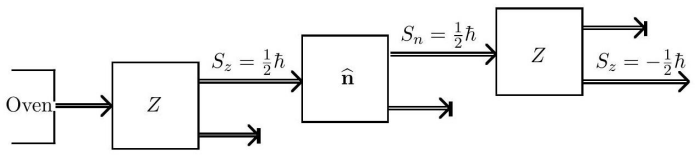
Show that

$$e^{i\frac{\phi}{2}\sigma_x}\sigma_z e^{-i\frac{\phi}{2}\sigma_x} = \sigma_z\cos\phi + \sigma_y\sin\phi\,,$$

with σ_x, σ_y and σ_z the Pauli matrices.

7. **Stern-Gerlach experiments and probabilities**

Consider a beam of spin 1/2 particles passing through a series of three Stern-Gerlach devices, as illustrated in the figure.



The first device filters out particles with $S_z = -\hbar/2$. The remaining $S_z = \hbar/2$ component proceeds them towards a second device, whose magnetic field is oriented in a direction $\vec{n}$ in the xz plane, making an angle θ with the z axis. This device transmits particles with $S_{\vec{n}} = \vec{S}\cdot\vec{n} = \hbar/2$, filtering out those with $S_{\vec{n}} = -\hbar/2$. Finally, the last device transmits only particles with $S_z = -\hbar/2$.

a) Determine the fraction of particles going through the first Stern-Gerlach device that survive to the third measurement. How must the second-device magnetic field be oriented in order to maximize the number of particles transmitted by the final device? Which fraction of particles survive for this value of θ?

b) What fraction of the particles survive the last measurement if the second Stern-Gerlach device is removed?

8. **A toy-model of the Young's double-slit experiment**

Consider a quantum particle that can reach a detection screen by traveling along two or more possible paths. To model its dynamics, we introduce three observables A, B, and C, each with a non-degenerate spectrum. The observable A corresponds to the preparation of the system: the particle is initially in the eigenstate $|a\rangle$, representing emission from the source. The observable B, with spectrum $\sigma_p(B)$ of eigenstates $\{|b\rangle\}$, encodes which-path information, identifying through which slit the particle passes. Finally, the observable C, with spectrum $\sigma_p(C)$ of eigenstates $\{|c\rangle\}$, represents the measurement performed at the detection screen. To ensure the presence of quantum interference and measurement incompatibility, we assume $[A, B] \neq 0, \quad [B, C] \neq 0, \quad [A, C] \neq 0$.

a) Suppose a measurement of B is performed after preparation in $|a\rangle$, followed by a measurement of C. Show that the probability of obtaining outcome $c \in \sigma_p(C)$ is

$$\text{Prob}(a \to b \to c) = \sum_{b \in \sigma_p(B)} |\langle c|b\rangle|^2 |\langle b|a\rangle|^2 .$$

b) If no measurement of B is performed, and the system evolves coherently from $|a\rangle$ to a measurement of C, show that the probability is

$$\text{Prob}(a \to c) = \Big| \sum_{b \in \sigma_p(B)} \langle c|b\rangle \langle b|a\rangle \Big|^2 .$$

c) Expand the last expression and relate it to the result in part a a). Identify the interference terms and discuss the conditions under which they vanish.

9. **More on probabilities**

An Hermitian operator Q representing a physical observable for a quantum system with Hilbert space $\mathcal{H}$ and orthonormal basis $\{|1\rangle, |2\rangle\}$ is defined as $Q|1\rangle = 2|1\rangle - 2i|2\rangle$ and $Q|2\rangle = 2i|1\rangle - |2\rangle$. Which are the possible outcomes of measuring Q if the system is initially prepared in the state $|\varphi\rangle = \frac{1}{\sqrt{3}}|1\rangle + \frac{1+i}{\sqrt{3}}|2\rangle$. And the associated probabilities? Which are the states of the system after the measurement?

10. **Compatibility of anticonmuting observables**

Can two Hermitian operators satisfying the anticommutation relation $\{A, B\} = AB + BA = 0$ be compatible? Illustrate the answer with an example.

11. **Compatibility of observables and spectra**

Consider three observables H, A, B.

a) Prove that if $[H, A] = [H, B] = 0$, but $[A, B] \neq 0$, the spectrum of H is degenerate. Can some of the eigenvalues of H be non-degenerate?

b) Given

$$H = \begin{pmatrix} n_1 & 0 & 0 \\ 0 & n_2 & 0 \\ 0 & 0 & n_3 \end{pmatrix}, \qquad n_1 = n_2 \neq n_3 ,$$

in some orthogonal basis of a three-dimensional Hilbert space, construct two Hermitian operators A and B such that $[H, A] = [H, B] = 0$, but $[A, B] \neq 0$. Determine the corresponding eigenvalues and eigenvectors. Find a basis with respect to which your operators H and A are simultaneously diagonal.

12. **Spin uncertainty relation**

Verify explicitly the validity of the relation $(\Delta S_x)^2(\Delta S_y)^2 \geq \frac{1}{4}\langle i[S_x, S_y]\rangle^2 = \frac{\hbar^2}{4}\langle S_z\rangle^2$ for a spin 1/2 particle in the state $|+\rangle$.

13. **The fall of the ice pick**

Estimate the order of magnitude for how long an ice pick could, in principle, remain balanced upright on its tip if the only factor limiting its stability were the Heisenberg uncertainty principle. Assume the tip is perfectly sharp, and that both the ice pick and the surface it rests on are rigid.
You may make reasonable physical approximations (e.g., mass, dimensions), as long as they do not affect the order of magnitude of the result. Use the uncertainty principle to estimate the timescale for the onset of instability, and express your final answer in seconds.

14. **Observables vs. Symmetries**

Determine which of the following operators can represent a symmetry operation, and which can represent an observable: a) $A \equiv |1\rangle\langle 2|$, b)$B \equiv |1\rangle\langle 2| + |2\rangle\langle 3| + |3\rangle\langle 1|$, c) $C \equiv |1\rangle\langle 2| + |2\rangle\langle 1|$.

15. **Unitary operators**

Consider a Hermitian operator A and a unitary operator U. How does the spectra of A compares to that $UAU^\dagger$?

16. **An alternative universe with modified Schrödinger equation**

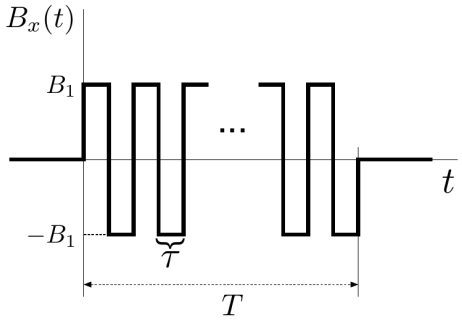
Consider an imaginary universe where quantum mechanics is controlled by a second order Schrödinger equation $i\hbar\frac{\partial^2}{\partial t^2}|\varphi, t\rangle = H|\varphi, t\rangle$. Assuming a time-independent Hamiltonian with discrete eigenvalues E_n, determine the evolution of $|\varphi, t\rangle = \sum_n c_n(t)|E_n\rangle$. Are probabilities conserved?

17. **Coherent states evolution**

Show that a coherent state $|\lambda\rangle = \sum_{n=0}^{\infty} e^{-|\lambda|^2/2}\frac{\lambda^n}{\sqrt{n!}}|n\rangle$ remains coherent when time evolution is given by the harmonic oscillator Hamiltonian. Calculate the time-evolved state $|\lambda, t\rangle$.

18. **An electron in a square-wave magnetic field**

Consider an electron at rest in a magnetic field $\vec{B}$ with constant z-component B_3 and time-dependent x-component $B_x(t)$ vanishing for $t < 0$ and $t > T$ and taking alternating constant values B_1 and $-B_1$ in time intervals τ for $0 \leq t \leq T$, as illustrated in the figure.



If the the spin is initially oriented in the positive z-axis, determine the expectation value of the magnetic moment after N oscillations, i. e. at a time $t = 2\pi N/\omega$, with ω the precession frequency. Calculate the total duration T of the pulse for which the angle between the magnetic moment expectation value at T and the z-axis is maximal. Particularize the result for $B_1 \ll B_3$.

19. **Conserved quantities**
Show that if two observables A and B are conserved, their commutator $[A, B]$ remains constant.

20. **Landau levels in Landau gauge**
Consider a charged particle (such as an electron) moving in three-dimensional space, subject to a uniform magnetic field pointing in the z-direction: $\vec{B} = (0, 0, B)$. We take the electric field to be zero. The quantum dynamics are governed by the Hamiltonian $H = \frac{1}{2m}(\vec{P} - q\vec{A})^2$, where $\vec{A}$ is a vector potential satisfying $\nabla \times \vec{A} = \vec{B}$.

a) Show that in the so-called Landau gauge $\vec{A} = (0, xB, 0)$, the Hamiltonian becomes

$$H = \frac{1}{2m}\left(P_x^2 + (P_y - qBx)^2 + P_z^2\right).$$

b) Because the Hamiltonian is invariant under translations in the y- and z-directions, we can look for solutions to the Schrödinger equation of the form $\varphi(\vec{x}) = e^{ik_y y + ik_z z}\chi(x)$. Substitute this ansatz into the Schrödinger equation and show that $\chi(x)$ must satisfy

$$\left[\frac{1}{2m}P_x^2 + \frac{m\omega_B^2}{2}\left(x - k_y l_B^2\right)^2\right]\chi(x) = \left(E - \frac{\hbar^2 k_z^2}{2m}\right)\chi(x),$$

where $\omega_B = \frac{qB}{m}$ and $l_B = \sqrt{\frac{\hbar}{qB}}$ is the magnetic length.

c) Identify the effective potential acting on the particle in the x-direction. What kind of familiar quantum mechanical system does this resemble?

d) Write down the energy levels of this system. What values can the quantum number n take? What role does k_z play?

e) Take $k_z = 0$ to focus on motion in the x-y plane. Show that the energy levels become

$$E_n = \hbar\omega_B\left(n + \frac{1}{2}\right), \quad n = 0, 1, 2, \ldots$$

These are called *Landau levels*. How does the spacing between levels depend on the magnetic field B?

f) The energy levels do not depend on k_y, yet the wavefunctions do. Explain why this means that each Landau level has a large degeneracy: many different quantum states with the same energy.

g) Suppose that the particle is confined to a rectangular region of size $L_x \times L_y$. Using the fact that $\chi(x)$ is localized near $x = k_y l_B^2$, estimate the allowed range of k_y. Then, derive the total number of distinct quantum states $\mathcal{N}$ that can fit into a single Landau level in this region. Show that $\mathcal{N} = \frac{qBA}{2\pi\hbar}$, where $A = L_x L_y$ is the area. Why is this degeneracy so large in practice?

21. **Landau levels in symmetric gauge**

Reconsider the motion of the charged particle in the previous exercise, but this time using the *symmetric gauge* for the vector potential $\vec{A} = \frac{B}{2}(-y, x, 0)$. This choice respects rotational symmetry in the plane, and the angular momentum about the origin becomes a useful quantum number.

a) Show that in the given gauge the Hamiltonian becomes

$$H = \frac{1}{2m}\left(P_x + \frac{qBy}{2}\right)^2 + \frac{1}{2m}\left(P_y - \frac{qBx}{2}\right)^2 + \frac{1}{2m}P_z^2\,,$$

and verify that it can be rewritten (for $k_z = 0$) as

$$H = -\frac{\hbar^2}{2m}\nabla^2 + \frac{qB}{2m}L_z + \frac{q^2B^2}{8m}(x^2 + y^2),$$

where $L_z = xP_y - yP_x$ is the angular momentum operator.

b) Switch to complex coordinates $w = x + iy$ and $\bar{w} = x - iy$ and define complex derivatives

$$\partial = \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right), \quad \bar{\partial} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right).$$

Verify the identities $\partial w = \bar{\partial}\bar{w} = 1$, and $\partial\bar{w} = \bar{\partial}w = 0$.

c) Show that the Hamiltonian for $k_z = 0$ becomes

$$H = -\frac{2\hbar^2}{m}\partial\bar{\partial} - \frac{\omega_B}{2}L_z + \frac{m\omega_B^2}{8}w\bar{w},$$

and that the angular momentum operator becomes $L_z = \hbar(w\partial - \bar{w}\bar{\partial})$.

d) Consider wavefunctions of the form $\varphi_0(w, \bar{w}) = f(w)e^{-|w|^2/4l_B^2}$, where $f(w)$ is any holomorphic function. Show that these wavefunctions all lie in the lowest Landau level and satisfy $H\varphi_0 = \frac{\hbar\omega_B}{2}\varphi_0$.

e) Show that the monomials $f(w) = w^M$ yield eigenfunctions of L_z with eigenvalue $\hbar M$, and describe the physical significance of the quantum number M.

f) In symmetric gauge, the wavefunction $\varphi_0 = w^M e^{-|w|^2/4l_B^2}$ is concentrated around a ring of radius $r = \sqrt{2M}l_B$. Estimate the number of distinct angular momentum states that can fit inside a circular region of radius R, and show that $\mathcal{N} = \frac{R^2}{2l_B^2} = \frac{qBA}{2\pi\hbar}$, where $A = \pi R^2$ is the area of the region. How does this result compare with the degeneracy you found in Landau gauge?

g) Although the wavefunctions look very different in Landau and symmetric gauges (strips vs. rings), explain why this difference does not imply a physical distinction.

CHAPTER 4

TWO-LEVEL SYSTEMS

Study hard what interests you the most in the most undisciplined, irreverent and original manner possible.

Richard Feynmann in Perfectly Reasonable Deviations from the Beaten Track

The realm of quantum mechanics encompasses several interesting problems that can be simplified and brought down to two-state problems similar to spin 1/2 systems, such as tunneling states in a double-well potential, the ammonia molecule, the Microwave Amplification by Stimulated Emission of Radiation or MASER, the two-kaon problem or neutrino oscillations. To illustrate the amazing things that are already within your reach, we will explore in this chapter some of these scenarios, leaving others as homework exercises. When doing so, it is extremely important that you don't miss the forest for the trees.

4.1 Don't miss the forest for the trees

While the physics involved in the formulation of the specific Hamiltonians we are going to consider can be very diverse and in some cases far beyond your current training, it will be always[1] possible to express the result in a given orthonormal basis $|\varphi_1\rangle, |\varphi_2\rangle$ as a two-by-two Hermitian matrix characterized by four real constants $g_0, g_1, g_2, g_3 \in \mathbb{R}$, [2] namely

$$H = \begin{pmatrix} g_0 + g_3 & g_1 - ig_2 \\ g_1 + ig_2 & g_0 - g_3 \end{pmatrix} = g_0 \mathbb{I} + g_1\sigma_1 + g_2\sigma_2 + g_3\sigma_3 = g_0\mathbb{I} + \vec{g}\cdot\vec{\sigma}\,, \tag{4.1}$$

with σ_i the linearly independent Hermitian Pauli matrices $[i = 1, 2, 3\ (x, y, z)]$, $\vec{\sigma}$ a vector $\sigma = (\sigma_x, \sigma_y, \sigma_z)^T$ and $\vec{g} = g\vec{n}$ a vector with components (g_1, g_2, g_3), magnitude $g = \sqrt{g_1^2 + g_2^2 + g_3^2}$

[1] See, however, Section 4.5.

[2] Remember that the diagonal elements must be real and the off-diagonal elements must be each-other's complex conjugate, so H can be completely identified by 4 independent real numbers.

and direction $\vec{n}$ ($\vec{n}\cdot\vec{n}=1$). The evolution of the system will be then given by a pair of Schrödinger equations,

$$i\hbar\frac{d}{dt}\begin{pmatrix}|\varphi_1\rangle\\|\varphi_2\rangle\end{pmatrix}=H\begin{pmatrix}|\varphi_1\rangle\\|\varphi_2\rangle\end{pmatrix}, \qquad i\hbar\frac{d}{dt}\begin{pmatrix}|\varphi_1\rangle\\|\varphi_2\rangle\end{pmatrix}=\begin{pmatrix}g_0+g_3 & g_1-ig_2\\ g_1+ig_2 & g_0-g_3\end{pmatrix}\begin{pmatrix}|\varphi_1\rangle\\|\varphi_2\rangle\end{pmatrix}, \tag{4.2}$$

which will be generically coupled if $g_1\pm ig_2\neq 0$. For this reason, we will refer to $\{|\psi_1\rangle,|\varphi_2\rangle\}$ as the *coupled basis* in what follows.

As a mathematical problem, solving the above system of coupled differential equations involves i) the diagonalization of the Hamiltonian and ii) the solution of the resulting uncoupled differential equations. The first task provides, through the characteristic equation $\det(H-\varepsilon_i\mathbb{I})=0$, the eigenvalues and normalized eigenvectors of the Hamiltonian, which are given by,

$$\begin{aligned} &H|\Phi_1\rangle=\varepsilon_1|\Phi_1\rangle\,, && H|\Phi_2\rangle=\varepsilon_2|\Phi_2\rangle\,,\\ &\varepsilon_1\equiv g_0+g\,, && \varepsilon_2=g_0-g\,. \end{aligned} \tag{4.3}$$

$$\begin{aligned} |\Phi_1\rangle&=+\sqrt{\frac{1+g_3/g}{2}}|\varphi_1\rangle+\sqrt{\frac{1-g_3/g}{2}}e^{i\tan^{-1}(g_1/g_2)}|\varphi_2\rangle\,,\\ |\Phi_2\rangle&=-\sqrt{\frac{1-g_3/g}{2}}e^{-i\tan^{-1}(g_1/g_2)}|\varphi_1\rangle+\sqrt{\frac{1+g_3/g}{2}}|\varphi_2\rangle\,. \end{aligned} \tag{4.4}$$

Note that defining

$$\tan\phi\equiv\frac{g_1}{g_2}\,, \qquad \cos\theta\equiv\frac{g_3}{g}\,, \qquad \sin^2\theta=1-\frac{g_3^2}{g^2}\,, \qquad \tan^2\theta\equiv\frac{g^2-g_3^2}{g_3^2}\,, \tag{4.5}$$

and moving the half-angle variables

$$\cos\frac{\theta}{2}=\sqrt{\frac{1+\cos\theta}{2}}=\sqrt{\frac{1+g_3/g}{2}}\,, \qquad \sin\frac{\theta}{2}=\sqrt{\frac{1-\cos\theta}{2}}=\sqrt{\frac{1-g_3/g}{2}}\,, \tag{4.6}$$

it is possible to rewrite these eigenvectors in a way matching the general form of vectors in the Bloch sphere.

$$|\Phi_1\rangle=\cos\frac{\theta}{2}|\varphi_1\rangle+\sin\frac{\theta}{2}e^{i\phi}|\varphi_2\rangle\,, \qquad |\Phi_2\rangle=-\sin\frac{\theta}{2}e^{-i\phi}|\varphi_1\rangle+\cos\frac{\theta}{2}|\varphi_2\rangle\,. \tag{4.7}$$

In this new basis, the Schrödinger equations in (4.2) effectively decouple,

$$i\hbar\frac{d}{dt}\begin{pmatrix}|\Phi_1\rangle\\|\Phi_2\rangle\end{pmatrix}=\begin{pmatrix}g_0+g & 0\\ 0 & g_0-g\end{pmatrix}\begin{pmatrix}|\Phi_1\rangle\\|\Phi_2\rangle\end{pmatrix}, \tag{4.8}$$

allowing us to determine the evolution of *uncoupled basis* states $|\Phi_1\rangle$ and $|\Phi_2\rangle$ using the standard procedure for time-independent Hamiltonians,

$$|\Phi_1,t\rangle=e^{-i\varepsilon_1 t/\hbar}|\Phi_1,0\rangle\,, \qquad |\Phi_2,t\rangle=e^{-i\varepsilon_2 t/\hbar}|\Phi_2,0\rangle\,, \tag{4.9}$$

and with it the evolved state $|\Phi,t\rangle$ of any given initial state $|\Phi,0\rangle$ that could be rewritten in terms of them. As an example, let us assume that the system is initially prepared in the state $|\Phi,0\rangle=|\varphi_1\rangle$ and determine the probability of finding it in the same state at time t, namely $P_1(t)=|\langle\varphi_1|\Phi,t\rangle|^2$. Inverting the relations in (4.7), we have

$$|\Phi,0\rangle=|\varphi_1\rangle=\cos\frac{\theta}{2}|\Phi_1,0\rangle-\sin\frac{\theta}{2}e^{i\phi}|\Phi_2,0\rangle\,, \tag{4.10}$$

and therefore

$$|\Phi, t\rangle = e^{-i\varepsilon_1 t/\hbar} \cos\frac{\theta}{2}|\Phi_1, 0\rangle - e^{-i\varepsilon_2 t/\hbar} \sin\frac{\theta}{2} e^{i\phi}|\Phi_2, 0\rangle\,. \tag{4.11}$$

Factoring out the g_0 part in $\varepsilon_{1,2}$,

$$|\Phi, t\rangle = e^{-ig_0 t/\hbar}\left[e^{-igt/\hbar}\cos\frac{\theta}{2}|\Phi_1, 0\rangle - e^{igt/\hbar}\sin\frac{\theta}{2}e^{i\phi}|\Phi_2, 0\rangle\right], \tag{4.12}$$

and computing the overlap with $\langle\varphi_1|$,

$$\begin{aligned}\langle\varphi_1|\Phi, t\rangle &= e^{-ig_0 t/\hbar}\left[\cos^2\frac{\theta}{2}e^{-igt/\hbar} + \sin^2\frac{\theta}{2}e^{igt/\hbar}\right] = e^{-ig_0 t/\hbar}\left[\frac{1+g_3/g}{2}e^{-igt/\hbar} + \frac{1-g_3/g}{2}e^{igt/\hbar}\right] \\ &= e^{-ig_0 t/\hbar}\left[\frac{e^{igt/\hbar}+e^{-igt/\hbar}}{2} - i\frac{g_3}{g}\frac{e^{igt/\hbar}-e^{-igt/\hbar}}{2i}\right] = e^{-ig_0 t/\hbar}\left[\cos\frac{gt}{\hbar} - i\frac{g_3}{g}\sin\frac{gt}{\hbar}\right],\end{aligned}$$

we get

$$P_1(t) = |\langle\varphi_1|\Phi, t\rangle|^2 = \cos^2\frac{gt}{\hbar} + \frac{g_3^2}{g^2}\sin^2\frac{gt}{\hbar} = 1 - \left(1 - \frac{g_3^2}{g^2}\right)\sin^2\frac{gt}{\hbar}\,. \tag{4.13}$$

Note that the resulting probability is periodic, with the system returning to its initial state after a time interval $\pi\hbar/g$. The specifics of the evolution are governed by the dimensionless ratio $g_3^2/(g_1^2+g_2^2)$. For a fixed mixing amplitude $\sqrt{g_1^2+g_2^2}$ and $g_3 = 0$, the oscillations display their maximum unit amplitude and the lowest frequency. This behavior remains accurate as long as $g_3^2/(g_1^2+g_2^2) \ll 1$. As g_3 increases, the amplitude of the oscillations decreases, while the frequency rises. In the regime where $g_3^2 \gg g_1^2+g_2^2$, the system transitions between states much more rapidly than when $g_3^2 \lesssim g_1^2+g_2^2$, but the transition probability is small.

The precession analogy

As we have seen in the previous chapter, the spin states $|\pm\rangle_n$ are eigenstates of $\vec{n}\cdot\vec{\sigma}$. Recalling that the $\pm$ signs are just a label for the first and second states of a basis where $\vec{n}\cdot\vec{\sigma}$ looks diagonal, we could have easily anticipated that the spectrum of the general two-level Hamiltonian (4.1) will take the form

$$H|\pm\rangle_n = (g_0 \pm g)\,|\pm\rangle_n\,, \tag{4.14}$$

with the states $|\pm\rangle_n$ understood now as *general* states not necessarily having to do with spin. Continuing with this analogy, we could also determine the time evolution of general two-level states by simply rewriting the associated Hamiltonian H in terms of the spin operators $\vec{S} = \frac{\hbar}{2}\vec{\sigma}$,

$$H = g_0\,\mathbb{I} + \frac{2}{\hbar}\vec{g}\cdot\vec{S} \equiv g_0\mathbb{I} + \vec{\omega}\cdot\vec{S}\,. \tag{4.15}$$

The $g_0\,\mathbb{I}$ component in this expression multiplies the states by a time-dependent phase $\exp(-ig_0 t/\hbar)$. On the other hand, the second term allows to identify the equivalent

of the Larmor frequency determining the precession of the states, namely

$$\vec{\omega} = \frac{2}{\hbar}\vec{g}\,. \tag{4.16}$$

4.2 The quantum Zeno effect

In quantum mechanics, the act of measurement has a profound effect on the system being observed. In this section, we will illustrate the so-called *Quantum Zeno Effect* or *Turing paradox*, a quantum mechanics phenomenom where the evolution of a quantum system is inhibited or slowed down by frequent measurements. It is named after one Zeno's four famous paradoxes, particularly the paradox that suggests a moving arrow appears to be at rest when observed at every instant in time. First theorized by Beskow and Nilsson in the 1960s and later formalized by Misra and Sudarshan in 1977, it was experimentally confirmed by Itano, Heinzen, Bollinger and Wineland in 1990.

Consider a simple two-level system with effective Hamiltonian

$$H = \epsilon\sigma_z\,, \tag{4.17}$$

and initial state

$$|\varphi, 0\rangle = \frac{1}{\sqrt{2}}\left(|G\rangle + |E\rangle\right)\,, \tag{4.18}$$

where ϵ is a positive constant with the dimensions of energy, $|G\rangle$ is the ground state of the syatem and $|E\rangle$ its first excited state,

$$E_+ = \epsilon\,, \quad |E\rangle = \begin{pmatrix}1\\0\end{pmatrix}\,, \qquad E_- = -\epsilon\,, \quad |G\rangle = \begin{pmatrix}0\\1\end{pmatrix}\,. \tag{4.19}$$

Note that the state (4.18) is not only a superposition state of the eigenvectors of H, but also an eigenstate of the spin operator σ_x with eigenvalue $+1$,[3]

$$\sigma_x|\varphi,0\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}0&1\\1&0\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} = |\varphi,0\rangle\,.$$

Let us consider a situation where this quantity is repeatedly measured at times $t = \Delta t,\, 2\Delta t, \ldots, N\Delta t$ and determine the probability $P_N(\Delta t)$ of obtaining $+1$ in all performed measurements. Taking into account that the Hamiltonian is independent of time to determine the evolved state of the system at time t,

$$|\varphi, t\rangle = U(t,0)|\varphi,0\rangle = \frac{1}{\sqrt{2}}\left(e^{-i\epsilon t/\hbar}\,|E\rangle + e^{i\epsilon t/\hbar}\,|G\rangle\right) = \frac{1}{\sqrt{2}}\begin{pmatrix}e^{-i\epsilon t/\hbar}\\ e^{i\epsilon t/\hbar}\end{pmatrix}\,,$$

and recalling the σ_x eigenvectors in the z basis,

$$|+\rangle_x = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}\,, \qquad |-\rangle_x = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}\,,$$

[3]The σ_x operator is indeed a paritity operator.

the probability of obtaining $+1$ for σ_x at time Δt is given by

$$P\left(\sigma_x; +1; \Delta t\right) = |_x \langle +|\varphi, \Delta t\rangle|^2 = \left|\frac{1}{2}\left(e^{-i\epsilon\Delta t/\hbar} + e^{i\epsilon\Delta t/\hbar}\right)\right|^2 = \cos^2\left(\frac{\epsilon\Delta t}{\hbar}\right) .$$

After this first measurement, the system is in a state $|\varphi, \Delta t\rangle = |+\rangle_x$. Evolving this state until the time of the second measurement at $t = 2\Delta t$, we get

$$|\varphi, 2\Delta t\rangle = U(2\Delta t, \Delta t)|\varphi(\Delta t)\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} e^{-i\epsilon\Delta t/\hbar} \\ e^{i\epsilon\Delta t/\hbar}\end{pmatrix} .$$

The probability of obtaining again $+1$ for σ_x at $t = 2\Delta t$ is therefore

$$P\left(\sigma_x, +1, +1; 2\Delta t\right) = |_x \langle +|\varphi, \Delta t\rangle|^2 \cdot |_x \langle +|\varphi, 2\Delta t\rangle|^2 = \left[\cos^2\left(\frac{\epsilon\Delta t}{\hbar}\right)\right]^2 .$$

After this second measurement, the system is again in a state $|\varphi, \Delta t\rangle = |+\rangle_x$, which can be further evolved until the time of the next measurement at $t = 3\Delta t$. Iterating the process N times, the probability of obtaining $+1$ in all N measurements becomes

$$P\left(\sigma_x; +1, +1, \ldots, +1; N\Delta t\right) = \left[\cos^2\left(\frac{\epsilon\Delta t}{\hbar}\right)\right]^N = \left[\cos^2\left(\frac{\epsilon T}{\hbar N}\right)\right]^N ,$$

where in the last step we have replaced $\Delta t = T/N$, with T the total observation time, which we take to be finite. Taylor expanding this expression for short times $\Delta t \to 0$ and using the relation,

$$\left(1 - \frac{x^2}{n^2}\right)^n \simeq e^{-x^2/n} ,$$

we obtain

$$P\left(\sigma_x; +1, +1, \ldots, +1; N\Delta t\right) \simeq \left[1 - \left(\frac{\epsilon\Delta t}{\hbar}\right)^2 + O\left(\Delta t^4\right)\right]^N \simeq \left[1 - \left(\frac{\epsilon T}{\hbar N}\right)^2\right]^N \simeq \exp\left(-\frac{\epsilon^2 T^2}{\hbar^2 N}\right) ,$$

which is always smaller than one for any positive and finite T and N. For infinite N, however, the probability approaches 1,

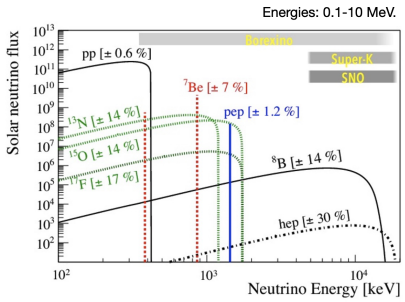
$$P\left(\sigma_x; \ +1, +1, \ldots, +1; N\Delta t\right) \simeq \left[1 - \left(\frac{\epsilon T}{\hbar N}\right)^2\right]^N = \left(1 - \frac{\epsilon T}{\hbar N}\right)^N \left(1 + \frac{\epsilon T}{\hbar N}\right)^N \simeq 1 .$$

Therefore, in the limit of infinite instantaneous measurements, the system is constantly forced to "reset" to the measured state, preventing it from evolving naturally according to the Schrödinger equation.

4.3 The solar neutrino problem

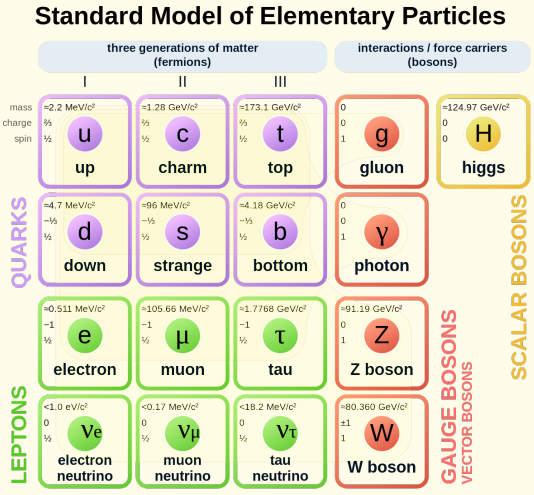
Nuclear reactions in the core of the Sun release an immense amount of energy in the form of light and other fundamental particles, including neutrinos. These *solar neutrinos* stream out from the Sun almost freely, providing a direct window into the interior nuclear activity.

$$p + p \rightarrow d + e^+ + \nu_e$$
$$^8B \rightarrow ^8Be^* + e^+ + \nu_e$$
$$p + e^- + p \rightarrow d + \nu_e$$
$$^7Be + e^- \rightarrow ^7Li + \nu_e$$
$$^3He + p \rightarrow ^4He + e^+ + \nu_e$$



Neutrinos in the Standard Model of particle physics

The Standard Model of particle physics is a fundamental theoretical framework that describes the electromagnetic, weak, and strong nuclear forces governing the behavior of subatomic particles. It classifies particles into two categories: fermions or half-integer spin particles and bosons or integer spin particles.



Fermions, subdivided in quarks and leptons, are the building blocks of matter. They are organized into three generations or families with increasing mass and complexity. The property distinguishing different generations of leptons is called *flavour*.

- The first generation consists of the lightest and most stable fermions, which form all the matter we encounter in our everyday live. In the quark sector, we have the up quark (u) and the down quark (d), which combine to form protons and neutrons, the building blocks of atomic nuclei. On the lepton side, we find the electron (e) and its associated neutrino (ν_e)
- The second generation includes the charm quark (c) and the strange quark (s). These particles, along with the first-generation quarks, form mesons and other exotic hadrons. On the lepton side, we have the muon (μ) and its associated neutrino (ν_μ).

- The third generation comprises the heaviest and least stable fermions. In the quark family, we encounter the heaviest elementary particle ever discovered, the top quark (t), as well as the bottom or beauty quark (b), which plays a crucial role in B meson decays. On the lepton side, we have the tau (τ) and its associated neutrino (ν_μ), the heaviest of its kind.

Bosons, on the other hand, are the force carriers of the fundamental forces of nature. There are four primary types, each of them associated with a specific force:

- The photon γ is the mediator of the electromagnetic force. It has no electric charge, no mass and travels at the speed of light.
- The W^+, W^- and Z bosons are responsible for the weak nuclear force, governing for instance the nuclear beta decay and neutrino interactions. These bosons are comparatively massive and have short lifetimes, resulting in short-range interactions.
- The gluon g is the force carrier of the strong nuclear force, also known as the strong interaction or the color force. It binds quarks together, forming protons, neutrons, and other hadrons. Gluons themselves carry color charge, making them unique among the Standard Model force carriers. In particular, unlike other bosons, gluons can interact with each other, leading to the phenomenon of color confinement.
- The Higgs boson H is a relatively new addition to the Standard Model of particle physics, where it provides mass to many of its constituents. Its discovery in 2012 validates the Standard Model theoretical framework, offering also the possibility of exploring new avenues beyond it.

Unlike electrons or protons, neutrinos do not have an electric charge, not interacting therefore with the electromagnetic forces responsible for the signals in most particle detectors. On top of that, the neutrinos interact very weakly with matter through the weak nuclear force. As a result, they can pass unaffected through large amounts of material, including the Earth itself.

When the first detectors capable of measuring the solar neutrinos' flux were installed in the 1960's, it was discovered, much to everyone's surprise, that there was a significant discrepancy between the predicted number of electron neutrinos and those detected on Earth. In particular, depending on the specific model under consideration, the measured electron neutrino flux was only around 30% to 50% of the anticipated figure. The issue was particularly perplexing because the various theoretical models of the Sun and the available detection techniques appeared to be rather robust and well-founded. This led to a deep reevaluation of our understanding of neutrinos, challenging in particular the initial assumption that neutrinos have zero mass. As we will demonstrate in what follows, the existence of neutrino masses allows for *neutrino oscillations*, a process thorugh which a neutrino of a given *flavour* transforms into a neutrino of a different *flavour*. This accounts for the deficiency observed in the detection of solar neutrinos.

Aiming to highlight the formal similarities with spin 1/2 systems, we will consider first a simplified scenario involving only two neutrinos flavours ν_e and ν_μ propagating in vacuum. Adapting the general notation in Section 4.1, $|\nu_e\rangle \equiv |\varphi_1\rangle$ and $|\nu_\mu\rangle \equiv |\varphi_2\rangle$, and taking into account that, as hypothesized above, this *flavour states* do not have a definite mass, the mass operator in this basis (which is our toy scenario plays the role of the Hamiltonian) would take the form

$$M = \begin{pmatrix} m_1 & \Delta m \\ \Delta m & m_2 \end{pmatrix} = \frac{1}{2}(m_1+m_2)\,\mathbb{I} + \Delta m\,\sigma_1 + \frac{1}{2}(m_1-m_2)\,\sigma_3\,, \tag{4.20}$$

with Δm a real quantity to be fixed experimentally. Comparing this expression with Eq. (4.1), we can easily identify the values of g_0, g_1, g_2 and g_3 and directly determine the eigenvalues and eigenvectors using the expressions in Section 4.1. Alternatively, we can explicitly compute them from the eigenvalue equation

$$\begin{pmatrix} m_1 & \Delta m \\ \Delta m & m_2 \end{pmatrix} \begin{pmatrix} \nu_e \\ \nu_\mu \end{pmatrix} = m \begin{pmatrix} \nu_e \\ \nu_\mu \end{pmatrix}\,. \tag{4.21}$$

In particular, isolating ν_μ from the first equation

$$m_1\nu_e + \Delta m\,\nu_\mu = m\nu_e \qquad \longrightarrow \qquad \nu_\mu = \frac{m-m_1}{\Delta m}\nu_e\,, \tag{4.22}$$

and substituting the result in the second one, we get

$$\Delta m\nu_e + m_2\left(\frac{m-m_1}{\Delta m}\nu_e\right) = m\left(\frac{m-m_1}{\Delta m}\nu_e\right) \qquad \longrightarrow \qquad m^2 - (m_1+m_2)m - \Delta m^2 + m_1 m_2 = 0\,, \tag{4.23}$$

or equivalently

$$m_\pm = \frac{1}{2}\left(m_1+m_2 \pm \sqrt{(m_1-m_2)^2 + 4\Delta m^2}\right)\,. \tag{4.24}$$

Adapting once again the general notation in Section 4.1 to the problem at hand, $|\nu_+\rangle \equiv |\Phi_1\rangle$ and $|\nu_-\rangle \equiv |\Phi_2\rangle$, the corresponding eigenstates are given by[4]

$$|\nu_+\rangle = \cos\theta\,|\nu_e\rangle + \sin\theta\,|\nu_\mu\rangle\,, \qquad\qquad |\nu_-\rangle = -\sin\theta\,|\nu_e\rangle + \cos\theta\,|\nu_\mu\rangle\,, \tag{4.25}$$

with

$$\tan\theta = \frac{\nu_\mu}{\nu_e} = \frac{m-m_1}{\Delta m} \tag{4.26}$$

and θ the so-called mixing angle. Using the double angle formula for the tangent,

$$\tan 2\theta = \frac{2\tan\theta}{1-\tan^2\theta}\,, \tag{4.27}$$

the last expression can be rewritten in terms of the mass difference $m_1 - m_2$,

$$\tan 2\theta = \frac{2\left(\frac{m-m_1}{\Delta m}\right)}{1-\left(\frac{m-m_1}{\Delta m}\right)^2} = 2\Delta m\left(\frac{m-m_1}{\Delta m^2 - (m-m_1)^2}\right) = \frac{2\Delta m}{m_1-m_2}\,, \tag{4.28}$$

[4]Note that, as compared with the general formulas in the previous section, the notation here effectively replaces $\theta \to 2\theta$ in order to match the most standard notation in the neutrino physics community. With this replacement, Eq. (4.4) becomes $\tan 2\theta = \sqrt{g_1^2+g_2^2}/g_3$, which upon comparison with Eq. (4.20) allows to directly find Eq. (4.28).

where in the last step we have taken into account the eigenvalue equation in (4.23), namely

$$\begin{aligned} &m^2-(m_1+m_2)m+m_1m_2=\Delta m^2\,,\\ &m^2-(m_1+m_2)m+m_1m_2-(m-m_1)^2=\Delta m^2-(m-m_1)^2\,,\\ &m^2-(m_1+m_2)m+m_1m_2-m^2-m_1^2+2mm_1=\Delta m^2-(m-m_1)^2\,,\\ &(m_1-m_2)(m-m_1)=\Delta m^2-(m-m_1)^2\,. \end{aligned} \tag{4.29}$$

Having diagonalized the mass matrix, we can easily determine the evolution of the uncoupled basis states by taking into account that the energies of the decoupled modes are determined by the usual relativistic dispersion relation in the $p \gg m_\pm c$ limit (neutrinos have very small masses)

$$\varepsilon_\pm = \sqrt{p^2c^2+m_\pm^2c^4} \simeq pc + \frac{c^4m_\pm^2}{2E}\,, \tag{4.30}$$

with $E \simeq pc$. Consider, for example, an electron neutrino ν_e produced in the Sun which propagates towards the Earth with momentum $\vec{p}$. By inverting the eigenstates' relations in (4.25), the associated state vector at $t=0$ can be expressed as a function of the definite-mass/energy states $|\nu_\pm\rangle$, namely

$$|\nu,0\rangle \equiv |\nu_e,0\rangle = \cos\theta\,|\nu_+\rangle - \sin\theta\,|\nu_-\rangle\ , \tag{4.31}$$

where we have again adapted the general notation in Section 4.1, i.e. $|\nu,0\rangle \equiv |\Phi,0\rangle$. At a given time t, this state becomes

$$|\nu,t\rangle = \cos\theta e^{-i\varepsilon_+t/\hbar}\,|\nu_+\rangle - \sin\theta e^{-i\varepsilon_-t/\hbar}\,|\nu_-\rangle\,, \tag{4.32}$$

with the energies $\varepsilon_\pm$ determined by Eq. (4.30). This expression can be rewritten in a more enlightening way by eliminating the *uncoupled basis* states $|\nu_\pm\rangle$ in favour of the *coupled basis* $|\nu_e\rangle$ and $|\nu_\mu\rangle$,

$$|\nu,t\rangle = \left(\cos^2\theta e^{-i\varepsilon_+t/\hbar}+\sin^2\theta e^{-i\varepsilon_-t/\hbar}\right)|\nu_e\rangle + \frac{1}{2}\sin 2\theta\left(e^{-i\varepsilon_+t/\hbar}-e^{-i\varepsilon_-t/\hbar}\right)|\nu_\mu\rangle\ , \tag{4.33}$$

making explicit that the electron neutrinos produced in the Sun can transform, during their journey to Earth, into muon neutrinos. The probability of finding a muon neutrino at time t is then given by

$$\begin{aligned} P_\mu(t) = |\langle\nu_\mu|\nu,t\rangle|^2 &= \frac{1}{4}\sin^2 2\theta\left(e^{i\varepsilon_+t/\hbar}-e^{i\varepsilon_-t/\hbar}\right)\left(e^{-i\varepsilon_+t/\hbar}-e^{-i\varepsilon_-t/\hbar}\right)\\ &= \frac{1}{4}\sin^2 2\theta\left(1-e^{i(\varepsilon_+-\varepsilon_-)t/\hbar}-e^{-i(\varepsilon_+-\varepsilon_-)t/\hbar}+1\right) = \sin^2 2\theta\left[\frac{1}{2}-\frac{1}{2}\,\mathrm{Re}\left(e^{i(\varepsilon_+-\varepsilon_-)t/\hbar}\right)\right]\,, \end{aligned}$$

or equivalently

$$P_\mu(t) = |\langle\nu_\mu|\nu,t\rangle|^2 = \sin^2 2\theta\sin^2\frac{\Delta Et}{2\hbar}\,, \tag{4.34}$$

where we have introduced the energy difference $\Delta E \equiv \varepsilon_+ - \varepsilon_-$. As a result, the survival probability that it has remained an electron neutrino is

$$P_e(t) = 1-\sin^2 2\theta\sin^2\frac{\Delta Et}{2\hbar}\,. \tag{4.35}$$

Taking into account Eq. (4.30) and denoting by L the distance $L = ct$ travelled by the (approximately massless) neutrino, the above expressions become

$$P_\mu(t) = \sin^2 2\theta \sin^2 \frac{c^4 \Delta M^2 L}{4E\,\hbar c}\,, \qquad P_e(t) = 1 - \sin^2 2\theta \sin^2 \frac{c^4 \Delta M^2 L}{4E\,\hbar c}\,, \tag{4.36}$$

showing clearly that if $\Delta M^2 \equiv m_+^2 - m_-^2 \neq 0$, oscillations can take place. This *flavour-changing* behavior, governed by neutrino mass difference and the mixing angle θ, explains why some of the expected neutrinos were missing when detected on Earth, since the detectors in question were originally designed to exclusively identify electron neutrinos. Neutrino oscillations were later confirmed through other experiments, such as the Sudbury Neutrino Observatory (SNO) and Super-Kamiokande, which detected neutrinos from various sources, not just the Sun, and observed oscillations between different neutrino *flavours*.

A rough lower bound of the neutrino mass difference can be obtained by considering the maximal oscillation probability for the time needed for a neutrino to travel from the Sun to the Earth, namely

$$\frac{c^4 \Delta M^2 t}{4\hbar E} \leq \frac{\pi}{2} \quad \longrightarrow \quad \Delta M^2 = \frac{2\pi\hbar E}{c^4 t} \leq \frac{2\pi\hbar E}{c^3 L}\,, \tag{4.37}$$

with $L = ct \simeq 1.5 \times 10^8$ km the Sun to Earth distance. For a neutrino energy of $E \simeq 8$MeV, this translate into a tiny value

$$\sqrt{\Delta M^2} \lesssim 8 \times 10^{-6} \mathrm{eV}/c^2\,. \tag{4.38}$$

The above toy-model can be generalized to accommodate the 3 neutrino *flavours* existing in the Standard Model of particle physics. In this case, the definite *flavour* states $|\nu_\alpha\rangle$ and the states of definite mass $|\nu_i\rangle$ are connected through a 3x3 unitary matrix, known as the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix, the MNS matrix or the neutrino mixing matrix,[5]

$$|\nu_\alpha\rangle = \sum_i U_{\alpha i} |\nu_i\rangle\,, \quad |\nu_i\rangle \sum_\alpha U^*_{\alpha i} |\nu_\alpha\rangle\,, \tag{4.39}$$

with $\alpha = e, \mu, \tau$, and $i = 1, 2, 3$. The corresponding probability for a transition from a *flavour* state α at $t = 0$ to a *flavour* state β at time t is then given by

$$P_{\alpha\to\beta}(t) = \sum_{i,j} U_{\alpha i} U^*_{\beta i} U^*_{\alpha j} U_{\beta j} e^{-i\Delta E_{ij} t/\hbar}\,, \tag{4.40}$$

with

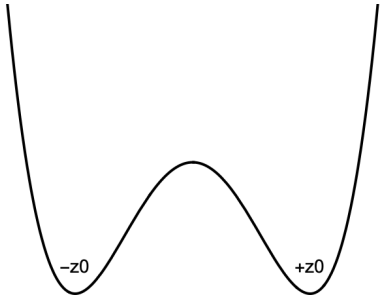
$$\Delta E_{ij} \simeq \frac{c^4 \Delta M^2_{ij}}{2E} = 1.27 \frac{\Delta M^2_{ij}}{(\mathrm{eV}/c^2)^2} \frac{L}{\mathrm{km}} \frac{\mathrm{GeV}}{E}\,. \tag{4.41}$$

4.4 The ammonia molecule

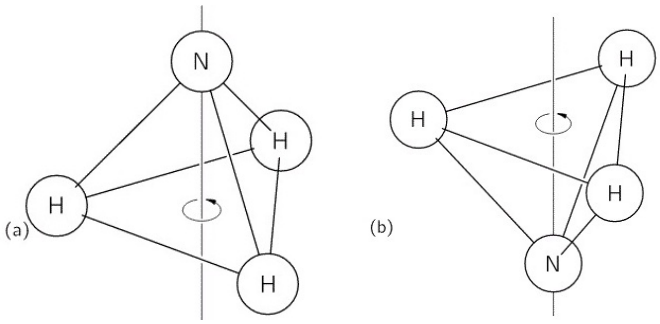
The ammonia molecule (NH_3) is another instructive example of a two-level system. This molecule is composed of one nitrogen atom and three hydrogen atoms, taking the shape of a

[5] This is analogous to the Cabibbo-Kobayashi-Maskawa (CKM) matrix in the context of quark *flavour* mixing.

pyramid or flattened tetrahedron. Neglecting for simplicity potential rotation effects in the molecule and assuming the nitrogen atom to be motionless, the potential energy of the system will be only a function of the distance between the nitrogen atom and the plane defined by the three hydrogen atoms. The potential experienced by the nitrogen atom can be schematically represented as a double-well potential like the one the figure,



The intermediate barrier parametrizes the fact that the nitrogen atom is effectively repelled by the hydrogen atoms when approaching the plane containing them. The two minima at $\pm z_0$ correspond to the two equilibrium configurations relevant for our discussion, nitrogen up and nitrogen down, and are classically stable and degenerate in energy,



Provided that the energy barrier is not infinite, this degeneracy will be unavoidably broken by quantum effects. To formalize this, let us denote by

$$|\varphi_1\rangle \quad \text{(nitrogen up)}\,, \qquad |\varphi_2\rangle \quad \text{(nitrogen down)}\,, \tag{4.42}$$

the quantum states associated to the wave functions around z_0 and $-z_0$. As in the classical picture, if the energy barrier is infinite, the Hamiltonian is diagonal

$$H = \begin{pmatrix} E_0 & 0 \\ 0 & E_0 \end{pmatrix}\,, \tag{4.43}$$

and the states $|\varphi_1\rangle$ and $|\varphi_2\rangle$ are degenerate in energy,

$$H|\varphi_1\rangle = E_0|\varphi_1\rangle\,, \qquad H|\varphi_2\rangle = E_0|\varphi_2\rangle\,. \tag{4.44}$$

However, if the energy of the barrier is not infinite, quantum tunneling can take place. Incorporating this effect through the off-diagonal elements of H and requiring the result to be Hermitian, we obtain

$$H = \begin{pmatrix} E_0 & -\Delta \\ -\Delta & E_0 \end{pmatrix} = E_0\mathbb{I} - \Delta\,\sigma_1\,, \tag{4.45}$$

with $-\Delta$ a real quantity to be fixed experimentally. Comparing again this expression with Eq. (4.1), we can easily identify the values of g_0, g_1, g_2 and g_3 (namely $g_0 = E_0$, $g_1 = -\Delta$, $g_1 = g_2 = 0$) and directly determine the eigenvalues and stationary states of the system. Alternatively, we can explicitly compute them using the characteristic equation $\det(H - \varepsilon_i \mathbb{I}) = 0$, obtaining

$$\begin{vmatrix} E_0 - \varepsilon & -\Delta \\ -\Delta & E_0 - \varepsilon \end{vmatrix} = 0 \quad \rightarrow \quad (E_0 - \varepsilon)^2 = \Delta^2 \quad \rightarrow \quad \varepsilon_\pm \quad = E_0 \pm \Delta \,. \tag{4.46}$$

The associated eigenstates are no longer degenerate in energy but rather correspond to an excited state (E) and a ground state (G),

$$|E\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}} (|\varphi_1\rangle - |\varphi_2\rangle) \,, \qquad |G\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} (|\varphi_1\rangle + |\varphi_2\rangle) \tag{4.47}$$

Denoting by $|\Phi\rangle$ the generic states in this new basis and if the ammonia molecule is initially in the state $|\varphi_1\rangle$,

$$|\Phi, 0\rangle = |\varphi_1\rangle = \frac{1}{\sqrt{2}} (|G\rangle + |E\rangle) \,, \tag{4.48}$$

the time-evolved state at time t will take the form

$$|\Phi, t\rangle = \frac{1}{\sqrt{2}} \left(e^{-i \frac{(E_0 - \Delta)t}{\hbar}} |G\rangle + e^{-i \frac{(E_0 + \Delta)t}{\hbar}} |E\rangle \right) \,, \tag{4.49}$$

which, as we did in the neutrino case, can be conveniently rewritten in terms of the states $|\varphi_1\rangle$ and $|\varphi_2\rangle$,

$$|\Phi, t\rangle = \frac{e^{-i \frac{E_0 t}{\hbar}}}{2} \left(e^{i \frac{\Delta t}{\hbar}} (|\varphi_1\rangle + |\varphi_2\rangle) + e^{-i \frac{\Delta t}{\hbar}} (|\varphi_1\rangle - |\varphi_2\rangle) \right) = e^{-i \frac{E_0 t}{\hbar}} \left[\cos \frac{\Delta\, t}{\hbar} |\varphi_1\rangle + i \sin \frac{\Delta\, t}{\hbar} |\varphi_2\rangle \right] \,, \tag{4.50}$$

making explicit that $|\Phi, t\rangle$ oscillates between nitrogen up and nitrogen down configurations, with a frequency $\omega = \Delta/\hbar \simeq 23.87\,\text{GHz}$ determined by the experimental energy difference $2\Delta = 0.9872 \times 10^{-4}$ eV. The associated time-dependent probabilities of finding the molecule in the states $|\varphi_1\rangle$ (nitrogen up) and $|\varphi_2\rangle$ (nitrogen down) are finally given by

$$P_1(t) = |\langle \varphi_1 | \varphi, t \rangle|^2 = \cos^2 \left(\frac{\Delta t}{\hbar} \right) \,, \qquad P_2(t) = |\langle \varphi_2 | \Phi, t \rangle|^2 = \sin^2 \left(\frac{\Delta t}{\hbar} \right) \,. \tag{4.51}$$

4.4.1 The ammonia molecule in an electric field

In the ammonia molecule the electrons tend to cluster towards the nitrogen, making its surroundings slightly negative and the hydrogen plane slightly positive. This produces an electric dipole moment $\vec{\mu}$ that points down when the nitrogen is up. When placed in an electric field pointing along the positive z direction, the dipole energy $E = -\vec{\mu} \cdot \vec{\mathcal{E}}$ induces two new contributions in (4.45), namely

$$H = \begin{pmatrix} E_0 + \mu \mathcal{E} & -\Delta \\ -\Delta & E_0 - \mu \mathcal{E} \end{pmatrix} = E_0 \, \mathbb{I} - \Delta \, \sigma_1 + \mu \, \mathcal{E} \sigma_3 \,. \tag{4.52}$$

The modified energy eigenvalues and eigenvectors take the form

$$\varepsilon_E(\mathcal{E}) = E_0 + \sqrt{\mu^2 \mathcal{E}^2 + \Delta^2} \,, \qquad \varepsilon_G(\mathcal{E}) = E_0 - \sqrt{\mu^2 \mathcal{E}^2 + \Delta^2} \,, \tag{4.53}$$

with the subscripts G and E referring again to the ground and excited states, respectively. According to these expressions and as a function of $\mathcal{E}$, ε_E and ε_G correspond to the branches of a hyperbola that approach asymptotically the values $E_0 \pm \mathcal{E}$. Therefore, small electric fields minimize the energy of the $|E\rangle$ state, while large ones modify the energy of the $|G\rangle$ state. This device acts therefore as an beam splitter, as can be seen explicitly by considering the weak field expansion ($\mu\mathcal{E}/\Delta \ll 1$) of (4.53),

$$\varepsilon_E(\mathcal{E}) \simeq E_0 + \Delta + \frac{1}{2}\frac{\mu^2\mathcal{E}^2}{\Delta} + \dots\,, \qquad \varepsilon_G(\mathcal{E}) \simeq E_0 - \Delta - \frac{1}{2}\frac{\mu^2\mathcal{E}^2}{\Delta} + \dots \tag{4.54}$$

In particular, if the ammonia molecules moves in a region where the electric field $\mathcal{E}$ is weak but the gradient $d\mathcal{E}^2/dz$ along the axis of the molecules is large, the molecules in the states $\varepsilon_E(\mathcal{E})$ and $\varepsilon_E(\mathcal{E})$ will experience opposite forces,

$$F_E = -\frac{d\varepsilon_E}{dz} \simeq -\frac{1}{2}\frac{\mu^2}{\Delta}d\frac{\mathcal{E}^2}{dz}\,, \qquad F_G = -\frac{d\varepsilon_G}{dz} = -F_E\,, \tag{4.55}$$

in complete analogy with a Stern-Gerlach apparatus.

4.4.2 The ammonia MASER

Consider now a beam of ammonia molecules in the $|E\rangle$ state entering a resonant electromagnetic cavity tuned to the frequency of 23.87 GHz through a small hole and exiting it through another one in the opposite side. If the ammonia molecules in the cavity interact with an electric field $\mathcal{E}$ oscillating with the resonant frequency 23.87 GHz, the interaction with that field will induce the transition $|E\rangle \to |G\rangle$, feeding extra energy into the field.

The mathematics of this transition become clearer if we express the Hamiltonian (4.52) in the $|E\rangle$ and $|G\rangle$ basis. Taking into account the corresponding matrix elements

$$\begin{aligned}
\langle E|H|E\rangle &= \frac{1}{2}(\langle\varphi_1| - \langle\varphi_2|)H(|\varphi_1\rangle - |\varphi_2\rangle) = \frac{1}{2}(\langle\varphi_1|H|\varphi_1\rangle - \langle\varphi_1|H|\varphi_2\rangle - \langle\varphi_2|H|\varphi_1\rangle + \langle\varphi_2|H|\varphi_2\rangle) \\
&= E_0 + \Delta\,, \\
\langle G|H|G\rangle &= E_0 - \Delta\,, \\
\langle E|H|G\rangle &= \frac{1}{2}(\langle\varphi_1| - \langle\varphi_2|)H(|+\rangle + |\varphi_2\rangle) = \frac{1}{2}(\langle+|H|\varphi_1\rangle + \langle\varphi_1|H|\varphi_2\rangle - \langle\varphi_2|H|\varphi_1\rangle - \langle\varphi_2|H|\varphi_2\rangle) \\
&= \mu\,\mathcal{E}\,,
\end{aligned}$$

we have

$$H = \begin{pmatrix} E_0 + \Delta & \mu\mathcal{E} \\ \mu\mathcal{E} & E_0 - \Delta \end{pmatrix}. \tag{4.56}$$

The associated Schrödinger equation for the amplitudes C_E and C_G to be in the $|E\rangle$ or $|G\rangle$ states is then given by

$$i\hbar\frac{d}{dt}\begin{pmatrix} C_E(t) \\ C_G(t) \end{pmatrix} = \begin{pmatrix} \Delta & \mu\mathcal{E} \\ \mu\mathcal{E} & -\Delta \end{pmatrix}\begin{pmatrix} C_E(t) \\ C_G(t) \end{pmatrix}, \tag{4.57}$$

where, for the sake of simplicity, we have set the constant energy E_0 to zero. Note that if $\mu\mathcal{E} = 0$, this equation admit an exact solution

$$\begin{pmatrix} C_E(t) \\ C_G(t) \end{pmatrix} = \begin{pmatrix} e^{-i\Delta t/\hbar} & \beta_E \\ e^{+i\Delta t/\hbar} & \beta_G \end{pmatrix}, \tag{4.58}$$

with time-independent β_E and β_G. Therefore, for small $\mu\mathcal{E}$, one can expect solutions

$$\begin{pmatrix} C_E(t) \\ C_G(t) \end{pmatrix} = \begin{pmatrix} e^{-i\Delta t/\hbar} & \beta_E(t) \\ e^{+i\Delta t/\hbar} & \beta_G(t) \end{pmatrix}, \tag{4.59}$$

with slowly varying β_E and β_G. Substituting this expression into Eq. (4.57) and performing some trivial algebra and cancellations, we get

$$i\hbar\frac{d}{dt}\begin{pmatrix} \beta_E(t) \\ \beta_G(t) \end{pmatrix} = \begin{pmatrix} 0 & e^{i\omega_0 t}\mu\mathcal{E} \\ e^{-i\omega_0 t}\mu\mathcal{E} & 0 \end{pmatrix}\begin{pmatrix} \beta_E(t) \\ \beta_G(t) \end{pmatrix}, \tag{4.60}$$

with $\omega_0 \equiv 2\Delta/\hbar$ the frequency of a photon associated to the transition $|E\rangle \to |G\rangle$. If the electric field in the cavity is set in resonance, we have

$$\mathcal{E}(t) = 2\mathcal{E}_0 \cos\omega_0 t = \mathcal{E}_0\left(e^{i\omega_0 t} + e^{-i\omega_0 t}\right) \qquad \begin{cases} e^{i\omega_0 t}\mu\mathcal{E} = \mu\mathcal{E}_0\left(1 + e^{2i\omega_0 t}\right), \\ e^{-i\omega_0 t}\mu\mathcal{E} = \mu\mathcal{E}_0\left(1 + e^{-2i\omega_0 t}\right), \end{cases} \tag{4.61}$$

and

$$i\dot{\beta}_E(t) = \frac{\mu\mathcal{E}_0}{\hbar}\left(1 + e^{2i\omega_0 t}\right)\beta_G(t) \simeq \frac{\mu\mathcal{E}_0}{\hbar}\beta_G(t), \tag{4.62}$$

$$i\dot{\beta}_G(t) = \frac{\mu\mathcal{E}_0}{\hbar}\left(1 + e^{-2i\omega_0 t}\right)\beta_E(t) \simeq \frac{\mu\mathcal{E}_0}{\hbar}\beta_E(t). \tag{4.63}$$

where in the last step, we have dropped the exponentials since these have zero time-averaged values. Taking another time derivative of the top equation we find

$$\ddot{\beta}_E(t) = -\left(\frac{\mu\mathcal{E}_0}{\hbar}\right)^2\beta_E(t), \tag{4.64}$$

which for a molecule in the state $|E\rangle$ at $t = 0$ has the simple solution

$$\beta_E(t) = \cos\left(\frac{\mu\mathcal{E}_0}{\hbar}t\right). \tag{4.65}$$

The time-dependent probability to be in the $|E\rangle$ state at time t is then given by

$$P_E(t) = |C_E(t)|^2 = \left|e^{-i\Delta t/\hbar}\beta_E(t)\right| = \cos^2\left(\frac{\mu\mathcal{E}_0}{\hbar}t\right). \tag{4.66}$$

Requiring that the travel time of the molecules entering the cavity in the state $|E\rangle$ is such that they leave it in the state $|G\rangle$,

$$\cos\left(\frac{\mu\mathcal{E}_0}{\hbar}T\right) = 0 \quad \to \quad \frac{\mu\mathcal{E}_0}{\hbar}T = \frac{\pi}{2}, \frac{3\pi}{2}, \ldots \tag{4.67}$$

each molecule gives an energy 2Δ to the electromagnetic field in the cavity, building up a coherent field or MASER (Microwave Amplification by Stimulated Emission of Radiation).

The first ammonia maser marked a groundbreaking achievement in the realm of quantum electronics. For their groundbreaking contributions to the development of this revolutionary technology, Charles H. Townes, James P. Gordon, and H. J. Zeiger were awarded the Nobel Prize in Physics in 1964.

4.5 Unstable states: an effective non-Hermitian Hamiltonian

In all the previous examples we have implicitly assumed the states in the coupled basis to be absolutely stable. Let us consider now the situation in which $|\varphi_1\rangle$ represents the excited state of a non-isolated two-level system able to radiate energy to the environment through the transition to a lower energy level $|\varphi_2\rangle$.

Atoms and radiation

When studying the hydrogen atom in *Quantum Physics I* you implicitly assumed the system to be completely isolated from the electromagnetic field, obtaining absolutely stable or stationary energy levels. While this idealized description serves as a valuable starting point, a more comprehensive understanding of atomic physics requires the inclusion of electromagnetic interactions proportional to the fine structure constant. This interactions allow for the decay of "stationary" states.

In order to incorporate the instability of the state, we will replace the associated energy eigenvalue by

$$E_1 \;\rightarrow\; E_1 - i\frac{\Gamma_1}{2}\,, \tag{4.68}$$

with $\Gamma_1 \equiv \hbar\gamma_1$ the so-called decay width, $\gamma_1 \equiv 1/\tau_1$ and τ_1 the lifetime of the state. The matrix representing the resulting non-Hermitian "Hamiltonian" in the coupled basis $\{|\varphi_1\rangle\,, |\varphi_2\rangle\}$ can be then written as

$$H = \begin{pmatrix} E_1 - i\frac{\Gamma_1}{2} & \Delta \\ \Delta^* & E_2 \end{pmatrix}\,, \tag{4.69}$$

with Δ a weak coupling between states. Note that the phenomenological replacement (4.68) effectively captures the idea of instability. In particular, for $\Delta = 0$, the probability of measuring $|\varphi_1\rangle$ at a given time t is no longer conserved,[6] but rather decays exponentially within a characteristic time τ_1,

$$P_{11}(t) = |\langle\varphi_1|\varphi_1\rangle|^2 = e^{-t/\tau_1}\,. \tag{4.70}$$

In the presence of coupling, the situation becomes slightly more involved and interesting. To determine the evolution of the system in this case, we must first diagonalize the "Hamiltonian" (4.69), determining the corresponding eigenvalues and eigenvectors in the uncoupled basis $\{|\Phi_1\rangle, |\Phi_2\rangle\}$. The eigenvalues of the matrix (4.69) follow directly from the characteristic equation $\det(H - \varepsilon\mathbb{I}) = 0$,

$$\varepsilon^2 - \varepsilon\left(E_1 + E_2 - i\frac{\Gamma_1}{2}\right) + E_1E_2 - i\frac{\Gamma_1}{2}E_2 - |\Delta|^2 = 0\,, \tag{4.71}$$

[6] This contrasts with the situation in strictly conservative, isolated systems, where the Hamiltonian is a constant of motion and all eigenstates remain stationary,

$$H\,|\varphi_n\rangle = E_n\,|\varphi_n\rangle\,, \qquad |\varphi_n, t\rangle = e^{-iE_nt/\hbar}\,|\varphi_n\rangle\,.$$

The non-conservation of probability is, however, not surprising, since the effective "Hamiltonian" (4.69) is clearly non-Hermitian, as we are explicitly neglecting the energy driven away from the system.

which, in the weak coupling regime

$$|\Delta| \ll \sqrt{(E_1 - E_2)^2 + \frac{\Gamma_1^2}{4}}\,, \tag{4.72}$$

admits two solutions

$$\varepsilon_1 \simeq E_1 - i\frac{\Gamma_1}{2} + \frac{|\Delta|^2}{E_1 - E_2 - i\frac{\Gamma_1}{2}}\,, \qquad \varepsilon_2 \simeq E_2 + \frac{|\Delta|^2}{E_2 - E_1 + i\frac{\Gamma_1}{2}}\,. \tag{4.73}$$

The energies and inverse lifetimes of the states in the presence of the coupling are given by the real and imaginary parts of ε_1 and ε_2. In particular, due to the mixing Δ, the initially stable state $|\varphi_2\rangle$ acquires a finite imaginary part

$$\mathrm{Im}\,\epsilon_2 = \frac{\Gamma_1}{2}\frac{|\Delta|^2}{(E_2 - E_1)^2 + \frac{\Gamma_1^2}{4}}\,. \tag{4.74}$$

and therefore a finite lifetime! Note also that the smaller the difference between the unperturbed energies E_1 and E_2, the more effectively the perturbation acts on the energies and lifetimes. The limiting case $E_1 = E_2$ is studied in the next section.

4.5.1 Coupling of states of the same energy

When the energies E_1 and E_2 are equal, the operator (4.69) admits a decomposition

$$H = \left(E_1 - i\frac{\Gamma_1}{4}\right)\mathbb{I} + K\,, \qquad K = \begin{pmatrix} -i\frac{\Gamma_1}{4} & \Delta \\ \Delta^* & i\frac{\Gamma_1}{4} \end{pmatrix}\,. \tag{4.75}$$

The eigenvalues of H are then given by

$$\varepsilon_1 = E_1 - i\frac{\Gamma_1}{4} + k_1\,, \qquad \varepsilon_2 = E_1 - i\frac{\Gamma_1}{4} - k_1\,, \tag{4.76}$$

with k_1 and $-k_1$ the two solutions of the characteristic equation for K, $\det(K - k\mathbb{I}) = 0$, namely

$$k^2 = |\Delta|^2 - \frac{\Gamma_1^2}{16} \quad \longrightarrow \quad k_1 = -k_2\,. \tag{4.77}$$

The eigenvectors of H and K are given by

$$|\Phi_1\rangle = \Delta\,|\varphi_1\rangle + \left(k_1 + i\frac{\Gamma_1}{4}\right)|\varphi_2\rangle\,, \qquad |\Phi_2\rangle = \Delta\,|\varphi_1\rangle + \left(-k_1 + i\frac{\Gamma_1}{4}\right)|\varphi_2\rangle\,. \tag{4.78}$$

In the uncoupled basis $\{|\Phi_1\rangle, |\Phi_2\rangle\}$, the Hamiltonian of the system becomes diagonal, allowing to determine the evolution of a given initial state using the standard procedure for time-independent Hamiltonians. Assuming, for instance, that the system is initially in the state

$$|\Phi, 0\rangle = |\varphi_2\rangle = \frac{1}{2k_1}\left[|\Phi_1\rangle - |\Phi_2\rangle\right]\,, \tag{4.79}$$

the state vector at time t becomes

$$|\Phi, t\rangle = \frac{1}{2k_1}e^{-iE_1t/\hbar}e^{-\frac{1}{4}\gamma_1 t}\left[e^{-ik_1t/\hbar}\,|\Phi_1\rangle - e^{ik_1t/\hbar}\,|\Phi_2\rangle\right]\,. \tag{4.80}$$

The probability of finding the system at time t in the state $|\varphi_1\rangle$ is therefore

$$\begin{aligned} P_{21}(t) = |\langle\varphi_1|\Phi(t)\rangle|^2 &= \frac{1}{4\,|k_1|^2} e^{-\gamma_1/2} \left| e^{-ik_1//\hbar} \langle\varphi_1|\Phi_1\rangle - e^{ik_1/\hbar} \langle\varphi_1|\Phi_2\rangle \right|^2 \\ &= \frac{1}{4\,|k_1|^2} e^{-\gamma_1 t/2} \,|\Delta|^2 \left| e^{-ik_1 t/\hbar} - e^{ik_1 t/\hbar} \right|^2 . \end{aligned} \tag{4.81}$$

A simple inspection of (4.77) reveals two different regimes, in close analogy with those appearing in a classical damped harmonic oscillator:

- $|\Delta| > \frac{\Gamma_1}{4}$, $\quad k_1 = -k_2 = \sqrt{|\Delta|^2 - \left(\frac{\Gamma_1}{4}\right)^2}$. In this case the eigenvalues of the system have the same imaginary part, but different real parts,

$$\varepsilon_1 = E_1 + \sqrt{|\Delta|^2 - \left(\frac{\Gamma_1}{4}\right)^2} - i\frac{\Gamma_1}{4}\,, \qquad \varepsilon_2 = E_1 - \sqrt{|\Delta|^2 - \left(\frac{\Gamma_1}{4}\right)^2} - i\frac{\Gamma_1}{4}\,, \tag{4.82}$$

  meaning that the two states $|\Phi_1\rangle$ and $|\Phi_2\rangle$ display the same lifetime, $2\tau_1$, but different energies. The associated probability of finding the system in the state $|\varphi_1\rangle$ is then represented by a damped sinusoid with time constat $2\tau_1$,

$$P_{21}(t) = \frac{|\Delta|^2}{|\Delta|^2 - \left(\frac{\Gamma_1}{4}\right)^2} e^{-t/(2\tau_1)} \sin^2\left(\sqrt{|\Delta|^2 - \left(\frac{\Gamma_1}{4}\right)^2}\, \frac{t}{\hbar} \right) . \tag{4.83}$$

  In other words, the coupling is sufficiently strong to make the system oscillate between the states $|\varphi_1\rangle$ and $|\varphi_2\rangle$ before the instability of $|\varphi_1\rangle$ start playing a role.

- $|\Delta| < \frac{\Gamma_1}{4}$, $\quad k_1 = -k_2 = i\sqrt{\left(\frac{\Gamma_1}{4}\right)^2 - |\Delta|^2}$. In this case, the eigenvalues of the system have the same real part, but different imaginary parts,

$$\varepsilon_1 = E_1 - i\left[\frac{\Gamma_1}{4} - \sqrt{\left(\frac{\Gamma_1}{4}\right)^2 - |\Delta|^2}\right] , \qquad \varepsilon_2' = E_1 - i\left[\frac{\Gamma_1}{4} + \sqrt{\left(\frac{\Gamma_1}{4}\right)^2 - |\Delta|^2}\right] , \tag{4.84}$$

  meaning that the two states $|\Phi_1\rangle$ and $|\Phi_2\rangle$ have the same energy E_1 but different lifetimes. The associated probability of finding the system in the state $|\varphi_1\rangle$ is represented by a sum of exponentials,

$$P_{21}(t) = \frac{|\Delta|^2}{\left(\frac{\Gamma_1}{4}\right)^2 - |\Delta|^2} e^{-t/(2\tau_1)} \sinh^2\left(\sqrt{\left(\frac{\Gamma_1}{4}\right)^2 - |\Delta|^2}\, \frac{t}{\hbar} \right) . \tag{4.85}$$

  In other words, the lifetime τ_1 is so short that the system becomes completely damped before it has time to oscillate between the states $|\varphi_1\rangle$ and $|\varphi_2\rangle$.

4.6 Exercises

1. **Eigenvalues of a general two-level system**

   Determine the eigenvalues and eigenstates of the Hamiltonian

   $$H = g_0\,\mathbb{I} + g_1\,\sigma_1 + g_2\,\sigma_2 + g_3\,\sigma_3 = g_0\mathbb{I} + \vec{g}\cdot\vec{\sigma}$$

   with $\mathbb{I}$ the identity matrix, σ_i the linearly independent Hermitian Pauli matrices [$i = 1, 2, 3$ (x, y, z)], $\vec{\sigma}$ a vector $\sigma = (\sigma_x, \sigma_y, \sigma_z)^T$ and $\vec{g} = g\,\vec{n}$ a vector with components (g_1, g_2, g_3), magnitude $g = \sqrt{g_1^2 + g_2^2 + g_3^2}$ and direction $\vec{n}$ ($\vec{n}\cdot\vec{n} = 1$). Express the eigenstates in terms of the ratios g_3/g and g_1/g_2 only. Can you find the connection with the form of a general vector in the Bloch's sphere? This is a good result to remember.

2. **Neutral K-meson Oscillations**

   There exist two types of neutral K-mesons that can be produced using down and strange quarks, $|K^0\rangle = |d\bar{s}\rangle$ and $|\overline{K}^0\rangle = |s\bar{d}\rangle$ [We call them "kay-zero" and the "kay-zero-bar"– the generic term for a K-meson is a kaon, pronounce "kay-on"]. If the weak interaction did not exist, both species would possess equal masses and the Hamiltonian governing the behavior of a kaon with zero momentum would be simply

   $$H_0 = \begin{pmatrix} M & 0 \\ 0 & M \end{pmatrix},$$

   with $M = 497.6$ MeV. The presence of the weak interaction leads, however, to an extra term

   $$H_m = \begin{pmatrix} 0 & \epsilon \\ \epsilon & 0 \end{pmatrix}.$$

   effectively mixing the aforementioned states. The two resulting eigenstates

   $$|K_S\rangle = \frac{1}{\sqrt{2}}\left(|K^0\rangle + |\overline{K}^0\rangle\right), \qquad |K_L\rangle = \frac{1}{\sqrt{2}}\left(|K^0\rangle - |\bar{K}^0\rangle\right).$$

   are respectively known as K-short and K-long, since they have different decay lifetimes τ ($\tau_{K_S} = 0.0896$ ns, $\tau_{K_L} = 51.2$ ns).

   a After introducing the mixing term, the masses of kaons differ by 3.56 μeV. Which is the value of ϵ?

   b If a kaon is created in the K_0 state at time $t = 0$, determine the probability of measuring it as a $\bar{K}^0$ as a function of time.

   c A beam of kaons is produced in the K_0 channel with a kinetic energy of 600 MeV per kaon. Plot the probability of the kaons being in the K_0 state as a function of the distance traveled x. Neglect the eventual decay of kaons.

   d Repeat the previous task taking into account the kaons decay. *Hint*: The wave function should be modified by a factor $e^{-t/(2\tau)}$ to take into account decays of lifetime τ. It is also convenient to remember that $\hbar c = 197.327$ MeV $\cdot$ fm.

3. **The ammonia molecule**

The molecular structure of NH_3 consists of four atoms, namely one nitrogen and three hydrogen. The nitrogen atom is positioned above a base formed by an equilateral triangle made up of three hydrogen atoms. Ignoring electronic, vibrational excitations and rotational excitations, this system admits a two-state description. The two states arise from transitions where the nitrogen atom undergoes a flipping motion, transitioning from a position z_0 above the fixed hydrogen plane to a position $-z_0$ below it. These states are separated by a potential energy barrier.

In a basis $|+\rangle, |-\rangle$ this system can be described by a Hamiltonian

$$H = \begin{pmatrix} E_0 & -\Delta \\ -\Delta & E_0 \end{pmatrix}$$

with the off-diagonal elements $\langle +|H|-\rangle = \langle -|H|+\rangle^* = -\Delta \simeq -0.4936 \times 10^{-4}\text{eV} \neq 0$ encoding the possibility to tunnel between the nitrogen up and nitrogen down configurations.

a) Determine the eigenvalues of the Hamiltonian H by using the spin precession analogy. Compare the result with that following from the direct diagonalization of H.

b) Determine the wavelength of the photon that may be emitted in the transtion between the previous states.

c) Determine the time evolution of the system if at $t = 0$ the ammonia molecule is in the state $|+\rangle$. Which are probabilities of finding the state in the up or down position at time t?

4. **Microwave Amplification by Stimulated Emission of Radiation (MASER)**

In the ammonia molecule, the electrons have a tendency to concentrate near the nitrogen atom, causing the nitrogen vertex to have a slightly negative charge and the hydrogen plane to have a slightly positive charge. Consequently, one obtains a downward-pointing electric dipole moment $\vec{\mu}$ when the nitrogen atom is up. In the presence of an electric field $\mathcal{E} = \mathcal{E}\vec{z}$ directed along the z axis and assuming the electric dipole moment to be $\vec{\mu} = -\mu\vec{z}$ with $\mu > 0$, the energy of the nitrogen-up state $|+\rangle$ increases by $\mu\mathcal{E}$, whereas the nitrogen-down state $|-\rangle$ decreases by the same amount. This leads to a modified Hamiltonian

$$H = \begin{pmatrix} E_0 + \mu\mathcal{E} & -\Delta \\ -\Delta & E_0 - \mu\mathcal{E} \end{pmatrix},$$

a) Determine the eigenvalues of H by using the spin precession analogy. Compare the result with that following from the direct diagonalization of H.

b) By considering the limit $\mu\mathcal{E}/\Delta \ll 1$, argue that this settings acts as a beam splitter.

c) Imagine that a beam of ammonia molecules in the $|E\rangle$ state enters a high-Q electromagnetic cavity through a small hole at one end. Argue that adjusting the velocity of the ammonia molecules, each molecule can give an energy 2Δ to the electromagnetic field of the cavity if the molecule travels a specific amount of time T ensuring that the molecule leaves the cavity in the state $|G\rangle$. This idea is the basis of the first ammonia maser, recognized with the Nobel Prize in Physics in 1964. Determine the travel time T.

5. **A biased double-well potential** ✐

   Consider a quantum particle confined in a one-dimensional double-well potential. Assume that the two lowest energy eigenstates are well separated from higher excited states, so that the system can be accurately described within a two-dimensional Hilbert space spanned by the localized states $|L\rangle$ and $|R\rangle$, representing the particle being confined in the left and right well, respectively.

   The wells are not identical: there is a small energy bias $\epsilon \in \mathbb{R}$, such that the left well has energy greater than the right well by ϵ. Additionally, the particle can tunnel between the wells with amplitude $\Delta > 0$.

   a) Construct the Hamiltonian operator H in the basis $\{|L\rangle, |R\rangle\}$, expressing it as a 2×2 matrix. Clearly state the physical interpretation of each term in your Hamiltonian, identifying the contributions from the energy bias and the tunneling amplitude.

   b) Rewrite the Hamiltonian in terms of Pauli matrices. Interpret the result as the Hamiltonian of a spin-1/2 particle in an effective magnetic field.

   c) Determine the eigenvalues and the corresponding normalized eigenvectors of the Hamiltonian. Discuss whether the eigenstates are localized in one well or delocalized across both wells.

   d) Suppose the particle is initially localized in the left well, that is, $|\varphi, 0\rangle = |L\rangle$. Determine the state $|\varphi, t\rangle$ at time t, and compute the probability of finding the particle in the right well as a function of time.

   e) Now consider an open version of the system in which the particle can irreversibly decay from the right well into the environment with rate $\Gamma > 0$. Propose a simple effective non-Hermitian Hamiltonian that models this process within the two-level approximation. Describe **qualitatively** how the presence of this decay channel modifies the system's dynamics. In particular, explain how the coherent oscillations between wells are affected and what the long-time behavior of the system is.

6. **Diagnosing a faulty two-level system Hamiltonian** ✐

   Suppose the poor Javier did a typographical error when writing a two-level system Hamiltonian

   $$H = a|1\rangle\langle 1| + b|2\rangle\langle 2| + c|1\rangle\langle 2|$$

   in the orthonormal basis $\{|1\rangle, |2\rangle\}$, with $a, b, c \in \mathbb{R}$.

   a) What fundamental principle of quantum mechanics does this Hamiltonian violate? Illustrate the issue by solving the time-dependent Schrödinger equation for $a = b = 0$.

   b) What condition should the Hamiltonian with $a = b = 0$ satisfy in order for the time evolution to be physically consistent?

7. **A qutrit** ✐

   Consider a quantum system satisfying

   $$H|+\rangle = +\hbar\omega|+\rangle\ , \qquad H|0\rangle = 0|0\rangle\ , \qquad H|-\rangle = -\hbar\omega|-\rangle\ ,$$

   with $\{|-\rangle, |0\rangle, |+\rangle\}$ an orthonormal basis and H the associated Hamiltonian.

a) Write an expression for H in terms of projection operators. Use this result to derive the matrix representation of H in the $\{|-\rangle, |0\rangle, |+\rangle\}$ basis. Identify the corresponding stationary states.

b) If the system is initially prepared in a state $|\varphi\rangle = \frac{1}{\sqrt{3}}\left(|-\rangle + i\,|0\rangle - |+\rangle\right)$, which are the possible outcomes of measuring H? And the associated probabilities?

c) Under which conditions could the matrix

$$A = \theta \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

represent a measurable quantity? Assume this condition to hold in what follows.

d) If the system is initially prepared in the state $|\varphi\rangle$, which are the possible outcomes of measuring A? And the associated probabilities?

e) Is there any physical state on which both A and H can be measured with certainty?

f) Which are the probabilities of obtaining the eigenvalues E_i and a_j if the associated observables H and A are measured in this specific order? What happens in we measure them in the opposite order?

g) If a measurement of A at $t = t_0$ results on a value $\sqrt{2}\theta$, which is the state of the system at $t > t_0$? And the expectation value of A?

h) Determine the probability of obtaining a value $\sqrt{2}\theta$ if we perform a second measurement of A at $t > t_0$.

8. **A three-level atom**

Consider a three-level atom involving a non-degenerate ground state and a two-fold degenerate excited state, defined by

$$H|g\rangle = 0\,, \qquad H|e_1\rangle = \epsilon\,|e_1\rangle\,, \qquad H|e_2\rangle = \epsilon\,|e_2\rangle\,,$$

with $\{|g\rangle, |e_1\rangle, |e_2\rangle\}$ an orthonormal basis, H the associated Hamiltonian and ϵ a positive constant.

a) Write H in terms of projection operators. Use this result to derive the matrix representation of H in the $\{|g\rangle, |e_1\rangle, |e_2\rangle\}$ basis. Identify the corresponding stationary states.

b) An observable A in this system is defined by its operation on the eigenstates of H,

$$A|g\rangle = \alpha|e_1\rangle\,, \qquad A|e_1\rangle = \alpha|g\rangle\,, \qquad A|e_2\rangle = -\alpha|e_2\rangle\,,$$

with $\alpha > 0$ a constant. What are all the possible outcomes of measuring A?

c) For each of the eigenstates of H, calculate the probability of obtaining each value of A if the system is in that eigenstate.

d) Do H and A have common eigenstates? Are they compatible observables? Explain.

e) If at time $t = 0$ the system is in the eigenstate of A with the largest eigenvalue, calculate the probabilities of obtaining the different possible results of a measurement of A as a function of time.

f) If at time $t = 0$ the system is in the state $|\varphi\rangle = \frac{1}{\sqrt{2}}(|e_1\rangle + |e_2\rangle)$, calculate the probabilities of obtaining the different possible results of a measurement of A as a function of time. Explain the differences between this result and the one found in part e).

CHAPTER 5

COMPOSITE SYSTEMS AND IDENTICAL PARTICLES

The Exclusion Principle is laid down purely for the benefit of the electrons themselves, who might be corrupted (and become dragons or demons) if allowed to associate too freely.

Alan Turing in a postcard to his logician friend Robin Gandy.
online at the Turing Archive

The career of a young theoretical physicist consists of treating the harmonic oscillator in ever-increasing levels of abstraction.

Sidney Coleman

So far we have been dealing with isolated quantum systems where no interactions with other systems are, or have ever been, present or problems that under specific circumstances can be factorized into independent single-particle systems. The free particle, the harmonic oscillator or the hydrogen atom you studied in your *Quantum Physics I* course are just some specific examples of this. This naive approach effectively disregards many relevant aspects in the description of realistic physical systems, like the potential interference between spins and wave functions or the uncertainties associated with lack of knowledge on a system, either because it involves numerous degrees of freedom, as happens for instance in solids or gases, or because we have only access to a subsystem of it. How should we proceed in those situations?

5.1 Composite systems

Postulate V

When dealing with composite quantum systems, the state space of the composite system is the tensor product of the state spaces of the individual systems. The evolution of the composite system is determined by the unitary transformations applied to the combined state vector.

Tensor products are used to describe physical systems consisting of multiple subsystems. In what follows we will focus on tensor products of two Hilbert spaces, but the most of the discussion can be easily generalized to tensor products of three or more Hilbert spaces.

Given two independent Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ with respective bases $\{|i\rangle\}$ and $\{|\alpha\rangle\}$, the so-called tensor product, Kronecker product or direct product $\mathcal{H}_1 \otimes \mathcal{H}_2$ is defined as the vector space over $\mathbb{C}$ spanned by all pairs of elements[1]

$$|i\rangle \otimes |\alpha\rangle \,, \tag{5.1}$$

with defining properties

$$\begin{aligned} &c(|i\rangle \otimes |\alpha\rangle) = (c|i\rangle) \otimes |\alpha\rangle = |i\rangle \otimes (c|\alpha\rangle) \,, && \text{(Complex multiplication linearity)} \\ &(|i\rangle + |j\rangle) \otimes |\alpha\rangle = |i\rangle \otimes |\alpha\rangle + |j\rangle \otimes |\alpha\rangle \,, && \text{(Vector addition distributivity)} \\ &|i\rangle \otimes (|\alpha\rangle + |\beta\rangle) = |i\rangle \otimes |\alpha\rangle + |i\rangle \otimes |\beta\rangle \,, \end{aligned} \tag{5.2}$$

for any $c \in \mathbb{C}$. The first property relates $(c|i\rangle) \otimes |\alpha\rangle$ and $|i\rangle \otimes (c|\alpha\rangle)$ to $c(|i\rangle \otimes |\alpha\rangle)$, ensuring that these objects are not linearly independent, which would give rise to a much larger space (remember that for instance $c|i\rangle$ and $|i\rangle$ are the same state in $\mathcal{H}_1$). The other properties are imposed for the sake of linearity. The bases states $|i\rangle$ and $|\alpha\rangle$ can be associated to independent properties of the same particle or of different particles. For instance, for a particle moving in three dimensions, the overall Hilbert space is constructed as the tensor product of three one-dimensional spaces, each corresponding to motion along one axis. When considering two distinct particles, their combined Hilbert space is the tensor product of the individual Hilbert spaces associated with each particle. If a particle has spin, the Hilbert space incorporates both the wave function for the center of mass, representing the spatial degrees of freedom, and the spin space, which is two-dimensional in the case of a spin-1/2 particle.

Notation

We will frequently drop the $\otimes$ symbol when expressing the tensor product of two vectors $|i\rangle$ and $|\alpha\rangle$,

$$|i\rangle \otimes |\alpha\rangle \equiv |i\rangle|\alpha\rangle \equiv |i\,, \alpha\rangle \,,$$

being any of these alternative notations appropriate as long as it is clear from the context that a tensor product is implied. Keep in mind, however, that while a double label in a ket, such as $|i, \alpha\rangle$, typically indicates a tensor product, this is not always true. For instance, in the case of the angular momentum, labels like $|l, m\rangle$ do not actually represent a tensor product.

The linear functional or bra vector corresponding to the product state $|i\rangle \otimes |\alpha\rangle$ is defined as

$$(|i\rangle \otimes |\alpha\rangle)^\dagger = \langle i| \otimes \langle \alpha| \,, \tag{5.3}$$

with the order of the factors on either side of $\otimes$ remaining unchanged when the dagger operation is applied. Given this functional, the inner product in $\mathcal{H}_1 \otimes \mathcal{H}_2$ is naturally defined in terms of the inner products $\langle|\rangle_1$ and $\langle|\rangle_2$ of the basis elements in $\mathcal{H}_1$ and $\mathcal{H}_2$,

$$(\langle i| \otimes \langle \alpha|)\,(|j\rangle \otimes |\beta\rangle) = \langle i|j\rangle_1 \langle \alpha|\beta\rangle_2 \,, \tag{5.4}$$

[1] The basis obtained by the tensor product of bases $|i\rangle$ and $|\alpha\rangle$ used in this definition is of course just one of the numerous bases of $\mathcal{H}_1 \otimes \mathcal{H}_2$. In particular, an orthonormal basis for $\mathcal{H}_1 \otimes \mathcal{H}_2$ can be composed entirely of generic product states without necessarily being a product of bases.

such that if the bases $\{|i\rangle\}$ and $\{|\alpha\rangle\}$ are orthonormal bases, $\{|i\rangle \otimes |\alpha\rangle\}$ is also an orthonormal basis of $\mathcal{H}_1 \otimes \mathcal{H}_2$,

$$(\langle i| \otimes \langle \alpha|)\,(|j\rangle \otimes |\beta\rangle) = \delta_{ij}\delta_{\alpha\beta}\,. \tag{5.5}$$

Consequently, if $\mathcal{H}_1$ and $\mathcal{H}_2$ are spaces of finite dimension n_1 and n_2, the dimension of $\mathcal{H}_1 \otimes \mathcal{H}_2$ is $n_1 n_2$.

Not a direct sum

Note that the tensor product operates quite differently from the "usual" product $\oplus$ or direct sum of two vector spaces, where for two vectors $|\varphi\rangle$ and $|\chi\rangle$ one has simply

$$|\varphi\rangle \oplus |\chi\rangle = \begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_{n_1} \\ \hline \chi_1 \\ \vdots \\ \chi_{n_2} \end{pmatrix} = \begin{pmatrix} \vec{v} \\ \hline \vec{w} \end{pmatrix}\,.$$

In particular, if $\mathcal{H}_1$ and $\mathcal{H}_2$ have finite dimensions n_1 and n_2, the direct sum of $\mathcal{H}_1$ and $\mathcal{H}_2$ has dimension $n_1 + n_2$, instead of $n_1 n_2$.

A general element of $\mathcal{H}_1 \otimes \mathcal{H}_2$ can be written as a linear combination

$$|\varphi\rangle = \sum_{i,\alpha} r_{i\alpha} |i\rangle \otimes |\alpha\rangle\,. \tag{5.6}$$

In particular, given two states $|\varphi_1\rangle = \sum_i a_i |i\rangle \in \mathcal{H}_1$ and $|\varphi_2\rangle = \sum_\alpha b_\alpha |\alpha\rangle \in \mathcal{H}_2$, the vector $|\varphi_1\rangle \otimes |\varphi_2\rangle$ in the basis $|i\rangle \otimes |\alpha\rangle$ can be written as

$$|\varphi_1\rangle \otimes |\varphi_2\rangle = \sum_{i,\alpha} a_i, b_\alpha\, |i\rangle \otimes |\alpha\rangle\,. \tag{5.7}$$

Therefore, *the components of a tensor product vector are the product of the components of the two vectors in the product.* Note, however, that not all vectors of the form (5.6) can be expressed as (5.7). States that can be expressed as the product of one state $|\varphi_1\rangle$ in $\mathcal{H}_1$ and one state $|\varphi_2\rangle$ in $\mathcal{H}_2$ are called *separable*, *product* or *pure* states. On the other hand, linear combinations of two or more separable states are called *entangled* or *non-separable* states and are only possible in quantum systems.

Example 15: Spin 1/2 and spatial degrees of freedom

Since spin is independent of position and momentum, the associated operators com-

mute,

$$[\vec{S},\vec{x}]=0\,,\qquad [\vec{S},\vec{p}]=0\,,\qquad [\vec{S},\vec{L}]=0\,. \tag{5.8}$$

The total state of a spin-1/2 particle is therefore constructed from the direct product of position and spin eigenstates,

$$|\vec{x},+\rangle\equiv|\vec{x}\rangle\otimes|+\rangle,\qquad |\vec{x},-\rangle\equiv|\vec{x}\rangle\otimes|-\rangle\,, \tag{5.9}$$

and the total wave function takes the form

$$\varphi(\vec{x})=\langle\vec{x}|\varphi\rangle=\underbrace{\langle\vec{x},+|\varphi\rangle}_{\varphi_+(\vec{x})}|+\rangle+\underbrace{\langle\vec{x},-|\varphi\rangle}_{\varphi_-(\vec{x})}|-\rangle\equiv\begin{pmatrix}\varphi_+(\vec{x})\\ \varphi_-(\vec{x})\end{pmatrix}. \tag{5.10}$$

The *Pauli spinors* $\varphi_\pm$ express the probability of finding the particle at the position $\vec{x}$ with spin in the positive $(+)$ and negative $(-)$ z-direction. The normalization of the wave function is obtained from

$$\langle\varphi|\varphi\rangle=\int d^3x|\varphi_+(\vec{x})|^2+\int d^3x|\varphi_-(\vec{x})|^2, \tag{5.11}$$

with $\int d^3x|\varphi_\pm(\vec{x})|^2$ the probability of finding a particle with $s_z=\pm\hbar/2$ in the volume V.

Example 16: Two spin-1/2 particles

Consider a system of two spin-1/2 particles, such an electron and a proton, with Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ of dimension $n_1=2$ and $n_2=2$. Taking into account that for each spin there are two basis states $|+\rangle=\binom{1}{0}$ and $|-\rangle=\binom{0}{1}$, the composite system $\mathcal{H}_1\otimes\mathcal{H}_2$ admits a basis

$$|+,+\rangle=|+\rangle\otimes|+\rangle\,,\quad |+,-\rangle=|+\rangle\otimes|-\rangle\,,\quad |-,+\rangle=|-\rangle\otimes|+\rangle\,,\quad |-,-\rangle=|-\rangle\otimes|-\rangle\,, \tag{5.12}$$

with the first and second entries referring respectively to the z component of the spin for the first and second particle. For instance, in the state $|-,+\rangle$ the first particle has spin $-\hbar/2$, while the second one has spin $+\hbar/2$. The associated bras are denoted by

$$\langle+,+|\,,\qquad \langle+,-|\,,\qquad \langle-,+|\,,\quad \langle-,-|\,,$$

maintaining therefore the original order in the ket, i.e. the first and second entries are still associated with the first and the second particle. The Hilbert $\mathcal{H}_1\otimes\mathcal{H}_2$ and its dual are therefore four-dimensional, $n_1n_2=4$.

Given the above basis states, the most general vector state in $\mathcal{H}_1\otimes\mathcal{H}_2$ takes the form

$$|\varphi\rangle=a|+,+\rangle+b|+,-\rangle+c|-,+\rangle+d|-,-\rangle\,, \tag{5.13}$$

with a, b, c, d complex coefficients satisfying the normalization condition

$$|a|^2 + |b|^2 + |c|^2 + |d|^2 = 1\,.$$

Note that, as stated after Eq. (5.7), is not generically possible to express $|\varphi\rangle$ as the tensor product of two states $|\varphi_1\rangle \in \mathcal{H}_1$ and $|\varphi_2\rangle \in \mathcal{H}_2$. In particular, for (5.13) to be of the form

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \otimes \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_1 b_1 \\ a_1 b_2 \\ a_2 b_1 \\ a_2 b_2 \end{pmatrix}\,, \tag{5.14}$$

or equivalently

$$|\varphi_1\rangle \otimes |\varphi_2\rangle = a_1 b_1 |+,+\rangle + a_1 b_2 |+,-\rangle + a_2 b_1 |-,+\rangle + a_2 b_2 |-,-\rangle\,, \tag{5.15}$$

we should have

$$\frac{a_1 b_1}{a_1 b_2} = \frac{a_2 b_1}{a_2 b_2} \qquad \leftrightarrow \qquad \frac{a}{b} = \frac{c}{d}\,.$$

A set of states that clearly do not satisfy this condition are, for instance, the seminal *Bell or EPR states*,

$$\frac{1}{\sqrt{2}}(|+,-\rangle - |-,+\rangle)\,, \qquad \frac{1}{\sqrt{2}}(|+,+\rangle - |-,-\rangle)\,, \tag{5.16}$$

$$\frac{1}{\sqrt{2}}(|+,+\rangle + |-,-\rangle)\,, \qquad \frac{1}{\sqrt{2}}(|+,-\rangle + |-,+\rangle)\,, \tag{5.17}$$

This orthonormal basis for the Hilbert space of two qubits is indeed instrumental in protocols like quantum teleportation, superdense coding, and entanglement swapping, where the manipulation and measurement of entangled states are key.

If A, B are operators acting respectively on $\mathcal{H}_1$ and $\mathcal{H}_2$, the operator $A \otimes B$ acting on $\mathcal{H}_1 \otimes \mathcal{H}_2$ is characterized by its action on the basis vectors $|i\rangle \otimes |\alpha\rangle$),

$$(A \otimes B)(|i\rangle \otimes |\alpha\rangle) = (A|i\rangle) \otimes (B|\alpha\rangle)\,. \tag{5.18}$$

The corresponding action of $A \otimes B$ on an arbitrary vector $|\varphi\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$ can be determined by expressing $|\varphi\rangle$ in terms of the basis vectors and applying the definition above. In particular, for the tensor product $|\varphi_1, \varphi_2\rangle = |\varphi_1\rangle \otimes |\varphi_2\rangle$, we have

$$\begin{aligned}(A \otimes B)|\varphi_1, \varphi_2\rangle &= \sum_{i,\alpha} a_i b_\alpha (A \otimes B)(|i\rangle \otimes |\alpha\rangle) = \sum_{i,\alpha} a_i b_\alpha (A|i\rangle) \otimes (B|\alpha\rangle) \\ &= \left(\sum_i a_i\, A|i\rangle\right) \otimes \left(\sum_\alpha b_\alpha B|\alpha\rangle\right)\,,\end{aligned} \tag{5.19}$$

or equivalently

$$(A \otimes B)|\varphi_1, \varphi_2\rangle = (A|\varphi_1\rangle) \otimes (B|\varphi_2\rangle)\,. \tag{5.20}$$

In other words, everything acts in the space where it can! Note, however, that in the case of vectors, an arbitrary linear operator on $\mathcal{H}_1 \otimes \mathcal{H}_2$ cannot be generally expressed in the form $A \otimes B$, with $A \in \mathcal{H}_1$ and $B \in \mathcal{H}_2$.

Among the properties of the tensor product of operators, we can highlight the following:

- The tensor product of two operators which are themselves the sum of other operators can be written as a sum of tensor product operators using the usual distributive rule,

$$\begin{aligned} \big(\alpha A + \alpha' B\big) \otimes C &= \alpha(A \otimes C) + \alpha' \left(B \otimes C\right) , \\ A \otimes \big(\beta B + \beta' C\big) &= \beta(A \otimes B) + \beta' \left(A \otimes C\right) , \end{aligned} \tag{5.21}$$

  with $\alpha, \alpha', \beta, \beta' \in \mathbb{C}$.

- The adjoint of a tensor product operator is the tensor product of the adjoint operators, written in the same order relative to the symbol $\otimes$,

$$(A \otimes B)^\dagger = A^\dagger \otimes B^\dagger \, . \tag{5.22}$$

  Note, however, that if the operators on A and B are themselves products of operators, one must reverse their order when taking the adjoint. For instance if $A = A_1 A_2$ and $B = B_1 B_2 B_3$, we have

$$(A_1 A_2 \otimes B_1 B_2 B_3)^\dagger = A_2^\dagger A_1^\dagger \otimes B_3^\dagger B_2^\dagger B_1^\dagger \, . \tag{5.23}$$

- The *ordinary* operator product of two tensor product operators is given by

$$(A \otimes B) \cdot (C \otimes D) = AC \otimes BD \, . \tag{5.24}$$

  Note that the order of the operators is preserved by this operation, with A to the left of C on both sides of the equation, and B to the left of D. Using the distributive law, we can derive an additional rule for the *ordinary* operator product of sums of tensor products, namely

$$(A \otimes B) \cdot (C \otimes D + F \otimes G) = AC \otimes BD + AF \otimes BG \, . \tag{5.25}$$

- The projector onto the tensor product vector $|i\,, \alpha\rangle \equiv |i\rangle \otimes |\alpha\rangle$ is obtained by taking the tensor product of the projectors in $|i\rangle$ and $|\alpha\rangle$, namely

$$|i\,, \alpha\rangle\langle i\,, \alpha| = |i\rangle\langle i| \otimes |\alpha\rangle\langle\alpha| \, , \tag{5.26}$$

  as follows immediately from the definition of the scalar product in $\mathcal{H}_1 \otimes \mathcal{H}_2$, cf. Eq. (5.4).

- The matrix elements of a product operator in the basis $|i\rangle \otimes |\alpha\rangle$ are the products of the matrix elements of the factors in the basis $|i\rangle$ and $|\alpha\rangle$,

$$\langle i, \alpha | A \otimes B | j, \beta \rangle = \langle i | A | j \rangle \cdot \langle \alpha | B | \beta \rangle = A_{ij} B_{\alpha\beta} \, , \tag{5.27}$$

  from which it follows also that the trace of a product operator is the product of the traces of its factors,

$$\mathrm{Tr}[A \otimes B] = \sum_i \langle i | A | i \rangle \cdot \sum_\alpha \langle \alpha | B | \alpha \rangle = \mathrm{Tr}_{\mathcal{H}_1} A \cdot \mathrm{Tr}_{\mathcal{H}_2} B \, , \tag{5.28}$$

  with the subscripts on $\mathrm{Tr}_{\mathcal{H}_1}$ and $\mathrm{Tr}_{\mathcal{H}_2}$ indicating *partial traces* over the subspaces $\mathcal{H}_1$ and $\mathcal{H}_2$.

- An operator A on a Hilbert space $\mathcal{H}_1$ can be naturally extended to an operator $A \otimes \mathbb{I}$ on a product Hilbert space $\mathcal{H}_1 \otimes \mathcal{H}_2$, with $\mathbb{I}$ the identity operator on $\mathcal{H}_2$. In particular, if $|\varphi_1\rangle$ is an eigenvector of A with eigenvalue a, then $|\varphi_1, \varphi_2\rangle$ is an eigenvector of $A \otimes \mathbb{I}$ with the same eigenvalue,

$$(A \otimes \mathbb{I})|\varphi_1, \varphi_2\rangle = (A|\varphi_1\rangle) \otimes |\varphi_2\rangle = a|\varphi_1\rangle \otimes |\varphi_2\rangle = a|\varphi_1, \varphi_2\rangle. \tag{5.29}$$

A simple calculation shows also that an operator $A \otimes \mathbb{I}$ commutes with one the form $\mathbb{I} \otimes B$,

$$\begin{aligned}(A \otimes \mathbb{I})(\mathbb{I} \otimes B)|\varphi_1, \varphi_2\rangle &= (A \otimes \mathbb{I})(|\varphi_1\rangle \otimes (B|\varphi_2\rangle)) = (A|\varphi_1\rangle) \otimes (B|\varphi_2\rangle)\,, \\ (\mathbb{I} \otimes B)(A \otimes \mathbb{I})|\varphi_1, \varphi_2\rangle &= (\mathbb{I} \otimes B)(A|\varphi_1\rangle \otimes |\varphi_2\rangle) = (A|\varphi_1\rangle) \otimes (B|\varphi_2\rangle)\,.\end{aligned} \tag{5.30}$$

reflecting the physical fact that operators that originate from different particles don't know anything about each other! For instance, this property allows to write the Hamiltonian of a system of two independent particles as

$$H = H_1 \otimes \mathbb{I} + \mathbb{I} \otimes H_2\,, \tag{5.31}$$

or the Hamiltonian of a free particle in three dimensions as

$$H = \frac{1}{2m}\left((P_x \otimes I \otimes I)^2 + (I \otimes P_y \otimes I)^2 + (I \otimes I \otimes P_z)^2\right)\,. \tag{5.32}$$

Example 17: Composing two-dimensional Hilbert spaces

Let us verify explicitly the relation (5.20) for two operators A and B represented by 2×2 matrices and two states $|\varphi_1\rangle$ and $|\varphi_2\rangle$ represented by 2×1 column vectors. Expanding these elements into components,

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \qquad B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}, \qquad |\varphi_1\rangle = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \qquad |\varphi_2\rangle = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix},$$

and taking into account the matrix elements of $A \otimes B$ in the $|i, \alpha\rangle$ basis,

$$\langle i', \alpha'|A \otimes B|i, \alpha\rangle = \langle i'|A|i\rangle\langle\alpha'|B|\alpha\rangle = A_{i'i}B_{\alpha'\alpha}\,,$$

we have

$$A \otimes B = \begin{pmatrix} A_{11}B & A_{12}B \\ A_{21}B & A_{22}B \end{pmatrix} = \begin{pmatrix} A_{11}B_{11} & A_{11}B_{12} & A_{12}B_{11} & A_{12}B_{12} \\ A_{11}B_{21} & A_{11}B_{22} & A_{12}B_{21} & A_{12}B_{22} \\ A_{21}B_{11} & A_{21}B_{12} & A_{22}B_{11} & A_{22}B_{12} \\ A_{21}B_{21} & A_{21}B_{22} & A_{22}B_{21} & A_{22}B_{22} \end{pmatrix},$$

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \otimes \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_1b_1 \\ a_1b_2 \\ a_2b_1 \\ a_2b_2 \end{pmatrix},$$

$$(A|\phi\rangle) = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} A_{11}a_1 + A_{12}a_2 \\ A_{21}a_1 + A_{22}a_2 \end{pmatrix},$$

$$(B|\chi\rangle) = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} B_{11}b_1 + B_{12}b_2 \\ B_{21}b_1 + B_{22}b_2 \end{pmatrix} .$$

The left-hand side of the given equation reads therefore

$$\begin{aligned}(A \otimes B)(|\phi\rangle \otimes |\chi\rangle) &= \begin{pmatrix} A_{11}B_{11} & A_{11}B_{12} & A_{12}B_{11} & A_{12}B_{12} \\ A_{11}B_{21} & A_{11}B_{22} & A_{12}B_{21} & A_{12}B_{22} \\ A_{21}B_{11} & A_{21}B_{12} & A_{22}B_{11} & A_{22}B_{12} \\ A_{21}B_{21} & A_{21}B_{22} & A_{22}B_{21} & A_{22}B_{22} \end{pmatrix} \begin{pmatrix} a_1b_1 \\ a_1b_2 \\ a_2b_1 \\ a_2b_2 \end{pmatrix} \\ &= \begin{pmatrix} A_{11}B_{11}a_1b_1 + A_{11}B_{12}a_1b_2 + A_{12}B_{11}a_2b_1 + A_{12}B_{12}a_2b_2 \\ A_{11}B_{21}a_1b_1 + A_{11}B_{22}a_1b_2 + A_{12}B_{21}a_2b_1 + A_{12}B_{22}a_2b_2 \\ A_{21}B_{11}a_1b_1 + A_{21}B_{12}a_1b_2 + A_{22}B_{11}a_2b_1 + A_{22}B_{12}a_2b_2 \\ A_{21}B_{21}a_1b_1 + A_{21}B_{22}a_1b_2 + A_{22}B_{21}a_2b_1 + A_{22}B_{22}a_2b_2 \end{pmatrix} ,\end{aligned}$$

while on the right-hand side we have

$$\begin{aligned}(A|\phi\rangle) \otimes (B|\chi\rangle) &= \begin{pmatrix} A_{11}a_1 + A_{12}a_2 \\ A_{21}a_1 + A_{22}a_2 \end{pmatrix} \otimes \begin{pmatrix} B_{11}b_1 + B_{12}b_2 \\ B_{21}b_1 + B_{22}b_2 \end{pmatrix} \\ &= \begin{pmatrix} (A_{11}a_1 + A_{12}a_2)(B_{11}b_1 + B_{12}b_2) \\ (A_{11}a_1 + A_{12}a_2)(B_{21}b_1 + B_{22}b_2) \\ (A_{21}a_1 + A_{22}a_2)(B_{11}b_1 + B_{12}b_2) \\ (A_{21}a_1 + A_{22}a_2)(B_{21}b_1 + B_{22}b_2) \end{pmatrix} \\ &= \begin{pmatrix} A_{11}B_{11}a_1b_1 + A_{11}B_{12}a_1b_2 + A_{12}B_{11}a_2b_1 + A_{12}B_{12}a_2b_2 \\ A_{11}B_{21}a_1b_1 + A_{11}B_{22}a_1b_2 + A_{12}B_{21}a_2b_1 + A_{12}B_{22}a_2b_2 \\ A_{21}B_{11}a_1b_1 + A_{21}B_{12}a_1b_2 + A_{22}B_{11}a_2b_1 + A_{22}B_{12}a_2b_2 \\ A_{21}B_{21}a_1b_1 + A_{21}B_{22}a_1b_2 + A_{22}B_{21}a_2b_1 + A_{22}B_{22}a_2b_2 \end{pmatrix} .\end{aligned}$$

The left-hand side matches therefore the right-hand side of the equation. The row and column size of $A \otimes B$ is 4×4. The row and column size of $|\varphi_1\rangle \otimes |\varphi_2\rangle$ is 4×1. The row and column size of $(A|\varphi_1\rangle) \otimes (B|\varphi_2\rangle)$ is also 4×1.

We could apply this general result to the system of two spin-1/2 particles introduced in Example 14. For instance, the operator

$$\sigma_z \otimes \sigma_x = \begin{pmatrix} \sigma_x & 0 \\ 0 & -\sigma_x \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix}$$

acts as a Pauli σ_z operator on the first spin, and a Pauli σ_x on the second, namely

$$\begin{aligned} \sigma_z \otimes \sigma_x|+,+\rangle &= |+,-\rangle , & \sigma_z \otimes \sigma_x|+,-\rangle &= |+,+\rangle , \\ \sigma_z \otimes \sigma_x|-,+\rangle &= -|-,-\rangle , & \sigma_z \otimes \sigma_x|-,-\rangle &= -|-,+\rangle . \end{aligned}$$

Be careful with the notation

In many quantum mechanics books, the symbol $\otimes$ indicating the tensor product of states and operators is frequently omitted, leading to expressions of the form $|i\rangle|\alpha\rangle$ or AB. While there is no possible confusion for states, this simplified notation can be ambiguous for operators, being sometimes necessary to specify the specific Hilbert space on which each of the operators act.

5.2 The no-cloning theorem

Throughout this lectures, we have already discover several ways of preparing identical quantum states. One possibility, explored in Sections 2.3 and 4.4.1, is *filtering*, either through a magnetic or electric field. Another one, explored in Section 4.5 is *waiting* for an unstable system to decay to its ground state. However, none of these methods allows to construct a replicas of a prototype state of our choice, as done for instance in classical physics when duplicating a key or copying a computer file.

To build a machine able to provide copies of a quantum state $|\psi\rangle$, we can only make use of unitary operations, since any measurement would inevitably induce the collapse of the original state into an eigenstate of the corresponding observable, corrupting the information contained in it. Given for instance a suitable Hamiltonian determining a unitary operator U and a blank piece of paper $|\chi\rangle$, we could imagine to construct a state like the one we are looking for, namely

$$|\psi\rangle \otimes |\chi\rangle \longrightarrow U(|\psi\rangle \otimes |\chi\rangle) = |\psi\rangle \otimes |\psi\rangle \,. \tag{5.33}$$

Unfortunately there not exist a single universal operator U able to clone a general quantum state constructed of an arbitrary superposition of orthogonal qubits. Indeed, if we apply the same procedure to another state $|\phi\rangle$,

$$|\phi\rangle \otimes |\chi\rangle \longrightarrow U(|\phi\rangle \otimes |\chi\rangle) = |\phi\rangle \otimes |\phi\rangle \,, \tag{5.34}$$

and compute the scalar product with (5.33), we obtain

$$(\langle\phi| \otimes \langle\chi|)U^\dagger U(|\psi\rangle \otimes ||\chi\rangle) = \langle\phi|\psi\rangle = (\langle\phi| \otimes \langle\phi|)(|\psi\rangle \otimes |\psi\rangle) = \langle\phi|\psi\rangle^2 \,, \tag{5.35}$$

which is only possible for $\langle\phi|\psi\rangle = 0$ or ± 1, i.e. if $|\psi\rangle$ and $|\phi\rangle$ are either orthogonal or the same state.[2]

5.3 Density matrix

Realistic quantum mechanical settings typically involve classical statistical uncertainties on top of the unavoidable quantum ones. Let us imagine that we ignore the quantum state of a system, being only able to assign classical probabilities $p_1, p_2, \ldots p_i \ldots$ within an ensemble of

[2] Note that, since classically states are necessarily orthogonal, the quantum no-cloning theorem is not in conflict with classical duplications.

normalized states $\{|\varphi_1\rangle, |\varphi_2\rangle, \dots |\varphi_i\rangle \dots\}$, not necessarily orthogonal to each other.[3] In order to account for this statistical mixture, we can make use of the the so-called density matrix,

$$\rho = \rho^\dagger = \sum_i p_i |\varphi_i\rangle\langle\varphi_i| \,, \qquad \sum_i p_i = 1 \,, \qquad \mathrm{Tr}\,\rho = 1 \,, \tag{5.36}$$

which is nothing else than a sum of projections onto the spaces spanned by the states $|\varphi_i\rangle$, weighted classically by the probabilities $p_i \geq 0$. This object allows us to calculate statistical properties of the system. In particular, given a basis $|n\rangle$ of the Hilbert space under consideration, the expectation value for the measurement of an observable is given by

$$\langle A\rangle = \sum_i p_i \langle\varphi_i|A|\varphi_i\rangle = \sum_{i,n} p_i \langle\varphi_i|n\rangle\langle n|A|\varphi_i\rangle = \sum_{i,n} p_i \langle n|\varphi_i\rangle\langle\varphi_i|A|n\rangle \,, \tag{5.37}$$

which taking into account (5.36) can be written as

$$\langle A\rangle = \sum_n \langle n|\rho A|n\rangle \equiv \mathrm{Tr}(\rho A) \,, \tag{5.38}$$

with the trace extending to all the states in the Hilbert space. This argument applies as well to any function of A, namely

$$\langle f(A)\rangle = \mathrm{Tr}(\rho f(A)) \,. \tag{5.39}$$

Similarly, the probability of obtaining the result a from the measurement of some physical quantity A on a mixed state is given

$$p(a) = \mathrm{Tr}\,(P_a \rho) \,, \qquad \text{with} \qquad P_a = |a\rangle\langle a| \,, \tag{5.40}$$

since

$$\begin{aligned} \mathrm{Tr}\,(P_a\,\rho) &= \sum_i \langle\varphi_i|P_a\,\rho|\varphi_i\rangle = \sum_i \langle\varphi_i|a\rangle\langle a|(\sum_j p_j|\varphi_j\rangle\langle\varphi_j|)|\varphi_i\rangle \\ &= \sum_{ij} p_j \langle\varphi_i|a\rangle\langle a|\varphi_j\rangle\langle\varphi_j|\varphi_i\rangle = \sum_i p_i \langle a|\varphi_i\rangle\langle\varphi_i|a\rangle = \sum_i p_i |\langle a|\varphi_i\rangle|^2 = p(a). \end{aligned} \tag{5.41}$$

[3] Indeed, if the states involved are not orthonormal, the concept of density matrix (a weighted sum of projection operators) still makes sense. In the orthonormal case, each projection operator $|\varphi_i\rangle\langle\varphi_i|$ corresponds to a distinct, non-overlapping pure state, making the density matrix easy to interpret in terms of distinct quantum states. For example, if the system is in a mixed state of two orthogonal states $|\varphi_1\rangle$ and $|\varphi_2\rangle$, the density matrix $\rho = p_1|\varphi_1\rangle\langle\varphi_1| + p_2|\varphi_2\rangle\langle\varphi_2|$ in the basis $\{|\varphi_1\rangle, \varphi_2\rangle\}$ with $\langle\varphi_1|\varphi_2\rangle = 0$ will look like as

$$\rho = \begin{pmatrix} p_1 & 0 \\ 0 & p_2 \end{pmatrix} .$$

In the non-orthonormal case, however, the density matrix in the basis $\{|\varphi_1\rangle, \varphi_2\rangle\}$ is typically not diagonal, since the matrix elements will involve not only the probabilities p_1, p_2 but also the non-zero overlap $\langle\varphi_1|\varphi_2\rangle \neq 0$, complicating the interpretation of the system's mixed state.

Example 18: Density matrix for a beam of electrons

Consider an electron beam with an isotropic distribution of spins,

$$\langle S_x\rangle = \langle S_y\rangle = \langle S_z\rangle = 0\,.$$

Since the spins are isotropically oriented the probabilities of finding S_z with values $\pm\hbar/2$ are both $1/2$. Therefore, in the representation where S_z is diagonal the density matrix (5.36) takes the form

$$\rho = \frac{1}{2}\left(|+\rangle\langle+| + |-\rangle\langle-|\right) = \frac{1}{2}\begin{pmatrix}1\\0\end{pmatrix}(1\,0) + \frac{1}{2}\begin{pmatrix}0\\1\end{pmatrix}(0\,1) = \frac{1}{2}\begin{pmatrix}1&0\\0&1\end{pmatrix}.$$

As can be trivially verified through Eq. (5.38) (note that ρ in this case is just the identity), this choice of ρ is consistent with the unpolarized character of the beam. For instance, we have

$$\langle S_z\rangle = \mathrm{Tr}(\rho S_z) = \mathrm{Tr}(S_z) = 0\,.$$

It is important to note that the density matrix (5.36) is not unique, since it can be put in the form of a linear combination of projectors in different ways. In particular, two density matrices ρ_1 and ρ_2 can be prepared with different probabilities λ and $(1-\lambda)$ and still describe the same state

$$\rho(\lambda) = \lambda\rho_1 + (1-\lambda)\rho_2\,. \tag{5.42}$$

For instance, the state

$$\rho = \frac{1}{2}\left(|+\rangle\langle+| + |-\rangle\langle-|\right)\,, \tag{5.43}$$

in the previous example can be alternatively written as

$$\rho = \lambda|\varphi_1\rangle\langle\varphi_1| + (1-\lambda)|\varphi_2\rangle\langle\varphi_2|\,, \tag{5.44}$$

with $|\varphi_1\rangle$ and $|\varphi_1\rangle$ orthogonal vectors of unit norm related to the $|\pm\rangle$ states by

$$|+\rangle = \sqrt{\lambda}|\varphi_1\rangle + \sqrt{1-\lambda}|\varphi_2\rangle\,, \qquad |-\rangle = \sqrt{\lambda}|\varphi_1\rangle - \sqrt{1-\lambda}|\varphi_2\rangle\,. \tag{5.45}$$

The main exception is, of course, the case in which ρ actually describes a pure state.

5.3.1 Pure vs Mixed States

A particularly nice feature of density matrices is that they allow us to treat pure states and mixed states with the same mathematical formalism. In particular, if one p_i is equal to one and the remaining ones zero, the state is pure. On the contrary, if the system is not in a pure state *and* we know *only* the statistical weights p_i, the system is in a mixed state.

To determine if a given system is in a pure or mixed state, one can consider, for instance, the square of the density matrix,

$$\rho^2 = \rho\,, \quad \text{(Pure State)}\,, \qquad \rho^2 \neq \rho\,, \quad \text{(Mixed State)}\,, \tag{5.46}$$

where we have taken into account that for a pure normalized state $|\varphi\rangle$, $|\varphi\rangle\langle\varphi|$ is a projector, i.e.

$$\rho^2 = |\varphi\rangle\langle\varphi||\varphi\rangle\langle\varphi| = |\varphi\rangle\langle\varphi| \,. \tag{5.47}$$

Alternatively, one could evaluate the so-called *purity* of the density matrix ρ, given by

$$\begin{aligned}\operatorname{Tr}\rho^2 &= \sum_i \sum_j \sum_n p_i p_j \langle n|\varphi_i\rangle\langle\varphi_i|\varphi_j\rangle\langle\varphi_j|n\rangle = \sum_i \sum_j \sum_n p_i p_j \langle\varphi_i|\varphi_j\rangle\langle\varphi_j|n\rangle\langle n|\varphi_i\rangle \\ &= \sum_i \sum_j p_i p_j \langle\varphi_i|\varphi_j\rangle\langle\varphi_j|\varphi_i\rangle = \sum_i \sum_j p_i p_j |\langle\varphi_i|\varphi_j\rangle|^2 \leq \sum_i \sum_j p_i p_j = \left(\sum_i p_i\right)^2 = 1 \,,\end{aligned}$$

or

$$\boxed{\operatorname{Tr}\rho^2 \leq 1 \,,} \tag{5.48}$$

with the equality taking place when only one of the eigenvalues of ρ is one and the rest are zero (pure state), i.e.

$$\boxed{\operatorname{Tr}\rho^2 = 1 \,, \quad \text{(Pure State)} \qquad \operatorname{Tr}\rho^2 < 1 \,, \quad \text{(Mixed State)} \,.} \tag{5.49}$$

In classical settings, the uncertainty inherent to a probability distribution p is associated to a *Shannon entropy*,

$$S(p) = -\sum_i p_i \log p_i \,, \tag{5.50}$$

with the convention $0 \log 0 = 0$. If there is no uncertainty, such as when $p_1 = 1$ and all other $p_i = 0$, then the entropy is $S = 0$. On the other hand, when there are N possible outcomes and no preference for any particular one, meaning $p_i = 1/N$ for each i, the entropy reaches its maximum value of $S = \log N$. In quantum mechanics, a density operator ρ is associated to a *von Neumann entropy*

$$\boxed{S(\rho) = -\operatorname{Tr}\rho \log \rho \,.} \tag{5.51}$$

Note that expression is just the Shannon entropy constructed out of the density matrix eigenvalues, as can be easily verified by expanding ρ in an orthonormal basis. Beyond its fundamental significance in statistical physics, the von Neumann entropy also serves as a measure of entanglement. Among its properties, we can highlight:

- **Non-negativity:** The von Neumann entropy satisfies $S(\rho) \geq 0$.
- **Minimum entropy:** The von Neumann entropy is zero, $S(\rho) = 0$, if and only if ρ represents a pure state.
- **Maximum entropy:** For a system described by an N-dimensional Hilbert space, the von Neumann entropy is maximized at $S = \log N$, i.e. when ρ is the maximally mixed state, $\rho = \mathbb{I}/N$.
- **Concavity:** Given a convex combination of density matrices, $\sum \lambda_i = 1$, the von Neumann entropy obeys the inequality

$$S\left(\sum_i \lambda_i \rho_i\right) \geq \sum_i \lambda_i S(\rho_i) \,. \tag{5.52}$$

This implies that entropy increases as our knowledge about the composition of the state decreases.

- **Subadditivity:** For a composite system with Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$, the von Neumann entropy is generally less than or equal to the sum of its parts, namely

$$S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B)\,, \tag{5.53}$$

with the equality holding if and only if ρ_{AB} is uncorrelated, i.e. if ρ_{AB} can be written as $\rho_{AB} = \rho_A \otimes \rho_B$.

- **Strong subadditivity:** Extending to a tripartite system $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$, entropy obeys

$$S(\rho_{ABC}) + S(\rho_B) \leq S(\rho_{AB}) + S(\rho_{BC})\,. \tag{5.54}$$

This result, though nontrivial to prove, can be understood intuitively by considering the subsystems AB and BC, which share an overlap on B. Strong subadditivity asserts that the total entropy of the combined systems cannot be smaller than the total entropy plus the entropy of their shared part.

Example 19: Density matrix for pure states

For the general spin state (3.33),

$$|\varphi\rangle = \begin{pmatrix} \cos\frac{\theta}{2} \\ e^{i\phi}\sin\frac{\theta}{2} \end{pmatrix},$$

the density matrix takes the form

$$\rho = \begin{pmatrix} \cos\frac{\theta}{2} \\ e^{i\phi}\sin\frac{\theta}{2} \end{pmatrix} \begin{pmatrix} \cos\frac{\theta}{2} & e^{-i\phi}\sin\frac{\theta}{2} \end{pmatrix} = \begin{pmatrix} \cos^2\frac{\theta}{2} & e^{-i\phi}\sin\frac{\theta}{2}\cos\frac{\theta}{2} \\ e^{i\phi}\sin\frac{\theta}{2}\cos\frac{\theta}{2} & \sin^2\frac{\theta}{2} \end{pmatrix}.$$

The associated trace is simply

$$\operatorname{Tr}\rho = \cos^2\frac{\theta}{2} + \sin^2\frac{\theta}{2} = 1\,,$$

and, as expected for a pure state, the eigenvalues λ are 0 and 1, as can be easily verified by explicit computation

$$|\rho - \lambda\mathbb{I}| = 0 = \begin{vmatrix} \cos^2\frac{\theta}{2} - \lambda & e^{-i\phi}\sin\frac{\theta}{2}\cos\frac{\theta}{2} \\ e^{i\phi}\sin\frac{\theta}{2}\cos\frac{\theta}{2} & \sin^2 - \lambda \end{vmatrix},$$

$$\lambda^2 + \cos^2\frac{\theta}{2}\sin^2\frac{\theta}{2} - \lambda\left(\cos^2\frac{\theta}{2} + \sin^2\frac{\theta}{2}\right) - \sin^2\frac{\theta}{2}\cos^2\frac{\theta}{2} = \lambda^2 - \lambda = 0 \to \lambda = 0, 1\,.$$

Example 20: Density matrix for mixed states

Given a system with Hilbert space $\mathbb{C}^2$ and orthonormal basis $\{|+\rangle, |-\rangle\}$, the operator

$$\rho = \frac{3}{4}|+\rangle\langle+| + \frac{1}{4}|-\rangle\langle-| \quad \longrightarrow \quad \rho = \begin{pmatrix} 3/4 & 0 \\ 0 & 1/4 \end{pmatrix}$$

satisfies all the conditions to be a density matrix, i.e. it has unit trace $\mathrm{Tr}\,\rho = 1$ (unit total probability), it is Hermitian $\rho^\dagger = \rho$, and it has semi-positive eigenvalues $\lambda = 3/4, 1/4$. More precisely, since

$$\rho^2 = \rho\rho = \left(\frac{3}{4}|+\rangle\langle+| + \frac{1}{4}|-\rangle\langle-|\right)\left(\frac{3}{4}|+\rangle\langle+| + \frac{1}{4}|-\rangle\langle-|\right) = \frac{9}{16}|+\rangle\langle+| + \frac{1}{16}|-\rangle\langle-| \neq \rho\,,$$

and

$$\mathrm{Tr}\,\rho^2 = \frac{9}{16}\langle+|+\rangle + \frac{1}{16}\langle-|-\rangle = \frac{9}{16} + \frac{1}{16} = \frac{5}{8} < 1\,,$$

it can be interpreted as the density operator for a mixed state, representing an ensemble where 75% of the time the system is in the state $|+\rangle$ and 25% of the time the system is in the state $|-\rangle$.

Example 21: General density matrix in $\mathcal{H} = \mathbb{C}^2$ and the Bloch's sphere

As any other operator in $\mathcal{H} = \mathbb{C}^2$, a two-by-two density matrix can be expressed in terms on the identity operator and the linearly independent Pauli matrices σ_i, namely

$$\rho = \frac{1}{2}(\mathbb{I} + \vec{P}\cdot\vec{\sigma})\,,$$

with $\vec{P} = (P_x, P_y, P_z)$ a real vector giving the position of a point in the unit *ball* $|\vec{P}| \leq 1$ and $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$. Note that this matrix is automatically Hermiticitian, $\rho = \rho^\dagger$, since all the Pauli matrices are Hermitian and the vector $\vec{P}$ is real. On top of that, it is positive definite and has unit trace, as follows directly from the characteristic equation determining to its eigenvalues,

$$\det(\rho - \lambda\mathbb{I}) = \det\left(\frac{1}{2}\begin{pmatrix} 1+P_z & P_x - iP_y \\ P_x + iP_y & 1-P_z \end{pmatrix} - \lambda\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) = 0\,,$$

or equivalently

$$(1 + P_z - 2\lambda)(1 - P_z - 2\lambda) - (P_x - iP_y)(P_x + iP_y) = 0\,.$$

Simplify each term in this expression,

$$(1 + P_z - 2\lambda)(1 - P_z - 2\lambda) = (1-2\lambda)^2 - P_z^2\,,$$

$$(P_x - iP_y)(P_x + iP_y) = P_x^2 + P_y^2\,,$$

and taking into account that $|\vec{P}|^2 = P_x^2 + P_y^2 + P_z^2$, we have $(1-2\lambda)^2 - |\vec{P}|^2 = 0$. Thus, the eigenvalues of the density matrix ρ are

$$\lambda_1 = \frac{1 - |\vec{P}|}{2}, \qquad \lambda_2 = \frac{1 + |\vec{P}|}{2}.$$

Since $0 \leq |\vec{P}| \leq 1$, both of them are non-negative. Also, by inspection, $\operatorname{Tr} \rho = 1$. Since for a pure state one of the previous eigenvalues must be zero and the other one, we must have $|\vec{P}| = 1$. On the other hand, for a mixed state, we must have $|\vec{P}| < 1$. In other words, pure states live on the surface of the Bloch sphere, while mixed states live in its interior, with the center corresponding to the *maximally mixed state*.

5.3.2 Ensembles in statistical physics

In statistical physics, the behavior of systems with many particles is described using various ensembles, each suited to different physical situations. The three main ensembles are the microcanonical, canonical, and grand canonical ensembles, each differing in the constraints imposed on the system.

- **Microcanonical Ensemble**: This ensemble describes an isolated system with fixed energy, volume, and number of particles. The system cannot exchange energy or particles with its surroundings. All accessible states have the same energy, and the probability distribution is uniform across these states. It is most useful for studying closed systems in equilibrium.
- **Canonical Ensemble**: The canonical ensemble is used for systems in thermal contact with a large heat reservoir, meaning the system can exchange energy but not particles with its environment. Here, the number of particles and the volume are fixed, while the energy can fluctuate. The probability of a system being in a particular state is governed by the Boltzmann distribution, which depends on the system's temperature. This ensemble is suitable for systems at constant temperature.
- **Grand Canonical Ensemble**: In this ensemble, the system can exchange both energy and particles with a reservoir, meaning that the number of particles, in addition to the energy, can fluctuate. The system is characterized by a fixed temperature, volume, and chemical potential. This ensemble is ideal for systems in contact with both a heat and particle reservoir, such as gases in open containers.

Each ensemble provides a different statistical description of a system, but they all converge to the same thermodynamic results in the thermodynamic limit (where the number of particles becomes very large). The choice of ensemble depends on the nature of the system's interaction with its surroundings and the constraints that define its equilibrium state.

Example 22: Density matrix for the canonical ensemble

Consider a classical system where each microstate i has an associated energy E_i. In the canonical ensemble, the probability of the system being in the i-th microstate is given by

$$p_i = \frac{1}{Z} e^{-\beta E_i} ,$$

with $\beta = 1/(k_B T)$, $k_B = 1.380649 \times 10^{-23}$ J/ K the Boltzmann constant, T the temperature, and

$$Z = \sum_i e^{-\beta E_i}$$

the so-called partition function, ensuring proper normalization. The equivalent quantum case is described by a Hamiltonian

$$H = \sum_i E_i \left|E_i\right\rangle \left\langle E_i\right| ,$$

and a density matrix

$$\rho = \frac{1}{Z} \sum_i p_i \left|E_i\right\rangle \left\langle E_i\right| = \frac{1}{Z} e^{-\beta H} , \qquad Z = \operatorname{Tr} e^{-\beta H} ,$$

diagonal in the basis $|E_i\rangle$ and with eigenvalues matching the classical canonical probabilities p_i. Note that at low temperatures ($T \to 0$, $\beta \to \infty$) this operator is strongly peaked around the lowest energy state $|E_0\rangle$,

$$\lim_{T\to 0} \rho = \left|E_0\right\rangle \left\langle E_0\right| ,$$

and hence becomes pure. On the contrary, at high energies ($T \to \infty$, $\beta \to 0$), all states contribute democratically,

$$\lim_{T\to \infty} \rho = \frac{1}{\dim \mathcal{H}} \sum_n \left|E_n\right\rangle \left\langle E_n\right| .$$

The partition function Z allows us to compute the expected energy of a thermal state,

$$E = \langle H \rangle = \operatorname{Tr}(\rho H) = \sum_j \langle E_j|\rho H|E_j\rangle = \frac{1}{Z} \sum_{i,j} e^{-\beta E_i/} \langle E_j|H|E_i\rangle\langle E_i|E_j\rangle$$
$$= \frac{1}{Z} \sum_i E_i \, e^{-\beta E_i} = \frac{1}{Z} \sum_i E_i e^{-\beta E_i} = -\frac{1}{Z} \frac{\partial}{\partial \beta} \sum_i e^{-\beta E_i} = -\frac{\partial}{\partial \beta} \ln Z .$$

For instance, a harmonic oscillator with energy levels $E_i = \hbar\omega(i + 1/2)$, we have

$$Z = \sum_{i=0}^{\infty} q^{i+1/2} = \frac{q^{1/2}}{1-q} , \qquad \rho = \frac{1}{Z} \sum_i q^{i+1/2} |E_i\rangle\langle E_i| = (1-q) \sum_i q^i |E_i\rangle\langle E_i| ,$$

with $q = e^{-\beta\hbar\omega}$, and therefore

$$\operatorname{Tr}(\rho H) = -\frac{\partial}{\partial \beta} \ln\left(\frac{q^{1/2}}{1-q}\right) = \frac{\hbar\omega}{2} \frac{1+q}{1-q} .$$

Note that the latest expression can be rewritten as

$$\mathrm{Tr}\,(\rho H) = \frac{\hbar\omega}{2} + \frac{\hbar\omega}{q^{-1}-1}\,,$$

with the first term standing for the ground state energy, which (in the absence of gravity) lacks any physical meaning and can be subtracted away. Doing so, we recover the one-dimensional version of the famous Planck formula for the spectrum of a black body of temperature T, namely

$$E = \frac{\hbar\omega}{q^{-1}-1} = \frac{\hbar\omega}{e^{\hbar\omega/k_BT}-1}\,.$$

Propagators, imaginary time and the emergence temperature

An additional connection between quantum and statistical mechanics appears when considering the propagator (3.234) in imaginary time. In particular, for time-independent Hamiltonians with eigenfunctions $\varphi_i(x)$ satisfying $H\varphi_i(x) = E_i\varphi_i(x)$, a (Wick) transformation $t - t_0 = -i\hbar\beta$, transforms the propagator

$$G(x,t;x',t_0) = \langle x|e^{-iH(t-t_0)/\hbar}|x'\rangle = \sum_i e^{-iE_i(t-t_0)/\hbar}\,\varphi_i(x)\varphi_i^*(x')\,,$$

into the thermal density matrix characterizing a canonical ensemble in the coordinate representation, namely

$$G(x,t_0-i\hbar\beta;x',t_0) = \sum_n e^{-\beta E_i}\,\varphi_i(x)\varphi_i^*(x')\,.$$

5.3.3 Composite systems and partial traces

Given two density matrices ρ_1 and ρ_2, we can construct an operator $\rho_1\otimes\rho_2$, which is Hermitian,

$$(\rho_1\otimes\rho_2)^\dagger = \rho_1^\dagger\otimes\rho_2^\dagger = \rho_1\otimes\rho_2\,, \tag{5.55}$$

has positive-definite eigenvalues $(\rho_1\otimes\rho_2)\,(|i\rangle\otimes|\alpha\rangle) = \lambda\,(|i\rangle\otimes|\alpha\rangle)$,

$$(\rho_1\otimes\rho_2)\,(|i\rangle\otimes|\alpha\rangle) = \rho_1|i\rangle\otimes\rho_2|\alpha\rangle = \lambda_1|i\rangle\otimes\lambda_2|\alpha\rangle = \lambda_1\lambda_2\,(|i\rangle\otimes|\alpha\rangle) \to \lambda = \lambda_1\lambda_2 \geq 0\,, \tag{5.56}$$

and unit trace

$$\mathrm{Tr}\,[\rho_1\otimes\rho_2] = \mathrm{Tr}\,[\rho_1]\otimes\mathrm{Tr}\,[\rho_2] = (1)(1) = 1\,, \tag{5.57}$$

qualifying therefore as a density matrix.

If we were just interested in the properties of the system 1 regardless of what happens with the system 2, we could average over all the possible states of the second system or, in the more standard terminology, *trace it out*. This operation result on the so-called *partial*

trace

$$\mathrm{Tr}_2\,[\rho_1 \otimes \rho_2] = \sum_\gamma \langle\gamma|\rho_1 \otimes \rho_2|\gamma\rangle = \sum_\gamma \rho_1 \otimes \langle\gamma|\rho_2|\gamma\rangle = \rho_1\,, \tag{5.58}$$

where in the last step we have taken into account that the trace of a density matrix is one. The partial trace Tr_2 can be therefore understood as the "inverse" of the tensor product: it provides smaller systems by "ignoring" one or more constituents of a composite system.

Example 23: Partial trace in a composite system of two spin 1/2 particles

Consider a general state in the composite system of two spin 1/2 particles with Hilbert space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 = \mathbb{C}^2 \otimes \mathbb{C}^2$ and dimension $\dim(\mathcal{H}_1 \otimes \mathcal{H}_2) = 2 \cdot 2 = 4$,

$$|\varphi\rangle = \sum_{i,\alpha} r_{i\alpha}|i\rangle \otimes |\alpha\rangle\,.$$

The density matrix associated to the state $|\varphi\rangle$ (pure in $\mathcal{H}$) takes the form

$$\rho = |\varphi\rangle\langle\varphi| = \sum_{i.\alpha,j,\beta} r_{i\alpha} r^*_{j\beta}|i\rangle\langle j| \otimes |\alpha\rangle\langle\beta|\,,$$

or more explicitly

$$\rho = \left(\begin{array}{cc|cc} |r_{00}|^2 & r_{00}r^*_{01} & r_{00}r^*_{10} & r_{00}r^*_{11} \\ r_{01}r^*_{00} & |r_{01}|^2 & r_{01}r^*_{10} & r_{01}r^*_{11} \\ \hline r_{10}r^*_{00} & r_{10}r^*_{01} & |r_{10}|^2 & r_{10}r^*_{11} \\ r_{11}r^*_{00} & r_{11}r^*_{01} & r_{11}r^*_{10} & |r_{11}|^2 \end{array}\right)\,.$$

The partial trace over the states of the second particle provides the density matrix for the first system,

$$\begin{aligned}\rho_1 = \mathrm{Tr}_2\,\rho &= \sum_\gamma \langle\gamma|\rho|\gamma\rangle = \sum_{i.\alpha,j,\beta,\gamma} r_{i\alpha} r^*_{j\beta}|i\rangle\langle j| \otimes \langle\gamma|\alpha\rangle\langle\beta|\gamma\rangle \\ &= \sum_{i.\alpha,j,\beta,\gamma} r_{i\alpha} r^*_{j\beta}|i\rangle\langle j|\delta_{\gamma\alpha}\delta_{\beta\gamma} = \sum_{i,j,\gamma} r_{i\gamma} r^*_{j\gamma}|i\rangle\langle j|\,,\end{aligned}$$

or in matrix form

$$\rho_1 = \sum_\gamma \begin{pmatrix} |r_{00}|^2 + |r_{01}|^2 & r_{00}r^*_{10} + r_{01}r^*_{11} \\ r^*_{00}r_{10} + r^*_{01}r_{11} & |r_{01}|^2 + |r_{11}|^2 \end{pmatrix}\,.$$

A simple comparison of this expression with that for ρ makes clear that the four elements of $\rho_1 = \mathrm{Tr}_2\,\rho$ are simply given by the sum of the diagonal elements of each of the four blocks in which ρ can be cast.

5.3.4 Time Evolution of the density matrix

Using the unitary evolution for each normalized state $\{|\varphi_i\rangle\}$ within the ensemble,

$$|\varphi_i, t\rangle = U(t, t_0)|\varphi_i, t_0\rangle| \,, \tag{5.59}$$

we can easily determine the evolution equation of the density matrix,

$$\begin{aligned} \rho(t) &= \sum_i p_i |\varphi_i, t\rangle\langle\varphi_i, t| = \sum_i U(t, t_0)|\varphi_i, t_0\rangle p_i \langle\varphi_i, t_0|U^\dagger(t, t_0) \\ &= U(t, t_0) \underbrace{\left(\sum_i |\varphi_i, t_0\rangle p_i \langle\varphi_i, 0| \right)}_{\rho(t_0)} U^\dagger(t, t_0) \,, \end{aligned} \tag{5.60}$$

namely

$$\rho(t) = U(t, t_0)\rho(t_0)U^\dagger(t, t_0) \,. \tag{5.61}$$

Differentiating with respect to time and using Eq. (3.194), we obtain the quantum version of the classical Liouville equation for the phase space density, the von Newmann equation

$$\frac{d\rho}{dt} = -\frac{i}{\hbar}[H, \rho] \,. \tag{5.62}$$

Note that, although this equation is formally similar to Eqs. (3.244) and (3.243), it differs by a minus sign following from the different order in the commutator. In anycase, any density matrixwhich depends only on the Hamiltonian H is also independent of time in both the Schrödinger and Heisenberg representations.

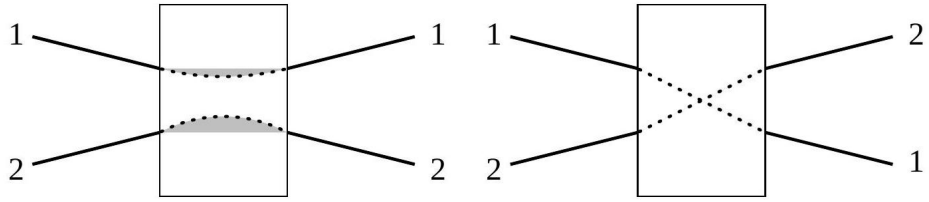
5.4 Identical particles

Postulate VI

The states of a system of identical particles must be either totally symmetric or antisymmetric under the exchange of any pair. Totally symmetric (antisymmetric) states are associated to integer (half-integer) spin [Spin-Statistics theorem].

In classical physics it is always possible to keep track of individual particles even if they are identical, i.e. if all their properties (mass,charge, ...) are the same. In particular, knowing the positions and velocities of a system of particles at a given time, we can assign them labels and follow their individual trajectories in phase space without disturbing the dynamics. The specific labels are of course irrelevant since all the possible configurations obtained by exchanging labels will follow exactly the same dynamics, making all the physical predictions insensitive to labels. Moreover, since there is a continuum of locations, the probability of finding two identical particles in exactly the same place is zero. Identical classical particles do not lose therefore their identity: they remain distinguishable.

Quantum mechanics introduces an extra twist that turns out to be of paramount significance: the uncertainty principle. It becomes impossible to monitor the trajectory of identical particles due to the fundamental inability to determine which particle is which when they come close enough to each other. On top of that, the trajectories of identical particles cannot be observed without modifying their behavior, i.e. even if the position of a particle is precisely known at a specific moment in time, it is impossible to determine its coordinates at a next instant. In quantum mechanics, identical particles lose therefore their individuality. The only measurements we can make are those to determine the probability of one particle being located at x_1, another one at x_2, etc..., but without aiming to distinguish which particle is which.



Example 24: Two spin 1/2 particles

Consider two static spin 1/2 particles lying infinitesimally close to each other, such that the spin is the only relevant label to be considered. In this situation, the state of two particles of opposite spin can be described by the tensor products

$$|+\rangle_{(1)} \otimes |-\rangle_{(2)} \quad \text{or} \quad |-\rangle_{(1)} \otimes |+\rangle_{(2)}\,, \tag{5.63}$$

with the understanding that the first ket is for particle 1 and the second for particle 2. Note that, although these states are orthogonal to each other, we cannot distinguish them by measuring any observable, since they have the same eigenvalues! Indeed, we cannot distinguish them from any linear combination

$$|\varphi\rangle = a|+\rangle_{(1)} \otimes |-\rangle_{(2)} + b|-\rangle_{(1)} \otimes |+\rangle_{(2)} \tag{5.64}$$

with complex coefficients a, b satisfying the normalization condition $|a|^2 + |b|^2 = 1$. This is known as *exchange degeneracy*. Although the physical states are identical, the mathematical descriptions differ due to the labels (1) and (2), which act essentially as quantum numbers, making the particles non-identical from a mathematical perspective. These labels, while physically unobservable, lead not only to distinct states associated with the same eigenvalues but also to different physical predictions. Indeed, given a system initially prepared in the state $|\varphi\rangle$, the probability of finding it in another state $|\chi\rangle$ depends on the coefficients a and b.

There are a priori different ways of dealing with exchange degeneracies. [4] A successful and empirically confirmed approach is to construct the full tensor product Hilbert space as if the particles were distinguishable, to subsequently restrict to states that are either symmetric or antisymmetric under hypothetical permutations of particles. As we will see in what follows, this construction gives rise to fermions and bosons, which, as far as we know, account for all known elementary particles. In other words, particles with mixed symmetry do not seem to exist.

5.4.1 The symmetric group

Since our whole discussion on the indistinguishability of particles rests on the concept of particle exchange, it will useful to consider some of the mathematical properties of permutations.

The symmetric group S_n is the group of permutations of n objects. A permutation $\pi \in S_n$ is usually represented by the two-line notation

$$\pi = \begin{pmatrix} 1 & 2 & \cdots & n \\ p_1 & p_2 & \cdots & p_n \end{pmatrix} , \tag{5.65}$$

understood this as a mapping of the standard ordered integers $1, \ldots, n$ into some arbitrary ordering of them

$$\pi = \begin{pmatrix} 1 & 2 & \cdots & n \\ \downarrow & \downarrow & \downarrow & \downarrow \\ p_1 & p_2 & \cdots & p_n \end{pmatrix} . \tag{5.66}$$

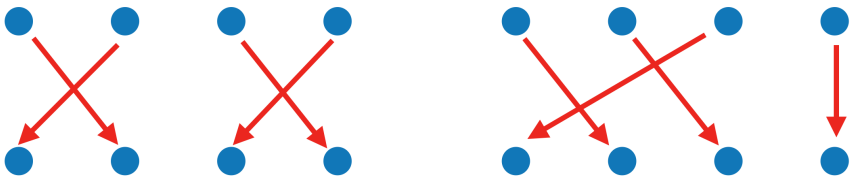
Alternatively, one can use the much more compact notation for cycles or product of cycles. A r-cycle or cycle of length r is written as $(i_1 i_2 \ldots i_r)$ with $i_j \in \{1, 2, \ldots, n\}$, $j = 1, \ldots, r$, and its action given by

$$i_1 \to i_2 \to i_3 \to \ldots \to i_{r-1} \to i_r \to i_1 \tag{5.67}$$

For instance, in S_4, we have

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix} \equiv (12)(34) , \qquad \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix} \equiv (123)(4) = (123) .$$

Finally, one can visualize the action of permutations using a graph where arrows point to where each entry is mapped, namely



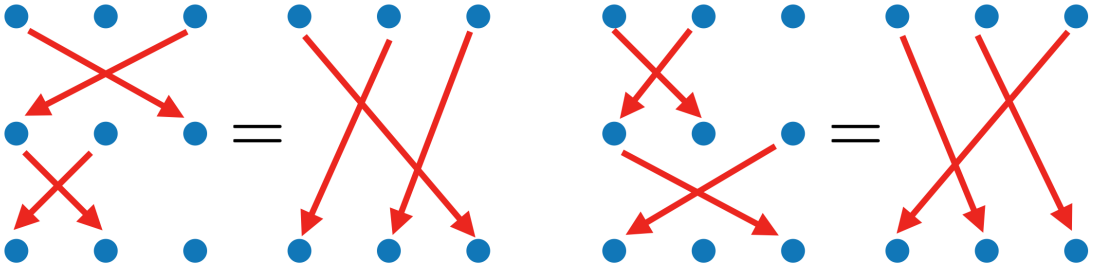
[4] For instance, one possibility, advocated by H.S. Green in 1953, would be to consider para-particles or multi-dimensional representations of the permutation group [H.S. Green, Phys. Rev. 90, 270 (1953)]. However, in this scenario any particle reaction involving a single para-particle is forbidden. This conservation law excludes the possibility that any of the many known particles (leptons, photons, hadrons) are paraparticles [O.W. Greenberg and A.M.L. Messiah, Phys. Rev. 138B, 1155 (1965), *Selection Rules for Para-Fields and the Absence of Para-Particles in Nature*].

From this pictorial approach, it is clear that if cycles are *disjoint* their product can be specified in any order,

$$(12)(34) = (34)(12)\,.$$

The permutations can be composed or multiplied, meaning that if we apply one permutation after the other (from right to left) we get a new permutation. Note, however, that in general permutations do not commute. For instance, we have

$$(12)(13) = (132), \qquad (13)(12) = (123)\,.$$



with the arrows on the graphs following the equal signs obtained by following the arrows in the left ones. Among the properties of cycles, we can also highlight:

- All 1-cycles are equal to the identity element, denoted by e.
- For a given positive integer n, S_n is a finite group with $n!$ elements, e.g.

$$\begin{aligned} S_3 &= \{e, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}\} \\ &= \{e, (12), (23), (13), (123), (132)\}. \end{aligned}$$

- Any rotation of a given cycle specifies the same cycle, e.g.

$$(123) = (231) = (312)\,.$$

- All cycles in S_n, except the identity, can be written as a product of 2-cycles or transpositions (pair exchanges). This result is intuitive since any set of integers can be reordered into an arbitrary configuration by the simple transposition of consecutive labels. The decomposition of a permutation into a product of transpositions is not unique, but it is unique (mod2), meaning that every permutation is either even or odd. A permutation is said to be even (odd) if it can be written as the product of an even (odd) number of transpositions e.g.

$$(123) = (13)(12) = (23)(13) \quad \text{is even}\,,$$

and

$$(13) = (12)(23)(12) = (12)(13)(23) \quad \text{is odd}\,.$$

5.4.2 Bosons and fermions

As we saw in Section 5.1, the Hilbert space of n copies of a single-particle Hilbert space $\mathcal{H}^{(1)}$ is defined by the tensor product

$$\mathcal{H}^{(n)} = \mathcal{H}^{(1)} \otimes \cdots \otimes \mathcal{H}^{(1)} \,. \tag{5.68}$$

As usual, the action of an element $\pi \in S_n$ on $\mathcal{H}^{(n)}$ is defined by its action on the associated basis elements,

$$\pi : |i_1, \ldots, i_n\rangle \mapsto |\pi\,(i_1), \ldots, \pi\,(i_n)\rangle \,, \tag{5.69}$$

and subsequently extended by linearity to the whole $\mathcal{H}^{(n)}$, namely

$$\pi|\varphi\rangle = \sum_{i_1,\ldots,i_n} c_{i_1,\ldots,i_n} \pi|i_1, \ldots, i_n\rangle = \sum_{i_1,\ldots,i_n} c_{i_1,\ldots,i_n} |\pi\,(i_1), \ldots, \pi\,(i_n)\rangle \,, \tag{5.70}$$

with

$$|\varphi\rangle = \sum_{i_1,\ldots,i_n} c_{i_1,\ldots,i_n} |i_1, \ldots, i_n\rangle \tag{5.71}$$

an arbitrary element of $\mathcal{H}^{(n)}$.

Since permutations do not commute, it is impossible to find complete basis of states that are eigenstates of all permutation operators. This doesn't mean, however, that we cannot find *some* states that are simultaneously eigenvectors of all permutation operators:

- Particles whose n-body Hilbert space is that of totally symmetric vectors,

$$\mathrm{Sym}^{(n)}(\mathcal{H}^{(1)}) = \{|\varphi\rangle \in \mathcal{H}^{(n)} : \pi|\varphi\rangle = |\varphi\rangle, \quad \forall \pi \in S_n\} \,, \tag{5.72}$$

  are called *bosons*, after the name of the indian physicist Satyendranath Bose. A bosonic state remains the same independently of how we shuffle around particles, it is totally symmetric. Having this in mind, it is convenient to introduce a unitary symmetrization operator or *symmetrizer*

$$P_S : |\varphi\rangle \mapsto \frac{1}{n!} \sum_{\pi \in S_n} \pi|\varphi\rangle \tag{5.73}$$

  leaving invariant every element $|\varphi\rangle \in \mathrm{Sym}^{(n)}\left(\mathcal{H}^{(1)}\right)$, namely

$$P_S|\varphi\rangle = \frac{1}{n!} \sum_{\pi \in S_n} \pi|\varphi\rangle = \frac{1}{n!} \sum_{\pi \in S_n} |\varphi\rangle = |\varphi\rangle \quad \longrightarrow \quad P_S\, \mathrm{Sym}^{(n)}(\mathcal{H}^{(1)}) = \mathrm{Sym}^{(n)}(\mathcal{H}^{(1)}) \,, \tag{5.74}$$

  where we have taken into account that $\pi|\varphi\rangle = |\varphi\rangle$ together with the fact that S_n has $n!$ elements. This operator is *de facto* the orthogonal projector onto $\mathrm{Sym}^{(n)}(\mathcal{H}^{(1)})$ and therefore we can made use of it to find a set of vector spanning $\mathrm{Sym}^{(n)}(\mathcal{H}^{(1)})$. Indeed, it is easy to see that $P_S^\dagger = P_S$ and $P_S^2 = P_S$, since

$$P_S^\dagger = \frac{1}{n!} \sum_{\pi \in S_n} \pi^\dagger = \frac{1}{n!} \sum_{\pi \in S_n} \underbrace{\pi^{-1}}_{\pi' \equiv \pi^{-1}} = \frac{1}{n!} \underbrace{\sum_{\pi' \in S_n^{-1}} \pi'}_{S_n^{-1} = S_n} = \frac{1}{n!} \sum_{\pi' \in S_n} \pi' = P_S \,. \tag{5.75}$$

and

$$P_S^2 = P_S \frac{1}{n!} \sum_{\sigma \in S_n} \sigma = \frac{1}{n!} \sum_{\sigma \in S_n} P_S\, \sigma = \frac{1}{n!} \sum_{\sigma \in S_n} P_S = P_S\,, \tag{5.76}$$

where we have taken into account that

$$P_S\, \sigma = \frac{1}{n!} \sum_{\pi \in S_n} \underbrace{\pi\sigma}_{\pi' \equiv \pi\sigma} = \frac{1}{n!} \underbrace{\sum_{\pi' \in S_n \sigma}}_{S_n \sigma = S_n} \pi' = \frac{1}{n!} \sum_{\pi' \in S_n} \pi' = P_S \tag{5.77}$$

and the fact that S_n has $n!$ elements.

Example 25: Triplet states

For a system of two static spin 1/2 particles lying infinitesimally close to each other, the symmetric group contains only the identity permutation and a transposition τ swapping the two particles, namely $S_2 = \{e, \tau\}$. As can be easily verified by applying the projector operator P_S, the states

$$|+,+\rangle, \qquad \frac{1}{\sqrt{2}}(|+,-\rangle + |-,+\rangle), \qquad |-,-\rangle\,, \tag{5.78}$$

are examples of totally symmetric vectors,

$$\begin{aligned}
P_S|+,+\rangle &= \frac{1}{n!} \sum_{\pi \in S_n} \pi|+,+\rangle = \frac{1}{2}(e|+,+\rangle + \tau|+,+\rangle) = \frac{1}{2}(|+,+\rangle + |+,+\rangle) = |+,+\rangle\,,\\
P_S|+,-\rangle &= \frac{1}{2}(e|+,-\rangle + \tau|+,-\rangle) = \frac{1}{2}(|+,-\rangle + |-,+\rangle)\,,\\
P_S|-,+\rangle &= \frac{1}{2}(e|-,+\rangle + \tau|-,+\rangle) = \frac{1}{2}(|-,+\rangle + |+,-\rangle) = \frac{1}{2}(|+,-\rangle + |-,+\rangle)\,,\\
P_S|-,-\rangle &= |-,-\rangle\,.
\end{aligned}$$

- Particles whose n-particle Hilbert space is that of totally antisymmetric vectors,

$$\wedge^{(n)}(\mathcal{H}^{(1)}) = \{|\varphi\rangle \in \mathcal{H}^{(n)} : \pi|\varphi\rangle = \mathrm{sgn}(\pi)|\varphi\rangle\}\,, \qquad \mathrm{sgn}(\pi) = \begin{cases} + & \text{even}\,,\\ - & \text{odd}\,, \end{cases} \tag{5.79}$$

are called *fermions*, after the name of the italian physicist Enrico Fermi. Similarly to what we did for bosons, we can introduce a unitary antisymmetrization operator or *antisymmetrizer*

$$P_A : |\varphi\rangle \mapsto \frac{1}{n!} \sum_{\pi \in S_n} \mathrm{sgn}(\pi)\pi|\varphi\rangle \tag{5.80}$$

leaving invariant every element $|\varphi\rangle \in \wedge^{(n)}\left(\mathcal{H}^{(1)}\right)$, namely

$$P_A|\varphi\rangle = \frac{1}{n!}\sum_{\pi\in S_n}\mathrm{sgn}(\pi)\pi|\varphi\rangle = \frac{1}{n!}\sum_{\pi\in S_n}\mathrm{sgn}^2(\pi)|\varphi\rangle = |\varphi\rangle \;\longrightarrow\; P_A\wedge^{(n)}\left(\mathcal{H}^{(1)}\right) = \wedge^{(n)}\left(\mathcal{H}^{(1)}\right), \tag{5.81}$$

where we have taken into account that $\pi|\varphi\rangle = \mathrm{sgn}(\pi)|\varphi\rangle$ together with the fact that S_n has $n!$ elements. Note also that, as happens for the symmetrizer, $P_A^\dagger = P_A$ and $P_A^2 = P_A$.

Example 26: The singlet state

In our beloved system of spin 1/2 particles, the singlet state

$$\frac{1}{\sqrt{2}}(|+,-\rangle - |-,+\rangle) \tag{5.82}$$

is a totally anti-symmetric vector. Indeed, we have

$$\tau\frac{1}{\sqrt{2}}(|+,-\rangle - |-,+\rangle) = \frac{1}{\sqrt{2}}(|-,+\rangle - |+,-\rangle) = -\frac{1}{\sqrt{2}}(|+,-\rangle - |-,+\rangle).$$

Alternatively, we could apply the projector operator P_A. Taking into account for a single transposition $\mathrm{sgn}(\tau) = -1$, we obtain

$$\begin{aligned} P_A|+,-\rangle &= \frac{1}{2}(|+,-\rangle - |-,+\rangle)\,, \\ P_A|-,+\rangle &= \frac{1}{2}(|-,+\rangle - |+,-\rangle) = -\frac{1}{2}(|+,-\rangle - |-,+\rangle)\,, \end{aligned}$$

and therefore

$$P_A\frac{1}{\sqrt{2}}(|+,-\rangle - |-,+\rangle) = \frac{1}{\sqrt{2}}(|+,-\rangle - |-,+\rangle)\,.$$

Note also that

$$\begin{aligned} P_A|+,+\rangle &= \frac{1}{n!}\sum_{\pi\in S_n}\mathrm{sgn}(\pi)\pi|+,+\rangle = \frac{1}{2}(|+,+\rangle - |+,+\rangle) = 0\,, \\ P_A|-,-\rangle &= 0\,. \end{aligned}$$

This illustrates a very important property of identical fermions following directly from the symmetrization postulate: two of them cannot occupy the same single particle state.[a] This is the so-called *Pauli exclusion principle*, which we will prove in full generality in Section 5.4.3.

[a] Note that this (very common) sentence (that we will also use without caring too much) is indeed a bit loose, since particles belonging to an n-particle system do not have a state on its own, only *the system* has a wave function.

When considering a system of identical particles, we are either in $\mathrm{Sym}^{(n)}(\mathcal{H}^{(1)})$ or in $\wedge^{(n)}(\mathcal{H}^{(1)})$. Note, however, that in general for $n > 2$, the n-particle Hilbert space is not completely spanned by purely symmetric or antisymmetric states, i.e.

$$\mathrm{Sym}^{(n)}(\mathcal{H}^{(1)}) \oplus \wedge^{(n)}(\mathcal{H}^{(1)}) \subset \mathcal{H}^{(n)} \,. \tag{5.83}$$

Example 27: Hilbert space reduction

For a system of three particles occupying three distinct states $|a\rangle$, $|b\rangle$, and $|c\rangle$, there are six possible states,

$$|a,b,c\rangle \quad |a,c,b\rangle \,, \quad |b,c,a\rangle \,, \quad |b,a,c\rangle \,, \quad |c,a,b\rangle \,, \quad |c,b,a\rangle \,,$$

but only two possible combinations for bosons and fermions,

$$\begin{aligned} &|a,b,c\rangle + |a,c,b\rangle + |b,c,a\rangle + |b,a,c\rangle + |c,a,b\rangle + |c,b,a\rangle \qquad \text{(Bosons)}\,, \\ &|a,b,c\rangle - |a,c,b\rangle + |b,c,a\rangle - |b,a,c\rangle + |c,a,b\rangle - |c,b,a\rangle \qquad \text{(Fermions)}\,. \end{aligned}$$

The reduction of the size of Hilbert space becomes even more dramatic for increasing n. In general, for n particles filling n distinct states, we have $n!$ states to start with but only one totally symmetric state and one totally antisymmetric state. The other $n! - 2$ states are disregarded.

5.4.3 The Pauli exclusion principle

Imagine a basis element $|i_1, \ldots, i_n\rangle$ in $\mathcal{H}^{(n)}$ where two identical particles j and k with $j \neq k$ occupy the same single-particle state, $i_j = i_k$, such that, by construction, the state is invariant under the transposition τ_{jk} swapping j and k,

$$\tau_{jk}|i_1, \ldots, i_n\rangle = |i_1, \ldots, i_n\rangle. \tag{5.84}$$

From this we easily conclude that the state $|i_1, \ldots, i_n\rangle$ cannot be an element of $\wedge^{(n)}\left(\mathcal{H}^{(1)}\right)$, since if it was, we would have

$$\tau_{jk}|i_1, \ldots, i_n\rangle = \mathrm{sgn}\left(\tau_{jk}\right)|i_1, \ldots, i_n\rangle = -|i_1, \ldots, i_n\rangle, \tag{5.85}$$

contradicting Eq. (5.84). Indeed, for any state $|\varphi\rangle \in \mathcal{H}^{(n)}$ invariant under transpositions, $\tau|\varphi\rangle = |\varphi\rangle$, we have

$$\begin{aligned} P_A\tau =& \frac{1}{n!}\sum_{\sigma\in S_n}\mathrm{sgn}(\sigma)\sigma\tau = \frac{1}{n!}\sum_{\sigma\in S_n}\mathrm{sgn}\left(\sigma\tau\tau^{-1}\right)\sigma\tau = \underbrace{\mathrm{sgn}\left(\tau^{-1}\right)}_{=-1}\frac{1}{n!}\sum_{\sigma\in S_n}\mathrm{sgn}(\sigma\tau)\underbrace{\sigma\tau}_{\sigma'\equiv\sigma\tau} \\ =& -\frac{1}{n!}\underbrace{\sum_{\sigma'\in\tau S_n}}_{\tau S_n = S_n}\mathrm{sgn}(\sigma')\sigma' = -\frac{1}{n!}\sum_{\sigma'\in S_n}\mathrm{sgn}(\sigma')\sigma' = -P_A\,, \end{aligned} \tag{5.86}$$

and therefore

$$P_A|\varphi\rangle = P_A\tau|\varphi\rangle = -P_A|\varphi\rangle\,, \qquad \longrightarrow \qquad P_A|\varphi\rangle = 0\,, \tag{5.87}$$

meaning that $|i_1, \ldots, i_n\rangle$ is orthogonal to $\wedge^{(n)}\left(\mathcal{H}^{(1)}\right)$ whenever two particles occupy the same single particle state.

The Spin-Statistics theorem

Within non-relativistic quantum mechanics, the correlation between spin and statistics can be seen as an empirical law. However, the spin-statistics relation emerges naturally from the unification of quantum mechanics and special relativity (i.e. from quantum field theory). The so-called Spin-Statistics theorem, formulated by the physicist Julian Schwinger, Sin-Itiro Tomonaga, and Richard P. Feynman in the late 1940s, states that the statistics of a particle (Bose-Einstein or Fermi-Dirac) is directly determined by its spin.

- Fermions, with half-integer spin values, obey Fermi-Dirac statistics, which dictate that no two identical fermions can occupy the same quantum state simultaneously.
- Bosons, with integer spin values, obey Bose-Einstein statistics, which means any number of them can occupy the same quantum state.

The assumptions in the theorem are: (1) Lorentz invariance, (2) causality, and (3) unitarity (i.e., positivity of the norms).

A corollary of the symmetrization postulate and the spin-statistics theorem is that there are only fermions and bosons in Nature. In the Standard Model of particle physics, these particles play indeed a fundamental role.

5.4.4 Wave-function of n identical non-interacting particles with spin

The Schrödinger equation for n non-interacting particles with spin,

$$i\hbar\frac{\partial}{\partial t}\varphi(\vec{r}_1, \vec{S}_1; \vec{r}_2, \vec{S}_2; \ldots; \vec{r}_n, \vec{S}_n) = H\varphi(\vec{r}_1, \vec{S}_1; \vec{r}_2, \vec{S}_2; \ldots; \vec{r}_n, \vec{S}_n)\,, \tag{5.88}$$

can be written as the product of a spatial part and a spin part,

$$\varphi(\vec{r}_1, \vec{S}_1; \vec{r}_2, \vec{S}_2; \ldots; \vec{r}_n, \vec{S}_n) = \langle \vec{r}_1, \vec{r}_2, \ldots, \vec{r}_n|\varphi\rangle = \phi\left(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_n\right) \otimes |\vec{S}_1, \vec{S}_2, \ldots, \vec{S}_n\rangle\,. \tag{5.89}$$

For identical bosons, the overall wave function must be symmetric; forcing the spatial and spin parts to be simultaneously symmetric or antisymmetric

$$\varphi_S(\vec{r}_1, \vec{S}_1; \vec{r}_2, \vec{S}_2; \ldots; \vec{r}_n, \vec{S}_n) = \begin{cases} \phi_S\left(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_n\right) \otimes |\vec{S}_1, \vec{S}_2, \ldots, \vec{S}_n\rangle_S\,, \\ \phi_A\left(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_n\right) \otimes |\vec{S}_1, \vec{S}_2, \ldots, \vec{S}_n\rangle_A\,. \end{cases} \tag{5.90}$$

On the other hand, if the particles identical fermions, the spatial and spin parts must have opposite parities, resulting in an overall antisymmetric wave function,

$$\varphi_S(\vec{r}_1, \vec{S}_1; \vec{r}_2, \vec{S}_2; \ldots; \vec{r}_n, \vec{S}_n) = \begin{cases} \phi_A\left(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_n\right) \otimes |\vec{S}_1, \vec{S}_2, \ldots, \vec{S}_n\rangle_S\,, \\ \phi_S\left(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_n\right) \otimes |\vec{S}_1, \vec{S}_2, \ldots, \vec{S}_n\rangle_A\,. \end{cases} \tag{5.91}$$

Example 28: Atoms with two non-interacting electrons

Accounting explicitly for spin degrees of freedom, but neglecting the the interaction between electrons and the spin-orbit interaction, the Hamiltonian of a two-electron atom, like the Helium atom, the singly ionized Lithium ion Li^+ or the doubly ionized Berilium atom Be^{2+} is given by

$$H = H_1 + H_2 = \left(\frac{\vec{p}_1^{\,2}}{2m} - \frac{Ze^2}{r_1}\right) + \left(\frac{\vec{p}_2^{\,2}}{2m} - \frac{Ze^2}{r_2}\right) \tag{5.92}$$

with $m = Mm_e/\left(M + m_e\right)$ the reduced mass of the system, M the mass of the nucleus, m_e the mass of the electrons and $\vec{r}_1$ and $\vec{r}_2$ the location of the electrons with respect to the nucleus, which is placed at the origin of coordinates. The corresponding time-independent Schrödinger equation is given by

$$\left[H_1 + H_2\right]\varphi(\vec{r}_1, \vec{S}_1; \vec{r}_2, \vec{S}_2) = E_{n_1 n_2}\varphi(\vec{r}_1, \vec{S}_1; \vec{r}_2, \vec{S}_2) \tag{5.93}$$

with

$$E_{n_1 n_2} = E_{n_1} + E_{n_2} = -\frac{Z^2 e^2}{2a_0}\left(\frac{1}{n_1^2} + \frac{1}{n_2^2}\right), \tag{5.94}$$

the sum of the energies of the electrons, $a_0 = \hbar^2/\left(me^2\right)$ the Bohr radius and $e^2/(2a_0) = 13.6\,\text{eV}$. The associated wave function is given by

$$\varphi(\vec{r}_1, \vec{S}_1; \vec{r}_2, \vec{S}_2) = \phi\left(\vec{r}_1, \vec{r}_2\right) \otimes |\vec{S}_1, \vec{S}_2\rangle \tag{5.95}$$

Since this system consists of two identical fermions, its wave function must be anti-symmetric. Therefore, either the spatial part is antisymmetric and the spin part is symmetric,

$$\varphi(\vec{r}_1, \vec{S}_1; \vec{r}_2, \vec{S}_2) = \frac{1}{\sqrt{2}}[\phi_{n_1 l_1 m_1}(\vec{r}_1)\phi_{n_2 l_2 m_2}(\vec{r}_2) - \phi_{n_2 l_2 m_2}(\vec{r}_1)\phi_{n_1 l_1 m_1}(\vec{r}_2)]|\vec{S}_1, \vec{S}_2\rangle_S, \tag{5.96}$$

or the spatial part is symmetric and the spin part is antisymmetric,

$$\varphi(\vec{r}_1, \vec{S}_1; \vec{r}_2, \vec{S}_2) = \phi_{n_1 l_1 m_1}(\vec{r}_1)\phi_{n_2 l_2 m_2}(\vec{r}_2)|\vec{S}_1, \vec{S}_2\rangle_A\,, \tag{5.97}$$

with $|\vec{S}_1, \vec{S}_2\rangle_S$ and $|\vec{S}_1, \vec{S}_2\rangle_A$ given respectively by the symmetric triplet states in (5.78) and the antisymmetric singlet (5.82). The ground state corresponds to both electrons occupying the lowest state $n = 1$, $l = 0$, $m = 0$, i.e., $n_1 = n_2 = 1$. The associated energy and wave function are given by

$$E_{11} = -2\frac{Z^2 e^2}{2a_0} = -27.2\,Z^2\,\text{eV}\,, \qquad \varphi_{11}(\vec{r}_1, \vec{S}_1; \vec{r}_2, \vec{S}_2) = \phi_{100}\left(\vec{r}_1\right)\phi_{100}\left(\vec{r}_2\right)|\vec{S}_1, \vec{S}_2\rangle_A\,, \tag{5.98}$$

with

$$\phi_{100}(\vec{r}) = R_{10}(r)Y_{00}(\Omega) = \frac{1}{\sqrt{\pi}}\left(\frac{Z}{a_0}\right)^{3/2} e^{-\frac{Zr}{a_0}}\,. \tag{5.99}$$

In the first excited state, one electron occupies the lowest level $n = 1$, $l = 0$, $m = 0$ and the other one occupies the level $n = 2$, $l = 0$, $m = 0$. This corresponds either

to $n_1 = 1, n_2 = 2$ or to $n_1 = 2, n_2 = 1$. The associated energy and wave function are given by

$$\begin{aligned} E_{12} &= -\frac{Z^2 e^2}{2a_0}\left(1+\frac{1}{4}\right) = -17.0\, Z^2 \text{ eV}\,, \\ \varphi_{12}(\vec{r}_1, \vec{S}_1; \vec{r}_2, \vec{S}_2) &= \frac{1}{\sqrt{2}}\left[\phi_{100}(\vec{r}_1)\,\phi_{200}(\vec{r}_2) - \phi_{200}(\vec{r}_1)\,\phi_{100}(\vec{r}_2)\right] |\vec{S}_1, \vec{S}_2\rangle_S\,, \end{aligned} \tag{5.100}$$

with

$$\phi_{200}(\vec{r}) = R_{20}(r)Y_{00}(\Omega) = \frac{1}{\sqrt{8\pi}}\left(\frac{Z}{a_0}\right)^{3/2}\left(1-\frac{Zr}{2a_0}\right)e^{-\frac{Zr}{2a_0}}\,. \tag{5.101}$$

Note that these theoretical results fail miserably to describe the experimental values. For instance, according to Eq. (5.98), the ground state energy of helium ($Z = 2$) should be $E_{11,\text{th}} \simeq -108.8\,\text{eV}$, while the experimental value is rather a $\sim 40\%$ lower, $E_{11,\text{exp}} = -78.97\,\text{eV}$. A more accurate determination of the energy levels will be obtained in Chapter 6, where we will explicitly account for for the Coulomb force between electrons using a variational and perturbation theory approach.

5.4.5 The Slater determinant and the permanent

The use of the projection and antisymmetrization operators to built bosonic and fermionic states can become rather complicated, since the procedure requires a sufficiently efficient way of summing over permutations. There is a case, however, where things can be phrased more compactly.

Given a collection of n *non-interacting* particles,[5] we can easily construct a multi-particle wave function for fermions that is anti-symmetric under the exchange of any pair of particles by using the Slater determinant

$$|\varphi\rangle = \frac{1}{\sqrt{n!}}\begin{vmatrix} |\varphi_{n_1}(\vec{r}_1)\rangle & |\varphi_{n_1}(\vec{r}_2)\rangle & \dots & |\varphi_{n_1}(\vec{r}_n)\rangle \\ |\varphi_{n_2}(\vec{r}_1)\rangle & |\varphi_{n_2}(\vec{r}_2)\rangle & \dots & |\varphi_{n_2}(\vec{r}_n)\rangle \\ \vdots & & \ddots & \\ |\varphi_{n_n}(\vec{r}_1)\rangle & |\varphi_{n_n}(\vec{r}_2)\rangle & \dots & |\varphi_{n_n}(\vec{r}_n)\rangle \end{vmatrix}\,, \tag{5.102}$$

with $\varphi_{n_i}(\vec{r}_j)$ the single-particle wave functions, n_i collectively denoting the different eigenvalues and the overall factor of $1/\sqrt{n!}$ ensuring that the resulting state is normalised.[6] Note that different rows are associated to different eigenvalues, while different columns correspond to different $\vec{r}_i$ assignments In fact, swapping two particles is tantamount to swapping two columns in the determinant, changing it by a minus sign. Also, if two particles sit in the same state (e.g. $n_1 = n_2$), the corresponding rows become linearly dependent and the determinant vanishes, ensuring the Pauli exclusion principle. Interestingly enough, a multi-particle

[5] Even though the state derived using the Slater determinant is only exact for non-interacting fermions, it provides a useful starting point for the study of weakly interacting systems in a perturbative scheme.

[6] If you have doubts about how this operates, consider writing and expanding the Slater determinant for $n = 3$ and verify that the result is indeed fully antisymmetric under the particles' interchanges.

wave function for bosons can be similarly obtained by expanding all terms in the Slater determinant and setting all relative signs positive. The resulting object is called the *Slater permanent*.

Example 29: Three and non-interacting particles in a box with spin $+\hbar/2$

To illustrate the use of Slater determinants, let us consider the ground, first excited and second excited state of a system of three identical and non-interacting spin-1/2 particles, all of them in the same spin state $|+\rangle$ and confined to move in a one-dimensional infinite potential well of length L, namely

$$V(x) = \begin{cases} 0, & 0 < x < L \\ \infty, & \text{otherwise}\,. \end{cases}$$

To this end, we recall the single-particle energy and wave function of a particle confined within an infinite potential well,

$$\varepsilon_n = \frac{n^2\hbar^2\pi^2}{2mL^2}, \quad n = 1, 2, 3, \dots\,, \qquad \varphi_n(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right)\,.$$

Since the three particles considered are identical spin 1/2 fermions, the *total* wave function of the system must be completely antisymmetric. Furthermore, as all of them share the same spin projection ($+\hbar/2$) and the Pauli exclusion principle dictates that no two particles can occupy the same quantum state, each energy level can accommodate at most one particle. One must therefore have,

$$E = \frac{\hbar^2\pi^2}{2mL^2}\left(n_1^2 + n_2^2 + n_3^2\right)\,, \qquad n_1 \neq n_2 \neq n_3\,.$$

Ground State: In the ground state, the three particles occupy the lowest available energy levels, $n = 1, 2, 3$, with only one particle per level. The total energy of the ground state is therefore

$$E^{(0)} = \frac{\hbar^2\pi^2}{2mL^2}\left(1^2 + 2^2 + 3^2\right) = \frac{7\hbar^2\pi^2}{mL^2}\,.$$

Given that the *total* wave function of the system, $\varphi^{(0)}(x_1, x_2, x_3)$, must be antisymmetric and the spin part is already fully symmetric,

$$|+++\rangle = |+\rangle \otimes |+\rangle \otimes |+\rangle\,,$$

the corresponding spatial part is necessarily antisymmetric. Making use of the Slater determinant, we can write

$$\varphi^{(0)}\left(x_1, x_2, x_3\right) = \frac{1}{\sqrt{3!}} \begin{vmatrix} \varphi_1(x_1) & \varphi_1(x_2) & \varphi_1(x_3) \\ \varphi_2(x_1) & \varphi_2(x_2) & \varphi_2(x_3) \\ \varphi_3(x_1) & \varphi_3(x_2) & \varphi_3(x_3) \end{vmatrix} \otimes |+++\rangle\,.$$

First Excited State: In the first excited state, the third particle is promoted from $n = 3$ to $n = 4$, with the other particles remaining in $n = 1$ and $n = 2$. Substituting the values, we get

$$E^{(1)} = \frac{\hbar^2\pi^2}{2mL^2}\left(1^2 + 2^2 + 4^2\right) = \frac{21\hbar^2\pi^2}{2mL^2}\,.$$

Following the same reasoning as before, the corresponding *total* wave function takes the antisymmetric form

$$\varphi^{(1)}\left(x_1, x_2, x_3\right) = \frac{1}{\sqrt{3!}}\begin{vmatrix} \varphi_1\left(x_1\right) & \varphi_1\left(x_2\right) & \varphi_1\left(x_3\right) \\ \varphi_2\left(x_1\right) & \varphi_2\left(x_2\right) & \varphi_2\left(x_3\right) \\ \varphi_4\left(x_1\right) & \varphi_4\left(x_2\right) & \varphi_4\left(x_3\right) \end{vmatrix} \otimes |+++\rangle\,.$$

Second Excited State: In the second excited state, the second particle is promoted from $n = 2$ to $n = 3$, while the other particles occupy $n = 1$ and $n = 4$. The total energy and *total* wave function are therefore

$$E^{(2)} = \frac{\hbar^2\pi^2}{2mL^2}\left(1^2 + 3^2 + 4^2\right) = \frac{13\hbar^2\pi^2}{mL^2}\,,$$

and

$$\varphi^{(2)}\left(x_1, x_2, x_3\right) = \frac{1}{\sqrt{3!}}\begin{vmatrix} \varphi_1\left(x_1\right) & \varphi_1\left(x_2\right) & \varphi_1\left(x_3\right) \\ \varphi_3\left(x_1\right) & \varphi_3\left(x_2\right) & \varphi_3\left(x_3\right) \\ \varphi_4\left(x_1\right) & \varphi_4\left(x_2\right) & \varphi_4\left(x_3\right) \end{vmatrix} \otimes |+++\rangle\,.$$

5.4.6 Overlappings and correlations

If the symmetrization postulate were always necessary, the study of a system with a restricted number of particles would become an impossible task, since one would need to consider all identical particles in the entire Universe. However, under specific circumstances, identical particles may behave as if they were actually different. Consider for instance two identical particles with wave functions (spatial parts) $\varphi_1(\vec{x})$ and $\varphi_2(\vec{x})$. According to the symmetrization postulate, the corresponding system state must be either symmetric ($S \to +$) or antisymmetric ($A \to -$),

$$\varphi_{S/A}(\vec{x}_1, \vec{x}_2) = N\left[\varphi_1(\vec{x}_1)\varphi_2(\vec{x}_2) \pm \varphi_1(\vec{x}_2)\varphi_2(\vec{x}_1)\right]\,, \tag{5.103}$$

with N a normalization factor. Given these expressions, the probability density of finding a particle in a given region of space is *a priori* correlated with the position of the other one,

$$\begin{aligned} |\varphi_{S/A}(\vec{x}_1, \vec{x}_2)|^2 = & N^2\left[|\varphi_1\left(\vec{x}_1\right)|^2|\varphi_2\left(\vec{x}_2\right)|^2 + |\varphi_1\left(\vec{x}_2\right)|^2|\varphi_2\left(\vec{x}_1\right)|^2\right. \\ & \left. \pm 2\,\mathrm{Re}\left(\varphi_1\left(\vec{x}_1\right)\varphi_2\left(\vec{x}_2\right)\varphi_1^*\left(\vec{x}_2\right)\varphi_2^*\left(\vec{x}_1\right)\right)\right]\,, \end{aligned} \tag{5.104}$$

with the last-term the so-called *exchange density*. This correlation becomes, however, negligible for non-overlapping wave functions, $\varphi_1(x_2) \sim 0$ and $\varphi_2(x_1) \sim 0$. In this limit, the probability density of finding one particle at $\vec{x} = \vec{x}_1 \sim \langle x_1\rangle$ and the other at $\vec{x} = \vec{x}_2 \sim \langle x_2\rangle$ becomes simply the join probability density expected for classical particles,

$$|\varphi_{S/A}\left(\vec{x}_1, \vec{x}_2\right)|^2 = N^2|\varphi_1\left(\vec{x}_1\right)|^2|\varphi_2\left(\vec{x}_2\right)|^2\,. \tag{5.105}$$

Then, the question of identity does not arise. There is no need of symmetrizing or antisymmetrizing the states of non-overlapping localized particles.

The situation changes completely when the two particles start approaching each other. To illustrate this explicitly, let us consider a simplified one-dimensional scenario where two identical particles of mass m are well-localized around the points $x = +a$ and $x = -a$, with approximately Gaussian single-particle wave functions

$$\varphi_1(x) = \left(\frac{1}{\pi a_0^2}\right)^{1/4} e^{-\beta(x-a)^2}, \qquad \varphi_2(x) = \left(\frac{1}{\pi a_0^2}\right)^{1/4} e^{-\beta(x+a)^2}, \tag{5.106}$$

and $2\beta a^2 \gg 1$ (or equivalently $(\Delta x)^2_{\pm a} \ll a^2$). The two-particle wave function of the composite system takes the form (5.103) with

$$N = \frac{1}{\sqrt{2}}\left(1 \pm I^2\right)^{-1/2} = \frac{1}{\sqrt{2}}\left(1 \pm e^{-2\beta a^2}\right)^{-1/2} \tag{5.107}$$

and

$$I = \int_{-\infty}^{\infty} dx\, \varphi_1(x)\varphi_2(x) = e^{-\beta a^2} \tag{5.108}$$

a function quantifying the overlap of single-particle wave functions. Taking into account that the Hamiltonian of the system of the two non-interacting identical particles of mass m is just

$$H = \frac{p_1^2}{2m} + \frac{p_2^2}{2m}, \tag{5.109}$$

the expectation value of the energy in the above state becomes

$$\mathcal{E} = 4|N|^2(E \pm \Gamma I) = \left(\frac{\hbar^2\beta}{2m}\right)\frac{1 \pm e^{-2\beta a^2}\left(1 - 2\beta a^2\right)}{1 \pm e^{-2\beta a^2}}, \tag{5.110}$$

with

$$E = -\frac{\hbar^2}{2m}\int_{-\infty}^{\infty} dx \varphi_i(x)\varphi_i''(x) = \frac{\hbar^2\beta}{4m} \tag{5.111}$$

the energy expectation value in either of the localized one-particle states $i = 1, 2$, and Γ the mixed matrix element ($i = 1, 2$ and $j \neq i$)

$$\Gamma = -\frac{\hbar^2}{2m}\int_{-\infty}^{\infty} dx \varphi_i(x)\varphi_j''(x) = \frac{\hbar^2\beta}{4m}\left(1 - 2\beta a^2\right)e^{-\beta a^2}. \tag{5.112}$$

As explicit in Eq. (5.110), a variation in the distance a between the two particles results in a change in the energy of the system, leading with it to an effective force $F \equiv -\mathcal{E}/\partial a$ between them. If the particles are fermions, we have

$$F = \frac{2\hbar^2\beta^2 a}{m} e^{-2\beta a^2}\frac{-1 + 2\beta a^2 + e^{-2\beta a^2}}{\left(1 - e^{-2\beta a^2}\right)^2},$$

which is *always* positive since $e^{-2\beta a^2} + 2\beta a^2 - 1 > 0$. This *repulsive force* arises due to the antisymmetry of the total wave function or equivalently to the Pauli exclusion principle. Fermions cannot occupy the same quantum state, leading to a repulsive force that increases

as the particles get closer, preventing them from collapsing into the same region of space. In the case of two bosons, the situation is more complex. The effective force

$$F = \frac{2\hbar^2\beta^2 a}{m} e^{-2\beta a^2} \frac{e^{-2\beta a^2} + 1 - 2a^2\beta}{\left(1 + e^{-2\beta a^2}\right)^2}$$

is attractive at large separations $\left(2\beta a^2 \gg 1\right)$, where the overlap between the wave functions becomes negligible, and repulsive at very short separations $a \leq 0.8\beta^{-1/2}$. However, contrary to the case of fermions, this repulsion is not associated to the Pauli exclusion principle but rather to the increase of kinetic energy as the particles come closer to each other.

Example 30: Fermi Gases and white dwarfs

An important consequence of the Pauli exclusion principle is the existence of high-density stars known as white dwarfs. Within these objects the force preventing collapse is not provided by nuclear reactions but rather by the large numbers of electrons present in the star's material and their impossibility to occupy the same states. This, in turn, generates a strong repulsive force among the electrons, ensuring the star doesn't collapse under its gravity.

A simple white dwarf can be modelled as a collection of ordinary matter (nucleons and electrons) confined to a finite radius by some kind of a potential well. For the sake of simplicity, let us describe this potential as an infinite square well in 3-dimensions, even if the star has of course a spherical shape.[a] The associated energy levels are given by

$$E = \frac{\pi^2\hbar^2}{2mL^2}\left(n_x^2 + n_y^2 + n_z^2\right) , \tag{5.113}$$

with L the size of any of the box sides and n_x, n_y and n_z positive-integer quantum numbers, which for large numbers of electrons can be interpreted as the components of a vector $\vec{n}$,

$$E = \frac{\pi^2\hbar^2}{2m_eL^2}|\vec{n}|^2 . \tag{5.114}$$

If the temperature of the star is sufficiently low, all the electrons will occupy the lowest available energy states, filling a sphere of radius $n_{\max}$. This gives

$$N_e = 2 \times \frac{1}{8} \times \frac{4}{3}\pi n_{\max}^3 = \frac{\pi}{3} n_{\max}^3 , \tag{5.115}$$

with the factor 2 due to the fact two electrons can occupy a state with given n_x, n_y and n_z and the factor 1/8 following from the fact that the n_i can take only positive values, so that only 1/8 of the sphere actually represents distinct physical states. The energy of the highest-energy electrons, E_{F}, is known as the *Fermi energy*. Taking into account (5.115) this can be written as[b]

$$E_{\mathrm{F}} = \frac{\pi^2\hbar^2}{2m_eL^2} n_{\max}^2 = \frac{\pi^2\hbar^2}{2m_eL^2}\left(\frac{3N_e}{\pi}\right)^{2/3} = \frac{\hbar^2}{2m_e}\left(3\pi^2 n_e\right)^{2/3} . \tag{5.116}$$

with $n_e = N_e/V$ the electron density and $V = L^3$ the volume of the box. The associated wave number

$$k_F = \left(3\pi^2 n_e\right)^{1/3} \tag{5.117}$$

corresponds to a de Bloglie wavelenght $\lambda_F = 2\pi/k_F = 2.03 n_e^{-1/3}$. Taking into account that $n_e^{-1/3}$ is approximately the interparticle spacing, this implies

$$d \simeq \frac{\lambda_F}{2}\,, \tag{5.118}$$

meaning essentially that the exclusion principle forces electrons to be at least a half-wave apart.
The total energy of the physically distinct states is given by

$$E_e = \frac{1}{8}\int_0^{n_{\max}} dn\; 4\pi n^2 \frac{\pi^2\hbar^2}{2mL^2} n^2 = \frac{\pi^3\hbar^2}{10 m_e L^2} n_{\max}^5 = \frac{3}{5} N_e E_{\rm F}\,. \tag{5.119}$$

This energy can be evaluated alongside the gravitational potential energy E_g of the star. Assuming a uniform mass density ρ with $M(r) = (4/3)\pi r^3 \rho$, this is given by

$$E_g = -\int_0^R dr\, 4\pi r^2 \rho \frac{GM(r)}{r} = -\frac{16\pi^2}{3} G\rho^2 \int_0^R dr r^4 = -\frac{16\pi^2}{15} G\rho^2 R^5 = -\frac{3GM^2}{5R}. \tag{5.120}$$

Rewriting this result in terms of the volume $V = 4\pi R^3/3$ and taking into account that most of the star's mass is in the nucleons (namely $M \simeq N_n M_n$ with N_n the total number of nucleons and M_n the single nucleon mass), we obtain

$$E_g = -\frac{3}{5} G N_n^2 M_n^2 \left(\frac{4\pi}{3V}\right)^{1/3}. \tag{5.121}$$

Taking into account this expression together with Eq. (5.116), we can determine the equilibrium configuration from

$$\frac{\partial}{\partial V}\left(E_e + E_g\right) = 0\,. \tag{5.122}$$

After some algebra and taking into account that the electron number is related to the nucleon number by $N_e \simeq N_n/2$, we obtain

$$R = \frac{(9\pi)^{2/3}}{8} \frac{\hbar^2}{G m_e M_n^2} N_n^{-1/3}. \tag{5.123}$$

Note that the radius decreases as the total mass increases. For a solar mass star $M_\odot = 2\times 10^{30}$ kg and $M_n = 2\times 10^{-27}$ kg, this corresponds to a number of electrons $N_e \simeq 10^{57}$ concentrated in an Earth-size radius $R \simeq 8000$ km. The associated density is insane: a white dwarf is 2 million times denser than water! Note also that the numerical value of the Fermi energy, $E_F \simeq 2\times 10^5$eV, is comparable to the rest energy $E_{\rm rest} = m_e c^2 = 5.11\times 10^5$eV, so the most energetic electrons are getting fairly relativistic.
According to the above description more massive stars lead to smaller white dwarfs. This holds true until the mass of the progenitor star exceeds a critical value from which on inverse beta decays $e^- + p^+ \to n + \nu$ converts virtually all protons and electrons into neutrons, liberating neutrinos which carry off energy. The electron degeneracy

pressure is therefore no longer able to avoid gravitational collapse. The final outcome in this case will be a neutron star if the neutron degeneracy pressure is able to stop the collapse, or a black hole is that is not the case. The degeneracy pressure of the neutrons can be calculated in the same way as the electron pressure, but with N_e, n_e and m_e replaced by N_n, n_m and M_n, obtaining a radius

$$R = \frac{(9\pi)^{2/3}}{16^{1/3}} \frac{\hbar^2}{GM_n^3} N_n^{-1/3} \,. \tag{5.124}$$

For a two solar mass star, this corresponds to a small radius $R \simeq 10$ km. The associated density is astounding: all humanity can be squashed down to a sugar-sized piece of neutron star

[a] Although this is admittedly a lazy square-cow approximation, it will be enough for our purposes.

[b] Notice that this energy is inversely proportional to the mass, allowing us to neglect the energy of nucleons in the star in first approximation.

5.4.7 The occupation number representation

Even when formulated in terms of Slater determinants or permanents, the manipulation of symmetric or antisymmetric states become quite cumbersome for increasing numbers of particles. On the one hand, the computation of matrix elements between n-particle symmetrized/antisymmetrized wave functions involves generically a large number of integrals, $(n!)^2$. On the other hand, adding new particles to the system forces us to perform continuous re-symmetrizations/re-antisymmetrizations of the whole wave function, which is certainly not an optimized procedure for the treatment of grand-canonical ensembles, particle creation in relativistic quantum field theory, or the production of photons and phonons in quantum optic or solid state physics. The source of these complexities is related to the fact that we insisted on working which wave functions that contain a lot of useless information, such as the specific positions of the individual particles. Indeed, we know from the beginning that all possible permutations will appear anyway! Keeping only the minimal amount of necessary information is precisely what the so-called *occupation number formalism* does. In this setting, instead of asking which particle is on which state, we will be interested in knowing how many particles are there in each state. The essential step is the introduction of *creation* and *annihilation operators* with specific commutation relations able to accommodate the statistical properties of bosons and fermions. Because this description does not refer to the labeling of particles, it contains no redundant information, leading therefore to a precise and simpler description of the quantum many-body state.

Identical bosons

Consider a state $|n\rangle$ describing a system of n identical bosons in the same quantum state $|\omega\rangle$, which could be associated for instance with an energy level $E = \hbar\omega$. Introducing an *annihilation operator* a leading to a state with one particle less,

$$a|n\rangle = \sqrt{n}|n-1\rangle \,, \tag{5.125}$$

we can define a *number operator* $N = a^\dagger a$ counting the number of particles in $|n\rangle$,

$$\langle n|a^\dagger = \sqrt{n}\langle n-1| \quad \longrightarrow \quad \langle n|a^\dagger a|n\rangle = n \quad \longrightarrow \quad a^\dagger a|n\rangle = n|n\rangle\,. \tag{5.126}$$

This allows us to interpret the adjoint operator $a^\dagger$ as a *creation operator* leading to a state with one extra particle,

$$|n\rangle = \frac{1}{\sqrt{n+1}}a|n+1\rangle \quad \longrightarrow \quad a^\dagger|n\rangle = \frac{1}{\sqrt{n+1}}a^\dagger a|n+1\rangle\,, \tag{5.127}$$

namely

$$\boxed{a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle\,.} \tag{5.128}$$

A state with no particles is called the *vacuum state* and satisfies

$$\boxed{a|0\rangle = 0\,.} \tag{5.129}$$

The state with n particles is correspondingly given by

$$\boxed{|n\rangle = \frac{1}{\sqrt{n!}}(a^\dagger)^n|0\rangle\,,} \tag{5.130}$$

which, as expected for bosonic particles, is explicitly symmetric under particle exchange. The consistency of the procedure

$$aa^\dagger|n\rangle = \sqrt{n+1}a|n+1\rangle = (n+1)|n\rangle\,, \qquad a^\dagger a|n\rangle = \sqrt{n}a^\dagger|n-1\rangle = n|n\rangle\,, \tag{5.131}$$

leads to the commutation relations

$$\boxed{[a, a^\dagger] = 1\,.} \tag{5.132}$$

The above formalism can be easily extended to a system of n_1 bosons in the state $|\omega_1\rangle$, n_2 bosons in the state $|\omega_2\rangle$, etc... with the corresponding state $|n_1, n_2, \cdots\rangle$ understood as completely symmetric under the exchange of any pair of particles and the *occupation numbers* n_i satisfying the relation $n = \sum_i n_i$, with n the total number of bosons in the system.

As we did for the system with a single energy level, we can define the annihilation (a_i) and creation ($a_i^\dagger$) operators for particles in the state $|\omega_i\rangle$,

$$\boxed{a_i|n_1, \cdots, n_i, \cdots\rangle = \sqrt{n_i}|n_1, \cdots, n_i - 1, \cdots\rangle\,,} \tag{5.133}$$

$$\boxed{a_i^\dagger\,|n_1, \cdots, n_i, \cdots\rangle = \sqrt{n_i+1}\,|n_1, \cdots, n_i + 1, \cdots\rangle\,,} \tag{5.134}$$

with commutation relations

$$\boxed{[a_i, a_j^\dagger] = \delta_{ij}\,, \qquad [a_i, a_j] = [a_i^\dagger, a_j^\dagger] = 0\,.} \tag{5.135}$$

The associated number operator $N_i = a_i^\dagger a_i$ determines the occupation of $|\omega_i\rangle$,

$$a_i^\dagger a_i |n_1, \cdots, n_i, \cdots\rangle = n_i |n_1, \cdots, n_i, \cdots\rangle \,. \tag{5.136}$$

Then, any state of the system of n identical bosons can be written as

$$|n_1, n_2, \cdots\rangle = \frac{1}{\sqrt{n_1! n_2! \cdots}} (a_1^\dagger)^{n_1} (a_2^\dagger)^{n_2} \cdots |0\rangle \tag{5.137}$$

with $|0\rangle \equiv |0, 0, \cdots\rangle$ the vacuum state, which is again symmetric under the exchange of any pair of particles. For instance, if ω_1 is the lowest single-particle energy level, the ground-state of a non-interacting bosonic system is $|n, 0, 0, ...\rangle$ (all bosons in the lowest energy state).

The Hilbert space of the system of n identical bosons constructed out of the direct sum of Hilbert spaces of fixed particle number,[7]

$$\mathcal{H}_B = \mathcal{H}_S^0 \oplus \mathcal{H}_S^1 \oplus \mathcal{H}_S^2 \oplus \mathcal{H}_S^3 \oplus \cdots = \bigoplus_{n=0}^{\infty} \mathcal{H}_S^n \,, \tag{5.138}$$

is called the *Fock space of bosons*. In particular, this formalism describes a system of oscillators with a discrete spectrum of energies $E_i = \hbar\omega_i$. When there is a continuum of oscillator modes of energy E, we rather have $a_i = a(\omega_i) \to a(\omega)$. This defines a quantum field theory for bosonic fields. In this framework, the harmonic oscillator is no longer a simple mechanical oscillator, but a mode of excitation in a quantum field, which can be viewed as an infinite collection of harmonic oscillators.

Identical fermions

The *occupation number formalism* can also be adapted to describe a system of identical fermions whose states are antisymmetric under the exchange of particles. In order to satisfy the Pauli exclusion principle, we must require

$$|n_1, n_2, \ldots\rangle \,, \qquad \text{with} \qquad n_i = 0 \text{ or } 1 \,. \tag{5.139}$$

Introducing annihilation (η) and creation ($\eta^\dagger$) operators for fermionic states,

$$\eta|0\rangle = 0 \,, \qquad \eta|1\rangle = |0\rangle \,, \qquad \eta^\dagger|0\rangle = |1\rangle \,, \qquad \eta^\dagger|1\rangle = 0 \,, \tag{5.140}$$

$$\begin{aligned} \eta_i \, |n_1, \cdots n_i, \cdots\rangle &= (-1)^{\nu_i} n_i \, |n_1, \cdots, 1 - n_i, \cdots\rangle \\ \eta_i^\dagger \, |n_1, \cdots n_i, \cdots\rangle &= (-1)^{\nu_i} \, (1 - n_i) \, |n_1, \cdots, 1 - n_i, \cdots\rangle \,, \end{aligned} \tag{5.141}$$

with $\nu_i = \sum_{k=1}^{i-1} n_k$ the so-called Jordan–Wigner string depending on the ordering of the involved single-particle states, we can define the number operator $N_i = \eta_i^\dagger \eta_i$ giving the occupancy of state $|\omega_i\rangle$, namely

$$\eta_i^\dagger \eta_i \, |n_1, \cdots n_i, \cdots\rangle = n_i^2 \, |n_1, \cdots n_i, \cdots\rangle = n_i \, |n_1, \cdots n_i, \cdots\rangle \tag{5.142}$$

[7] Note how, by virtue of the direct sum, states containing a different number of particles are defined to be orthogonal.

and derive the anticommutation relations

$$\{\eta_i, \eta_j^\dagger\} = \delta_{ij}\,, \qquad \{\eta_i, \eta_j\} = \{\eta_i^\dagger, \eta_j^\dagger\} = 0\,, \tag{5.143}$$

with $\{A, B\} \equiv AB + BA$ the *anticommutator*. Any state of the system of identical fermions can be then written as

$$|n_1, n_2, \cdots\rangle = (\eta_1^\dagger)^{n_1} (\eta_2^\dagger)^{n_2} \cdots |0\rangle\,. \tag{5.144}$$

As expected, two fermions cannot occupy the same state,

$$\{\eta_i^\dagger, \eta_i^\dagger\} = 0\,, \qquad \longrightarrow \qquad (\eta_i^\dagger)^2 = 0\,, \tag{5.145}$$

ensuring the fulfillment of the Pauli exclusion principle. For instance, the ground-state for n non-interacting fermions is represented by $|1, 1, ..., 1, 0, 0, ...\rangle$.

When there is a continuum of modes of energy E, we have $\eta_i = \eta(\omega_i) \to \eta(\omega)$. This defines a quantum field theory for fermionic fields.

Superselection rules and Majorana fermions

Note that the seemingly innocent phase $(-1)^{\nu_i}$ in Eq. (5.141) seems to depend on the particle numbers n_k to the left of n_i and therefore on the specific order of fermions. This order should be, however, completely arbitrary, since our predictions of physical properties should not rely on how we decide to order the modes in our description. On top of that, the phase $(-1)^{\nu_i}$ seems to be something observable. Imagine for instance a *gedanken experiment* involving three modes, one on Earth and two on Jupiter. By measuring an operator containing an odd number of fermionic annihilation and creation operators, one could a priori determine by a phase whether a particle is present on Jupiter or not, achieving instantaneous communication and violating causality! The solution to these issues lies in the fact that not all observables are allowed for fermionic operators. In particular, any valid fermionic operator must commute with the parity operator

$$\Pi = i^n \prod_{j=1}^{2n} c_j\,, \tag{5.146}$$

written here in terms of $2n$ Majorana fermions, defined as

$$c_{2j-1} = f_j^\dagger + f_j\,, \qquad c_{2j} = -i(f_j^\dagger - f_j)\,, \tag{5.147}$$

for $j = 1, \ldots, n$. These fermions are Hermitian, $c_j = c_j^\dagger$, and satisfy the fermionic anticommutation relations $\{c_j, c_k\} = \delta_{j,k}$ for $j, k = 1, \ldots, 2n$. For obvious reasons, they are often referred to as *half fermions*.

5.5 Some worked-out exercises

Mixtures, superpositions, and basis transformations

Consider the following matrix expressed in an orthonormal basis $\{|\lambda_1\rangle, |\lambda_2\rangle\}$,

$$\rho = \frac{1}{4}\begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}.$$

a) Determine whether it can describe a physical state.
b) Determine whether it can describe a statistical mixture of the states $|\lambda_1\rangle$ and $|\lambda_2\rangle$ and, if so, calculate the weight or probability of obtaining each state in the mixture.
c) Determine whether it can represent a superposition of the states $|\lambda_1\rangle$ and $|\lambda_2\rangle$ and, if so, determine the coefficient of each state in the linear combination.
d) Determine whether it can represent a mixture of the states $|\kappa_1\rangle = \frac{\sqrt{3}}{2}|\lambda_1\rangle + \frac{1}{2}|\lambda_2\rangle$ and $|\kappa_2\rangle = \frac{\sqrt{3}}{2}|\lambda_1\rangle - \frac{1}{2}|\lambda_2\rangle$, and, if so, calculate the weight or probability of each state in the mixture.

a) To describe a physical state, the density matrix ρ must be Hermitian, semi-positive definite and have unit trace. Since the matrix elements are real, the eigenvalues $\lambda_1 = 3/4$ and $\lambda_2 = 1/4$ are non-zero and $\mathrm{Tr}\,\rho = \frac{1}{4}(3+1) = 1$, ρ describes a physical state..

b) The matrix ρ is diagonal in the $\{|\lambda_1\rangle, |\lambda_2\rangle\}$ basis and none of its eigenvalues are zero. It represents therefore a statistical mixture of these states with weights or probabilities $P(\lambda_1) = 3/4$ and $P(\lambda_2) = 1/4$.

c) Since ρ has two non-zero eigenvalues, it cannot represent a pure state. Therefore, it cannot represent a superposition of $|\lambda_1\rangle$ and $|\lambda_2\rangle$.

d) The states $|\kappa_1\rangle$ and $|\kappa_2\rangle$ form an orthonormal basis. Therefore, if ρ represents a mixture of these states, it should be diagonal in that basis, namely $\rho = p_1|\kappa_1\rangle\langle\kappa_1| + p_2|\kappa_2\rangle\langle\kappa_2|$, with $p_1 + p_2 = 1$, and p_1 and p_2 the probabilities of the respective pure states $|\kappa_1\rangle$ and $|\kappa_2\rangle$. Expanding $|\kappa_1\rangle\langle\kappa_1|$ and $|\kappa_2\rangle\langle\kappa_2|$,

$$|\kappa_1\rangle\langle\kappa_1| = \left(\frac{\sqrt{3}}{2}|\lambda_1\rangle + \frac{1}{2}|\lambda_2\rangle\right)\left(\frac{\sqrt{3}}{2}\langle\lambda_1| + \frac{1}{2}\langle\lambda_2|\right) = \frac{3}{4}|\lambda_1\rangle\langle\lambda_1| + \frac{\sqrt{3}}{4}|\lambda_1\rangle\langle\lambda_2| + \frac{\sqrt{3}}{4}|\lambda_2\rangle\langle\lambda_1| + \frac{1}{4}|\lambda_2\rangle\langle\lambda_2|,$$

$$|\kappa_2\rangle\langle\kappa_2| = \left(\frac{\sqrt{3}}{2}|\lambda_1\rangle - \frac{1}{2}|\lambda_2\rangle\right)\left(\frac{\sqrt{3}}{2}\langle\lambda_1| - \frac{1}{2}\langle\lambda_2|\right) = \frac{3}{4}|\lambda_1\rangle\langle\lambda_1| - \frac{\sqrt{3}}{4}|\lambda_1\rangle\langle\lambda_2| - \frac{\sqrt{3}}{4}|\lambda_2\rangle\langle\lambda_1| + \frac{1}{4}|\lambda_2\rangle\langle\lambda_2|,$$

and combining the terms, we get

$$\rho = \frac{3}{4}(p_1+p_2)|\lambda_1\rangle\langle\lambda_1| + \frac{\sqrt{3}}{4}(p_1-p_2)|\lambda_1\rangle\langle\lambda_2| + \frac{\sqrt{3}}{4}(p_1-p_2)|\lambda_2\rangle\langle\lambda_1| + \frac{1}{4}(p_1+p_2)|\lambda_2\rangle\langle\lambda_2|.$$

Since $p_1 + p_2 = 1$, this simplifies to

$$\rho = \frac{3}{4}|\lambda_1\rangle\langle\lambda_1| + \frac{\sqrt{3}}{4}(p_1-p_2)|\lambda_1\rangle\langle\lambda_2| + \frac{\sqrt{3}}{4}(p_1-p_2)|\lambda_2\rangle\langle\lambda_1| + \frac{1}{4}|\lambda_2\rangle\langle\lambda_2|.$$

Since the off-diagonal elements of ρ are zero, we must have $p_1 = p_2$, which, combined with the condition $p_1 + p_2 = 1$, implies $p_1 = p_2 = 1/2$. Therefore, ρ represents a mixture of $|\kappa_1\rangle$ and $|\kappa_2\rangle$, with statistical weights $1/2$.

Entangled states and density matrices

Consider a composite system of two spin 1/2 particles with Hilbert spaces $\mathcal{H}_1 = \mathbb{C}^2$ and $\mathcal{H}_2 = \mathbb{C}^2$. Show that the expectation value of any observable $A^{(1)} = A \otimes \mathbb{I}$ of particle 1 in a state $|\varphi\rangle = \frac{1}{\sqrt{2}}(|+\rangle \otimes |+\rangle - |-\rangle \otimes |-\rangle)$ can be expressed as $\langle A^{(1)}\rangle = \mathrm{Tr}[\rho^{(1)} A]$, with $\rho^{(1)} = \frac{1}{2}\mathbb{I}_{2\times 2}$. Compare this result with the expectation value $\langle A\rangle_{\mathrm{sup}} = \langle\phi|A|\phi\rangle$ of A in the superposition state $|\phi\rangle = \frac{1}{\sqrt{2}}(|+\rangle - |-\rangle)$. Find the observables for which both expectation values agree.

By direct computation, the expectation value $\langle A^{(1)}\rangle_{|\varphi\rangle}$ becomes

$$\begin{aligned}
\langle A^{(1)}\rangle_{|\varphi\rangle} &= \langle\varphi|A\otimes\mathbb{I}|\varphi\rangle = \frac{1}{\sqrt{2}}(\langle+|\otimes\langle+| - \langle-|\otimes\langle-|)(A\otimes\mathbb{I})\frac{1}{\sqrt{2}}(|+\rangle\otimes|+\rangle - |-\rangle\otimes|-\rangle) \\
&= \frac{1}{2}(\langle+|\otimes\langle+| - \langle-|\otimes\langle-|)(A|+\rangle\otimes|+\rangle - A|-\rangle\otimes|-\rangle) \\
&= \frac{1}{2}(\langle+|A|+\rangle\underbrace{\langle+|+\rangle}_{1} - \langle+|A|-\rangle\underbrace{\langle+|-\rangle}_{0} - \langle-|A|+\rangle\underbrace{\langle-|+\rangle}_{0} + \langle-|A|-\rangle\underbrace{\langle-|-\rangle}_{1}) \\
&= \frac{1}{2}(\langle+|A|+\rangle + \langle-|A|-\rangle) = \mathrm{Tr}[\rho^{(1)}A]\,,
\end{aligned}$$

where the last equality follows at this stage from pure inspection. Note, however, that $\rho^{(1)}$ is indeed the reduced density matrix for particle 1. To see this explicitly, let us write down explicitly the density matrix for the pure state $|\varphi\rangle$,

$$\rho = |\varphi\rangle\langle\varphi| = \frac{1}{2}\left[(|+\rangle\otimes|+\rangle - |-\rangle\otimes|-\rangle)(\langle+|\otimes\langle+| - \langle-|\otimes\langle-|)\right].$$

Expanding this expression yields four terms

$$\begin{aligned}
\rho &= \frac{1}{2}\left[(|+\rangle\otimes|+\rangle)(\langle+|\otimes\langle+|) - (|+\rangle\otimes|+\rangle)(\langle-|\otimes\langle-|) - (|-\rangle\otimes|-\rangle)(\langle+|\otimes\langle+|) + (|-\rangle\otimes|-\rangle)(\langle-|\otimes\langle-|)\right] \\
&= \frac{1}{2}\left[|+\rangle\langle+|\otimes|+\rangle\langle+| + |-\rangle\langle-|\otimes|-\rangle\langle-| - |+\rangle\langle-|\otimes|+\rangle\langle-| - |-\rangle\langle+|\otimes|-\rangle\langle+|\right].
\end{aligned}$$

The *reduced density matrix* $\rho^{(1)} = \mathrm{Tr}_2\,\rho$ for particle 1 is obtained by tracing out the degrees of freedom of particle 2 in this equation namely

$$\rho^{(1)} = \frac{1}{2}\left[|+\rangle\langle+|\,\mathrm{Tr}_2(|+\rangle\langle+|) + |-\rangle\langle-|\,\mathrm{Tr}_2(|-\rangle\langle-|) - |+\rangle\langle-|\,\mathrm{Tr}_2(|+\rangle\langle-|) - |-\rangle\langle+|\,\mathrm{Tr}_2(|-\rangle\langle+|)\right].$$

Taking into account that

$$\begin{aligned}
|+\rangle\langle+| &= \begin{pmatrix}1\\0\end{pmatrix}(1\quad 0) = \begin{pmatrix}1&0\\0&0\end{pmatrix} \quad\longrightarrow\quad \mathrm{Tr}_2(|+\rangle\langle+|) = 1\,, \\
|-\rangle\langle-| &= \begin{pmatrix}0\\1\end{pmatrix}(0\quad 1) = \begin{pmatrix}0&0\\0&1\end{pmatrix} \quad\longrightarrow\quad \mathrm{Tr}_2(|-\rangle\langle-|) = 1\,, \\
|+\rangle\langle-| &= \begin{pmatrix}1\\0\end{pmatrix}(0\quad 1) = \begin{pmatrix}0&1\\0&0\end{pmatrix} \quad\longrightarrow\quad \mathrm{Tr}_2(|+\rangle\langle-|) = 0\,, \\
|-\rangle\langle+| &= \begin{pmatrix}0\\1\end{pmatrix}(1\quad 0) = \begin{pmatrix}0&0\\1&0\end{pmatrix} \quad\longrightarrow\quad \mathrm{Tr}_2(|-\rangle\langle+|) = 0,
\end{aligned}$$

the previous expression reduces to the one given in the exercise

$$\rho^{(1)} = \frac{1}{2}\left[|+\rangle\langle+| + |-\rangle\langle-|\right] = \frac{1}{2}\mathbb{I}_{2\times 2}\,.$$

As for $\langle A\rangle_{|\phi\rangle}$, we have

$$\langle A\rangle_{|\phi\rangle} = \langle\phi|A|\phi\rangle = \frac{1}{\sqrt{2}}(\langle+| - \langle-|)A\frac{1}{\sqrt{2}}(|+\rangle - |-\rangle = \frac{1}{2}(\langle+|A|+\rangle - \langle+|A|-\rangle - \langle-|A|+\rangle + \langle-|A|-\rangle)\,.$$

For the expectation values $\langle A^{(1)}\rangle$ and $\langle A\rangle_{\text{sup}}$ to agree, the off-diagonal terms $\langle+|A|-\rangle$ and $\langle-|A|+\rangle$ must vanish. Hence, the observables for which both expectation values coincide are the ones that are diagonal in the $\{|+\rangle, |-\rangle\}$ basis.

Density matrix evolution for a three level system

Consider an ensemble of particles with a Hilbert space spanned by the orthonormal basis $\{|1\rangle, |0\rangle, |-1\rangle\}$. The ensemble is prepared in such a way that 30% of the particles are in the state $|1\rangle$, 50% in the state $|0\rangle$, and 20% in the state $\frac{1}{\sqrt{2}}(|1\rangle + |-1\rangle)$.

a) Write down the density matrix ρ for this ensemble in the $\{|1\rangle, |0\rangle, |-1\rangle\}$ basis.

b) Determine whether the initial state is pure or mixed.

c) Compute the expectation values of the operators

$$S_x = \frac{\hbar}{\sqrt{2}}\begin{pmatrix}0&1&0\\1&0&1\\0&1&0\end{pmatrix}, \quad S_y = \frac{\hbar}{\sqrt{2}}\begin{pmatrix}0&-i&0\\i&0&-i\\0&i&0\end{pmatrix}, \quad S_z = \hbar\begin{pmatrix}1&0&0\\0&0&0\\0&0&-1\end{pmatrix}$$

for this state.

d) Write down the time-dependent density matrix $\rho(t)$ using the unitary time evolution operator for a Hamiltonian $H = \hbar\omega(|1\rangle\langle 1| + |0\rangle\langle 0| + |-1\rangle\langle -1|)$.

e) Show that $\rho(t)$ remains a valid density matrix (i.e., it retains the properties of Hermiticity, unit trace, and positive semidefiniteness) at any time t.

a) The density matrix ρ is given by

$$\rho = 0.3|1\rangle\langle 1| + 0.5|0\rangle\langle 0| + 0.2\left(\frac{|1\rangle + |-1\rangle}{\sqrt{2}}\right)\left(\frac{\langle 1| + \langle -1|}{\sqrt{2}}\right),$$

or

$$\rho = 0.3\begin{pmatrix}1&0&0\\0&0&0\\0&0&0\end{pmatrix} + 0.5\begin{pmatrix}0&0&0\\0&1&0\\0&0&0\end{pmatrix} + 0.2\begin{pmatrix}\frac{1}{2}&0&\frac{1}{2}\\0&0&0\\\frac{1}{2}&0&\frac{1}{2}\end{pmatrix} = \begin{pmatrix}0.4&0&0.1\\0&0.5&0\\0.1&0&0.1\end{pmatrix}.$$

b) To diagonalize ρ, we solve for the eigenvalues,

$$\det(\rho - \lambda I) = \det\begin{pmatrix}0.4-\lambda&0&0.1\\0&0.5-\lambda&0\\0.1&0&0.1-\lambda\end{pmatrix} = 0\,,$$

obtaining

$$\lambda_1 = 0.5, \quad \lambda_2 = 0.3, \quad \lambda_3 = 0.2\,.$$

Since not all eigenvalues are 0 or 1, the state is a mixed state.

c) For S_x, S_y and S_z we have

$$\langle S_x\rangle = \mathrm{Tr}(\rho S_x) = \frac{\hbar}{\sqrt{2}}\,\mathrm{Tr}\begin{pmatrix} 0 & 0.4 & 0 \\ 0.4 & 0 & 0.4 \\ 0 & 0.4 & 0 \end{pmatrix} = 0\,,$$

$$\langle S_y\rangle = \mathrm{Tr}(\rho S_y) = \frac{\hbar}{\sqrt{2}}\,\mathrm{Tr}\begin{pmatrix} 0 & -0.4i & 0 \\ 0.4i & 0 & -0.4i \\ 0 & 0.4i & 0 \end{pmatrix} = 0\,,$$

$$\langle S_z\rangle = \mathrm{Tr}(\rho S_z) = \hbar\,\mathrm{Tr}\begin{pmatrix} 0.4 & 0 & 0.1 \\ 0 & 0 & 0 \\ 0.1 & 0 & -0.1 \end{pmatrix} = 0.4\hbar - 0.1\hbar = 0.3\hbar\,.$$

Degeneracies and total spin of four fermions in a one-dimensional box

Consider a composite system of four spin-1/2 particles confined in a one-dimensional infinite potential well, bounded by $x \in [0, L]$.

a) Determine the energies of the ground state and the first two excited states of the composite system.
b) Identify the degeneracies of the previous states and the possible values of the third component of the total spin in each case.
c) Write down the wavefunction representing the ground state of the composite system.

a) The energies and state vectors of the individual spin-1/2 particles take the form

$$E_n = n^2 E_1\,, \qquad E_1 = \frac{\pi^2\hbar^2}{2m_e L^2}\,, \qquad \langle x\,|n,\pm\rangle \equiv \sqrt{\frac{2}{L}}\sin\left(\frac{n\pi x}{L}\right)|\pm\rangle\,,$$

with $n \in \mathbb{N}$ the quantum number of the one-dimensional infinite potential well and the states $|\pm\rangle$ the spin states associated to the spin projections $m_s = \pm 1/2$. Since we are dealing with identical fermions, the Pauli exclusion principle prevents two spin-1/2 particles from having the same quantum numbers.

- Ground state: Two particles occupy the $n = 1$ level, with states $|1,+\rangle$ and $|1,-\rangle$. The other two particles occupy the $n = 2$ level, with states $|2,+\rangle$ and $|2,-\rangle$. The corresponding energy of the composite system is therefore

$$E_{\mathrm{G}} = 2(1^2 + 2^2)E_1 = 5\frac{\pi^2\hbar^2}{m_e L^2}\,.$$

- First excited state: Two particles occupy the $n = 1$ level, with states $|1,+\rangle$ and $|1,-\rangle$. The third particle occupies the $n = 2$ level with state $|2,\pm\rangle$. The fourth particle occupies the $n = 3$ level with state $|3\pm\rangle$. The corresponding energy of the composite system is then

$$E_{1\mathrm{st}} = (2\cdot 1^2 + 2^2 + 3^2)E_1 = \frac{15}{2}\frac{\pi^2\hbar^2}{m_e L^2}\,.$$

- Second excited state: One particle occupies the $n = 1$ level, with state $|1, \pm\rangle$, two particles occupy the $n = 2$ level, with states $|2, +\rangle$ and $|2, -\rangle$, and the fourth particle occupies the $n = 3$ level, with state $|3, \pm\rangle$. The corresponding energy of the composite system is then

$$E_{2\text{nd}} = (1 + 2 \cdot 2^2 + 3^2)E_1 = 9\frac{\pi^2\hbar^2}{m_e L^2} \,.$$

b) The ground state is non-degenerate. The sum of the individual m_s values determining the third projection of the total spin is 0.

The first excited state is 4 times degenerate, corresponding to the possible combinations of m_s values of the particles in the $n = 2$ and $n = 3$ levels. The possible values of the z component of the total spin are $\{-1, 0, 1\}$.

The second excited state is 4 times degenerate, corresponding to the possible combinations of m_s values of the particles in the $n = 1$ and $n = 3$ levels. The possible values of the third component of the total spin are $\{-1, 0, 1\}$.

c) In the ground state, the four particles occupy the individual states $|1, +\rangle$, $|1, -\rangle$, $|2, +\rangle$, and $|2, -\rangle$. Since we are dealing with identical fermions, the state vector of the composite system must be completely antisymmetric. Using the Slater determinant, we have

$$|\varphi\rangle = \frac{1}{\sqrt{4!}} \begin{vmatrix} |1,+\rangle_1 & |1,-\rangle_1 & |2,+\rangle_1 & |2,-\rangle_1 \\ |1,+\rangle_2 & |1,-\rangle_2 & |2,+\rangle_2 & |2,-\rangle_2 \\ |1,+\rangle_3 & |1,-\rangle_3 & |2,+\rangle_3 & |2,-\rangle_3 \\ |1,+\rangle_4 & |1,-\rangle_4 & |2,+\rangle_4 & |2,-\rangle_4 \end{vmatrix} .$$

Two interacting particles in a 1D harmonic trap: bosons vs fermions

Consider a system of two particles of equal mass m constrained to move in a one dimensional harmonic oscillator potential and interacting weakly through an attractive harmonic force proportional to the distance separating them, namely $F_{12} = -\kappa\,(x_1 - x_2)$. Determine the energies and wave functions of the ground state, the first excited state and the second excited state under the assumption that the trapped particles are a) identical bosons of spin 0 and b) identical spin 1/2 particles.

Denoting $k = m\omega^2$ for compactness, the Hamiltonian of this two-particle system can be written as

$$H = -\frac{\hbar^2}{2M}\left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}\right) + \frac{1}{2}k\left(x_1^2 + x_2^2\right) + \frac{\kappa}{2}\left(x_1 - x_2\right)^2 ,$$

with $\kappa > 0$. This expression can be diagonalized by introducing the set of variables

$$\xi = \frac{1}{\sqrt{2}}\left(x_1 + x_2\right), \qquad \eta = \frac{1}{\sqrt{2}}\left(x_1 - x_2\right) ,$$

which, up to normalization constants, can be understood either as a rotation of 45 degrees in the $x_1 - x_2$ plane or a change of coordinates to the reference frame defined by the center of mass ("CM") of the system and the relative coordinate between the particles ("rel"). Taking into account that the gradient part of the Hamiltonian is invariant under rotations, we get

$$H = -\frac{\hbar^2}{2M}\left(\frac{\partial^2}{\partial \xi^2} + \frac{\partial^2}{\partial \eta^2}\right) + \frac{1}{2}k\left(\xi^2 + \eta^2\right) + \kappa\eta^2 = -\frac{\hbar^2}{2M}\left(\frac{\partial^2}{\partial \xi^2} + \frac{\partial^2}{\partial \eta^2}\right) + \frac{1}{2}k\xi^2 + \frac{1}{2}(k + 2\kappa)\eta^2 \,.$$

Written in this form, the Hamiltonian is just the sum of two independent harmonic oscillators of angular frequencies

$$\omega_{\rm CM} = \sqrt{\frac{k}{M}}\,, \qquad \omega_{\rm rel} = \sqrt{\frac{k+2\kappa}{M}}\,.$$

The total energies and spatial wave functions of the system are therefore given by

$$E_{n_{\rm CM}n_{\rm rel}} = \left(n_{\rm CM}+\frac{1}{2}\right)\hbar\omega_{\rm CM} + \left(n_{\rm rel}+\frac{1}{2}\right)\hbar\omega_{\rm rel}\,, \qquad \varphi_{n_{\rm CM}n_{\rm rel}}(\xi,\eta) = \varphi^{(k)}_{n_{\rm CM}}(\xi)\varphi^{(k+2\kappa)}_{n_{\rm rel}}(\eta),$$

with $n_{\rm CM}, n_{\rm rel} = 0, 1, 2, \cdots$ non-negative integers representing the quantum numbers for the CM and relative motions and $\varphi^{(k)}_n$ the n-th eigenfunction of a harmonic oscillator of spring constant k.
Since the CM coordinate ξ depends symmetrically on both x_1 and x_2, and therefore it remains unchanged ($\xi \leftrightarrow \xi$) under particle exchange ($x_1 \leftrightarrow x_2$), the wave function $\varphi_{n_{\rm CM}}(\xi)$, which depends only on ξ, is also symmetric under particle exchange, $\varphi_{n_{\rm CM}}(\xi) \longrightarrow \varphi_{n_{\rm CM}}(\xi)$, for any quantum number $n_{\rm CM}$. On the contrary, the relative coordinate η changes sign ($\eta \leftrightarrow -\eta$) under particle interchange ($x_1 \leftrightarrow x_2$), meaning the wave function $\varphi_{n_{\rm rel}}(\eta)$ will have the following behavior:

- If $n_{\rm rel}$ is even, the relative wave function $\varphi_{n_{\rm rel}}(\eta)$ is symmetric with respect to η, $\varphi_{n_{\rm rel}}(\eta) = \varphi_{n_{\rm rel}}(-\eta)$, and therefore symmetric under particle exchange.
- If $n_{\rm rel}$ is odd, the relative wave function $\varphi_{n_{\rm rel}}(\eta)$ is antisymmetric with respect to η, $\varphi_{n_{\rm rel}}(\eta) = -\varphi_{n_{\rm rel}}(-\eta)$, and therefore antisymmetric under particle exchange.

The total spatial wave function $\varphi_{n_{\rm CM}n_{\rm rel}}(\xi,\eta)$ is therefore symmetric if $n_{\rm rel}$ is even and antisymmetric if $n_{\rm rel}$ is odd.
Taking into account that $\omega_{\rm CM} < \omega_{\rm rel}$, the three lowest energy levels of the system are, a priori,

$$\begin{array}{lll} n_{\rm CM}=0\,, \quad n_{\rm rel}=0\,, & n_{\rm CM}=1\,, \quad n_{\rm rel}=0\,, & n_{\rm CM}=0\,, \quad n_{\rm rel}=1\,, \\ E_{00} = \dfrac{1}{2}\hbar\left(\omega_{\rm CM}+\omega_{\rm rel}\right), & E_{10} = \dfrac{1}{2}\hbar\left(\omega_{\rm CM}+\omega_{\rm rel}\right) + \hbar\omega_{\rm CM}, & E_{01} = \dfrac{1}{2}\hbar\left(\omega_{\rm CM}+\omega_{\rm rel}\right) + \hbar\omega_{\rm rel}. \end{array}$$

For identical and spinless bosons, the associated *total* wave functions, containing in this specific case a purely spatial part only, must be symmetric with respect to the interchange of the two particles. Thus, only the symmetric states $\varphi_{00}(\xi,\eta) = \varphi^{(k)}_0(\xi)\varphi^{(k+2\kappa)}_0(\eta)$ (ground state) and $\varphi_{10}(\xi,\eta) = \varphi^{(k)}_1(\xi)\varphi^{(k+2\kappa)}_0(\eta)$ (1st excited state) are allowed. The state $\varphi_{01}(\xi,\eta) = \varphi^{(k)}_0(\xi)\varphi^{(k+2\kappa)}_1(\eta)$ is, on this occasion, forbidden, due to the antisymmetric contribution of the relative wave function $\varphi^{(k+2\kappa)}_1(\eta)$.

If the particles are identical with spin 1/2, the associated *total* wave functions, including now the product of spatial and spin parts, must be antisymmetric with respect to an interchange of the two particles. As the spatial parts $\varphi_{00}(\xi,\eta) = \varphi^{(k)}_0(\xi)\varphi^{(k+2\kappa)}_0(\eta)$ and $\varphi_{10}(\xi,\eta) = \varphi^{(k)}_1(\xi)\varphi^{(k+2\kappa)}_0(\eta)$ are symmetric, they must be combined with an antisymmetric spin contribution (the singlet state). Up to normalization factors, we have

$$\varphi_{00}(\xi,\eta)\left(|+-\rangle - |-+\rangle\right), \qquad \varphi_{10}(\xi,\eta)\left(|+-\rangle - |-+\rangle\right),$$

On the other hand, since the spatial part $\varphi_{01}(\xi,\eta)$ is antisymmetric due to the relative-motion contribution $\varphi^{(k+2\kappa)}_1(\eta)$, it must be combined with a antisymmetric spin contribution (the triplet state). Up to normalization factors, we have

$$\varphi_{01}(\xi,\eta)|\pm\pm\rangle\,, \qquad \varphi_{01}(\xi,\eta)\left(|+-\rangle + |-+\rangle\right).$$

5.6 Exercises

1. **How to test a bomb without exploding it?**

   Consider a stock of bombs equipped with a novel type of sensor, such that if a single photon strikes the sensor, it triggers an explosion. Some of these bombs are, however, malfunctioning, allowing photons to pass through the sensor's hole without any impact. As a result, these bombs do not explode. Is it possible to identify the bombs that are still operational? *Hint:* See arXiv:hep-th/9305002

2. **Composite systems of two spins 1/2 particles**

   a) Consider a system of two spin 1/2 particles with Hilbert spaces $\mathcal{H}_1 = \mathbb{C}^2$ and $\mathcal{H}_2 = \mathbb{C}^2$ and Hamiltonian

   $$H = \epsilon \left(\sigma_x \otimes \sigma_x + \sigma_y \otimes \sigma_y + \sigma_z \otimes \sigma_z\right) ,$$

   with ϵ a constant with dimensions of energy. Determine the associated stationary states in the basis of eigenstates of $\sigma_z \otimes \sigma_z$.

   b) Repeat the exercise for a Hamiltonian

   $$H = \epsilon \left(\sigma_z \otimes \mathbb{I} + \mathbb{I} \otimes \sigma_z\right) + J \left(\sigma_x \otimes \sigma_x + \sigma_y \otimes \sigma_y\right) ,$$

   with J a constant with dimensions of energy.

3. **A density matrix for an ensemble of spin 1/2 particles**

   The density matrix

   $$\rho = \begin{pmatrix} \frac{1}{4} & n \\ n^* & p \end{pmatrix}$$

   describes an ensemble of spin- $\frac{1}{2}$ particles in the S_z basis, with the asterisk representing complex conjugation.

   a) Which are the allowed values of p?

   b) What values must n have for ρ to represent a pure state?

   c) Determine the pure state associated to the maximum possible real value of n.

4. **Reconstructing the density matrix from spin expectation values**

   a) A pure ensemble of spin-1/2 particles is prepared identically. The expectation values $\langle S_x \rangle$ and $\langle S_z \rangle$ are known, along with the *sign* of $\langle S_y \rangle$. Show how this information allows reconstruction of the state vector $|\varphi\rangle$, up to an overall phase. Why is it unnecessary to know the full magnitude of $\langle S_y \rangle$ in this case?

   b) Now consider a *mixed* ensemble of spin-$\frac{1}{2}$ particles. Suppose the ensemble averages $\langle S_x \rangle$, $\langle S_y \rangle$, and $\langle S_z \rangle$ are all known. Show how to construct the 2×2 density matrix ρ that fully characterizes the ensemble.

5. **Density matrix evolution for a two level system**

   A two-level quantum system is initially in the state described by the density matrix

   $$\rho(0) = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

and evolves under a Hamiltonian

$$H = \frac{\hbar\omega}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

a) Write down the time-dependent density matrix $\rho(t)$ using the unitary time-evolution operator.

b) Show that $\rho(t)$ remains a valid density matrix (i.e., it retains the properties of Hermiticity, unit trace, and positive semidefiniteness) at any time t.

c) Calculate the expectation value of the Pauli matrix σ_z as a function of time, namely $\langle\sigma_z(t)\rangle$.

6. **Characterizing a spin-1 density matrix**

How many independent real parameters are required to fully specify the density matrix of an ensemble of three level system?

7. **Permutation symmetry and observables**

a) Verify, formally and graphically, that the set of permutations of n elements form a group, namely that i) all compositions are again a permutation, ii) there is an identity element leaving everything as it is, iii) there is an inverse for every permutation and iv) the product of permutations is associative. How many permutations of n letters are there?

b) Show that all observables of identical particles commute with the permutation operator, especially the Hamiltonian.

8. **The hydrogen atom is a boson**

The symmetrization postulate for elementary particles induces also a definite bosonic or fermionic character for composite particles, as long as the energy exchanges with their environment are much less than the difference between the first excited state and the ground state. By considering a system of two hydrogen atoms, show that the hydrogen atom as a whole is indeed a boson.

Generalize the result to a system of N elementary fermions, arguing that the composite particle is a fermion when it contains an odd number of fermions, and a boson when it contains an even number of fermions. Provide examples of composite particles that are bosons and fermions.

9. **Three non-interacting particles in a 1D infinite potential well**

Consider a system of three non-interacting particles constrained to move in a one dimensional infinite potential well of length L, i.e. $V(x) = 0$ for $0 < x < L$ and $V(x) = \infty$ otherwise. Determine the energies and wave functions of the ground state, the first excited state and the second excited state under the assumption that the trapped particles are a) spinless and distinguishable particles with masses $m_1 < m_2 < m_3$, b) identical bosons of spin 0, c) identical spin 1/2 particles and, iv) all in the spin state $|+\rangle$. Determine the associated degeneracies.

10. ***N* non-interacting particles in a 1D infinite potential well**

    Consider N noninteracting bosons confined in an infinite potential well of length L, i.e. $V(x) = 0$ for $0 < x < L$ and $V(x) = \infty$ otherwise. Determine the ground state energy of the system. How would the ground state energy change if the particles were fermions instead of bosons?

11. **Two non-interacting particles in a sphere**

    Consider a system of two identical and non-interacting spin 1/2 particles constrained to move inside a sphere of radius R. Determine the wave functions of the ground state and the first excited state.

12. **Five non-interacting particles in a 1D harmonic oscillator**

    Consider a system of five non-interacting particles constrained to move in a one dimensional harmonic oscillator potential. *Describe* the ground state of the system under the assumption that the trapped particles are a) identical fermions, b) identical bosons and c) distinguishable particles. Determine the associated degeneracy. Repeat the exercise for excited configurations involving one unit of energy above the ground state, two units of energy above the ground state and three units of energy above the ground state.

13. **Six non-interacting particles in a ring**

    Certain molecular systems can be modelled as a number of electrons confined to move on a circular ring of radius R. Assuming the electrons in the ring not to interact with each other, derive the normalized one-particle wave functions and energies for this scenario. Use this result to estimate the three lowest energy levels of benzene, an aromatic hydrocarbon composed of six carbon atoms joined in a planar hexagonal ring with one hydrogen atom attached to each and 6 electrons. Taking into account that the average C-C bond length in benzene is 1.40 Angstroms, estimate the minimum energy required to excite an electron from the ground state to a higher energy level. What is the corresponding wavelength at which this molecular system can absorb radiation to facilitate this excitation?

14. **Zero-point energy and scaling properties in a lattice**

    A particle of mass m moves through a lattice composed of delta function barriers, each with large dimensionless strength $g_0 \gg 1$. The lattice has spacing L and total length D, with $D \gg L$.

    a) Since $g_0 \gg 1$, the barriers are very strong, meaning transmission across them is highly suppressed. Particles are effectively localized within each individual well. Estimate the ground state energy by modeling the particle as confined to a single infinite square well of size L. Express your result.

    b) If instead $g_0 \ll 1$, the barriers would be weak, and the particle could freely explore the entire system of length D. Calculate the ground state energy in this limit and compare it to the result from part a). How does the ratio of the two energies scale with D/L?

    c) Based on your results, discuss why the essential physical properties of a solid (such as optical response or conductivity) remain unchanged when considering samples of increasing size, provided $D \gg L$.

CHAPTER 6

ANGULAR MOMENTUM

For a spin-one particle only—because it requires three amplitudes—the correspondence with a vector is very close. In each case, there are three numbers that must transform with coordinate changes in a certain definite way. In fact, there is a set of base states which transform just like the three components of a vector.

RICHARD FEYNMAN, CHAPTER 5, THE FEYNMAN LECTURES ON PHYSICS, VOLUME III, 1965

In this chapter, we will explore the fundamental role of symmetry in quantum mechanics through the study of the rotation group and the elegant algebraic theory of angular momentum, which facilitates the solution of problems with rotational symmetries like the hydrogen atom. The chapter also introduces to so-called Clebsch-Gordan coefficients, a crucial tool for combining angular momenta in composite systems.

6.1 Groups and algebras

A set G forms a *group* if it is endowed with an operation $\star$ and satisfies:

1. Closure: If $g_1, g_2 \in G$, then $g_1 \star g_2 \in G$.
2. Associativity: $g_1 \star (g_2 \star g_3) = (g_1 \star g_2) \star g_3$ for all $g_1, g_2, g_3 \in G$.
3. There exists a unit element I such that $I \star g = g \star I = g$ for all $g \in G$.
4. Every $g \in G$ has an inverse g^{-1} such that $g^{-1} \star g = g \star g^{-1} = I$.

Given a group G, a *subgroup* H is a subset of G that contains I such that $h_1 \star h_2 \in H$ for all $h_1, h_2 \in H$. A group is called *Abelian* (or commutative) if $g_1 \star g_2 = g_2 \star g_1$ for all $g_1, g_2 \in G$ and *non-Abelian* (or non-commutative) if there is at least one pair g_1, g_2 such that $g_1 \star g_2 \neq g_2 \star g_1$.

A parametric group in which the dependence of the group elements on all its parameters is continuous and differentiable is called a *Lie group*. The number of independent parameters is called the dimension D of G. The *Lie algebra* $\mathcal{G}$ associated with a Lie group G is obtained

by expanding the group element g around the identity element I, namely[1]

$$g(x_i) = I + i\sum_{i=1}^{D} x_i T^i + \mathcal{O}(x^2)\,, \tag{6.1}$$

with x_i $(i = 1, \ldots, D)$ a set of local parameters and T^i the *generators* of the Lie algebra $\mathcal{G}$. These satisfy

$$[T^i, T^j] = i\sum_{k=1}^{D} f^{ijk} T^k\,, \tag{6.2}$$

with f^{ijk} the so-called *structure constants*, and the *Jacobi identity*[2]

$$[[T^i, T^j], T^k] + [[T^j, T^k], T^i] + [[T^k T^i], T^j] = 0\,. \tag{6.3}$$

A Lie algebra satisfying Eqs. (6.2) and (6.3) can be integrated to obtain a Lie group (Lie's theorem). Up to subtleties, the study of Lie groups can be reduced to the study of Lie algebras.

6.2 The rotation group

In three dimensions, rotations can be described using various mathematical representations, such as axes and angles or the Euler-angle parametrization. For instance, in an orthonormal basis $\{x, y, z\}$, rotations of angles α, β and γ around the x, y and z axes can be written as

$$R_x(\alpha) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}\,, \qquad R_y(\beta) = \begin{pmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{pmatrix}\,, \tag{6.4}$$

$$R_z(\gamma) = \begin{pmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix}\,, \tag{6.5}$$

A general rotation can be then constructed out of three consecutive rotations around the z, y and new z axes,

$$R(\alpha, \beta, \gamma) = R_z(\alpha) R_y(\beta) R_z(\gamma)\,, \tag{6.6}$$

namely

$$R(\alpha, \beta, \gamma) = \begin{pmatrix} \cos\alpha\cos\beta\cos\gamma - \sin\alpha\sin\gamma & -\cos\alpha\cos\beta\sin\gamma - \sin\alpha\cos\gamma & \cos\alpha\sin\beta \\ \sin\alpha\cos\beta\cos\gamma + \cos\alpha\sin\gamma & -\sin\alpha\cos\beta\sin\gamma + \cos\alpha\cos\gamma & \sin\alpha\sin\beta \\ -\sin\beta\cos\gamma & \sin\beta\sin\gamma & \cos\beta \end{pmatrix}\,.$$

The Euler-angle and the axis-angle parametrizations are related by

$$\tan\theta = \frac{\tan\frac{\beta}{2}}{\sin\frac{\alpha+\gamma}{2}}, \qquad \varphi = \frac{\pi + \alpha - \gamma}{2}, \qquad \cos\phi = 2\cos^2\frac{\beta}{2}\cos^2\frac{\alpha+\gamma}{2} - 1\,. \tag{6.7}$$

[1] The factor i ensures the generators T^i to be Hermitian. In mathematics, this factor is usually dropped, taking $T_i = -T_i^\dagger$.

[2] Since matrix multiplication is associative, this last relation is automatically satisfied by matrices.

The set of rotations forms a non-Abelian *Lie group* with the composition of rotations as binary operation. In three-dimensions this group is denoted as $SO(3)$, which represents the special orthogonal group in three dimensions. For an infinitesimal rotation $R_z(\phi)$ around the z axis, we have

$$R_z(\delta\phi) = \begin{pmatrix} 1 & -\delta\phi & 0 \\ \delta\phi & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \mathbb{I} - i\ \delta\phi\, K_z\,, \tag{6.8}$$

with the *self-adjoint* matrix

$$K_z = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{6.9}$$

the generator of these rotations. Taking into account that $R_z(\phi + \delta\phi) = R_z(\phi)R_z(\delta\phi)$, we can then write

$$R_z(\phi) + \frac{\mathrm{d}R_z(\phi)}{\mathrm{d}\phi}\delta\phi = R_z(\phi)\,(\mathbb{I} - i\,\delta\phi K_z) \quad \longrightarrow \quad \frac{dR_z(\phi)}{d\phi} = -i\,R_z(\phi)K_z\,, \tag{6.10}$$

or equivalently

$$R_z(\phi) = \mathrm{e}^{-i\,K_z\phi}\,, \tag{6.11}$$

where we have taken into account the boundary condition $R_z(0) = \mathbb{I}$ when performing the integration. Similarly rotations around the x-axis and the y-axis are generated by

$$R_x(\phi) = \mathrm{e}^{-i\,K_x\phi}\,, \qquad R_y(\phi) = \mathrm{e}^{-i\,K_y\phi}\,, \tag{6.12}$$

with

$$K_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \qquad K_y = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}. \tag{6.13}$$

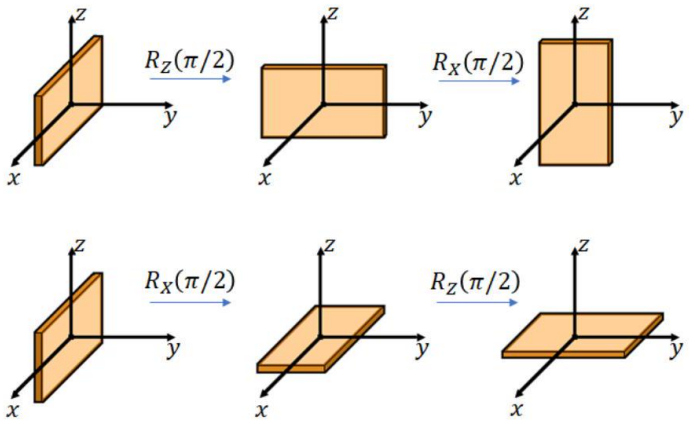
The three generators K_x, K_y and K_z satisfy the commutation relations

$$[K_i, K_j] = i\,\epsilon_{ijk}K_k\,, \tag{6.14}$$

and therefore the same algebra as the angular momentum

$$[J_i, J_j] = i\,\epsilon_{ijk}\hbar J_k\,, \qquad \vec{J} \equiv \hbar\vec{K}\,. \tag{6.15}$$

This is course a well-known property: while rotations around the same axis commute, rotations around different axes do not. For instance, a rotation of $\pi/2$ around the z axis followed by a rotation of $\pi/2$ around the x axis produces a different result from that obtained by a rotation of $\pi/2$ around the x axis followed by a rotation of $\pi/2$ around the z axis:



6.3 The angular momentum operator

Just as in classical mechanics, the angular momentum in quantum mechanics appears as the generator of rotations. Under a rotation of angle ϕ around an arbitrary axis $\vec{n} = (n_x, n_y, n_z)$ $\vec{n}$, the vectors $|\varphi\rangle \in \mathcal{H}$ change as[3]

$$|\varphi\rangle \mapsto R_n(\phi)|\varphi\rangle = \mathrm{e}^{-\frac{i}{\hbar}\phi\vec{n}\cdot\vec{J}}|\varphi\rangle\,, \tag{6.18}$$

with $\vec{J}$ the total angular momentum operator for the system. Note that this definition of the angular momentum as the generator of rotations superseeds definitions such as $\vec{L} = \vec{x} \times \vec{p}$ which are not as general, and, in particular, not applicable to spin systems. The angular momentum $\vec{J}$ is defined without any reference to the position or momentum of anything, it could be an orbital momentum L, a spin S, a combination of both or something else!

Example 31: Pauli matrices as an angular momentum representation

As we show in the first chapter, the Pauli matrices,

$$J_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad J_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad J_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{6.19}$$

satisfy the same commutation relations as the angular momentum (6.15), allowing us to interpret them as the generators of the rotations in the 3-dimensional space of a spin 1/2 system. They are indeed a 2-dimensional representation of the angular momentum algebra, in contrast to the 3-dimensional representation in (6.9) and (6.12).

[3] In general an infinitesimal rotation around $\vec{n}$,

$$R_{\vec{n}}(\delta\phi) = \mathbb{I} - i\,\delta\phi\,(n_x K_x + n_y K_y + n_z K_z) = \mathbb{I} - i\,\delta\phi\vec{n}\cdot\vec{K}\,, \tag{6.16}$$

is generated by

$$R_{\vec{n}}(\phi) = \mathrm{e}^{-i\,\phi\vec{n}\cdot\vec{K}}\,. \tag{6.17}$$

In our usual basis $|\pm\rangle$, and taking into account that $\sigma_i^2 = \mathbb{I}$, we have for instance,

$$R_z(\phi) = \mathrm{e}^{-\frac{i}{\hbar}J_z\phi} = \begin{pmatrix} \mathrm{e}^{-i\,\frac{\phi}{2}} & 0 \\ 0 & \mathrm{e}^{i\,\frac{\phi}{2}} \end{pmatrix}, \tag{6.20}$$

$$\begin{aligned} R_y(\phi) = \mathrm{e}^{-\frac{i}{\hbar}J_y\phi} &= \mathbb{I} - i\,\frac{\phi}{2}\sigma_y + \frac{1}{2!}\left(-i\,\frac{\phi}{2}\right)^2\sigma_y^2 + \frac{1}{3!}\left(-i\,\frac{\phi}{2}\right)^3\sigma_y^3\ldots \\ &= \mathbb{I}\left[1 - \frac{1}{2!}\left(\frac{\phi}{2}\right)^2 + \ldots\right] - i\,\sigma_y\left[\frac{\phi}{2} - \frac{1}{3!}\left(\frac{\phi}{2}\right)^3 + \ldots\right] \\ &= \mathbb{I}\cos\frac{\phi}{2} - i\,\sigma_y\sin\frac{\phi}{2} = \begin{pmatrix} \cos\frac{\phi}{2} & -\sin\frac{\phi}{2} \\ \sin\frac{\phi}{2} & \cos\frac{\phi}{2} \end{pmatrix}, \end{aligned} \tag{6.21}$$

or more generically

$$R(\vec{\phi}) = \exp\left[-\frac{i}{2}(\vec{\phi}\cdot\vec{\sigma})\right] = \sum_{n=0}^{\infty}\frac{1}{n!}\left[-\frac{i}{2}\,\vec{\phi}\cdot\vec{\sigma}\right] = \cos(\phi/2) - i\sin(\phi/2)(\vec{\phi}\cdot\vec{\sigma})\,. \tag{6.22}$$

Note, however, that rotations in spin 1/2 systems are somehow unusual since a rotation of 2π radians does not take us back to the original state, but rather produces a $e^{i\pi}$ phase,

$$R_z(2\pi)|\varphi\rangle = R_y(2\pi)|\varphi\rangle = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix} = -|\varphi\rangle\,. \tag{6.23}$$

In other words, the $s = 1/2$ representation does not satisfy the property $R_n(0) = R_n(2\pi) = \mathbb{I}$, being necessary to perform a 4π rotation for getting back to the same state! This unusual rotation group is called the $SU(2)$ group, the Special ($\det R = 1$) Unitary ($R^\dagger R = \mathbb{I}$) group with complex Hermitian generators

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \begin{pmatrix} o & u^* \\ u & 0 \end{pmatrix}. \tag{6.24}$$

Due to above behaviour, the SU(2) group is said to be double covering group of SO(3), to which it is locally isomorphic (same algebra) but globally different (smaller number of elements).

The fact that a 2π rotation is not the identity operation prompts us to consider its physical significance, more specifically whether it leads to any observable effects. Although an overall phase factor in a quantum state lacks any physical meaning, it's possible to split a beam of spin-1/2 particles into two segments and subject one of them to a 2π rotation. As a result, the phase shift becomes discernible in the interference pattern that emerges when the segments are rejoined. This experimental scenario has been conducted with neutron beams, showing the -1 phase shift associated with 2π rotations to be real [cf. Sakurai's book for a more detailed discussion].

The Dirac's belt trick

The so-called Dirac's belt trick or the Balinese cup trick demonstrate that when an object with attached strings is rotated by 360 degrees, it does not return to its initial state. A subsequent rotation of another 360 degrees does, however, the job.

This can be easily verified by taking a cup of coffee and executing two hand rotations while keeping the cup in the upright position. Following the initial hand rotation, your arm becomes twisted. However, during the second rotation, your arm returns to its original orientation. This is achieved by the hand making one rotation over the elbow, resulting in the arm's twist, followed by another rotation under the elbow, which effectively untwists it.

Alternatively, you could make use of a belt featuring a conventional frame buckle, with the prong acting as a reference point. While the end opposite of the buckle is secured to prevent any movement, the belt is extended without any initial twist, and the buckle is meticulously maintained in a horizontal position while undergoing a complete clockwise rotation of 360 degrees. At this point, the belt will visibly exhibit a twist. A second clockwise rotation of 360 degrees, while seemingly exacerbating the twist, enables the belt to be restored to its untwisted state while maintaining the buckle in a horizontal position and aligned in the same direction.

While the different component of $\vec{J}$ do not commute with each other,

$$[J_i, J_j] = i\hbar\epsilon_{ijk}J_k\,, \tag{6.25}$$

they do so with the Casimir operator

$$J^2 = J_x^2 + J_y^2 + J_z^2, \tag{6.26}$$

since this is a scalar and scalars are not changed by rotations. In particular, taking into account the Leibniz identity $[AB, C] = A[B, C] + [A, C]B$ in (3.44), we have for instance

$$\begin{aligned}\left[J^2, J_z\right] &= \left[J_x^2 + J_y^2 + J_z^2, J_z\right] = \left[J_x^2, J_z\right] + \left[J_y^2, J_z\right] = \\ &= J_x\left[J_x, J_z\right] + \left[J_x, J_z\right]J_x + J_y\left[J_y, J_z\right] + \left[J_y, J_z\right]J_y = \\ &= J_x\left(-i\hbar J_y\right) + \left(-i\hbar J_y\right)J_x + J_y\left(i\hbar J_x\right) + \left(i\hbar J_x\right)J_y = 0\,,\end{aligned} \tag{6.27}$$

or more generally

$$\left[J^2, J_i\right] = 0, \tag{6.28}$$

meaning that it is always possible to measure simultaneously the absolute value of the angular momentum and one of its components. Choosing for instance the commuting operators J^2 and J_z, we can write

$$J^2|j,m\rangle = \hbar^2 j(j+1)|j,m\rangle, \qquad J_z|j,m\rangle = m\hbar|j,m\rangle\,, \tag{6.29}$$

with $j(j+1)$ and m a convenient notation for the corresponding eigenvalues, with $j, m \in \mathbb{R}$ as must be for Hermitian operators.[4] In order to determine the possible values of these quantities, let us combine the operators J_x and J_y into the so-called raising and lowering operators

$$J_\pm = J_x \pm i J_y \tag{6.30}$$

satisfying the commutation relations

$$\left[J^2, J_\pm\right] = 0\,, \qquad [J_z, J_\pm] = \pm\hbar J_\pm\,, \qquad [J_+, J_-] = 2\hbar J_z\,. \tag{6.31}$$

The first expression follows directly from the definition of $J_\pm$ and the property (6.28). For the second and third expressions we have respectively

$$[J_z, J_\pm] = [J_z, J_x \pm i J_y] = [J_z, J_x] \pm i\,[J_z, J_y] = i\hbar J_y \pm i\,(-i\hbar J_x) = \pm\hbar\,(J_x \pm i J_y) = \pm\hbar J_\pm \tag{6.32}$$

$$\begin{aligned}[J_+, J_-] &= [(J_x + iJ_y), (J_x - iJ_y)] = [J_x, J_x] + [J_x, -iJ_y] + [iJ_y, J_x] + [iJ_y, -iJ_y] \\ &= 0 - i\hbar J_z + i\hbar J_z - (-1)\hbar^2 J_x = 2\hbar J_z + \hbar^2 J_x - \hbar^2 J_x = 2\hbar J_z\,.\end{aligned} \tag{6.33}$$

Other useful relations

Other useful relations among J^2, J_z and $J_\pm$ are

$$J^2 = \frac{1}{2}\,(J_+ J_- + J_- J_+) + J_z^2\,, \tag{6.34}$$

and

$$J_+ J_- = J^2 - J_z^2 + \hbar J_z\,, \qquad J_- J_+ = J^2 - J_z^2 - \hbar J_z\,. \tag{6.35}$$

The corresponding proof is also straightforward. Expanding the products $J_+ J_-$ and $J_- J_+$ as

$$\begin{aligned}J_+ J_- &= (J_x + iJ_y)(J_x - iJ_y) = J_x^2 + J_y^2 - i[J_x, J_y]\,, \\ J_- J_+ &= (J_x - iJ_y)(J_x + iJ_y) = J_x^2 + J_y^2 + i[J_x, J_y]\,,\end{aligned}$$

we get

$$J_+ J_- + J_- J_+ = 2(J_x^2 + J_y^2)\,,$$

and therefore (6.34),

$$J^2 = J_x^2 + J_y^2 + J_z^2 = \frac{1}{2}(J_+ J_- + J_- J_+) + J_z^2\,.$$

From this expression and using the commutator $[J_+, J_-] = 2\hbar J_z$, we find (6.35).

[4] Note also that the positivity of the J^2 requires $j(j+1) \geq 0$.

The name given to $J_\pm$ becomes obvious when considering the action of J^2 and J_z on the states $J_\pm|j,m\rangle$, namely

$$J^2 J_\pm|j,m\rangle = J_\pm J^2|j,m\rangle = \hbar^2 j(j+1)J_\pm|j,m\rangle \tag{6.36}$$

$$J_z J_\pm|j,m\rangle = (J_\pm J_z + [J_z,J_\pm])\,|j,m\rangle = (J_\pm J_z \pm \hbar J_\pm)\,|j,m\rangle = (m\pm 1)\hbar J_\pm|j,m\rangle\,, \tag{6.37}$$

where we have used the first two expressions in (6.31). The states $J_\pm|j,m\rangle$ are therefore eigenstates of the operators J^2 and J_z with eigenvalues j and $m+1$. This allows us to write

$$J_+|j,m\rangle = c^+_{jm}|j,m+1\rangle\,, \qquad J_-|j,m\rangle = c^-_{jm}|j,m-1\rangle\,, \tag{6.38}$$

with the constants c^+_{jm} and c^-_{jm} related through the condition $J_- = (J_+)^+$, namely

$$c^-_{jm} = \langle j,m-1|J_-|j,m\rangle = \langle j,m|J_+|j,m-1\rangle^* = (c^+_{jm-1})^*\,. \tag{6.39}$$

Given a state $|j,m\rangle$, the operators J_+ and J_- generate a whole tower of states

$$\ldots,\ |j,m-2\rangle,\ |j,m-1\rangle,\ |j,m\rangle,\ |j,m+1\rangle,\ |j,m+2\rangle\ \ldots \tag{6.40}$$

In particular, the application of the raising operator J_+ provides vectors with higher J_z,

$$|j,m\rangle \xrightarrow{J_+} |j,m+1\rangle \xrightarrow{J_+} |j,m+2\rangle \ldots \xrightarrow{J_+} |j,m_{\max}\rangle\,, \tag{6.41}$$

up to a limiting case $J_+|j,m_{\max}\rangle = 0$ (physical systems do not have infinite J_z). Analogously, the lowering operator J_- generates vectors with lower J_z,

$$|j,m\rangle \xrightarrow{J_-} |j,m-1\rangle \xrightarrow{J_-} |j,m-2\rangle \ldots \xrightarrow{J_-} |j,m_{\min}\rangle \tag{6.42}$$

with $J_-|j,m_{\min}\rangle = 0$. The proportionality constants in Eqs. (6.38) and (6.39) can be easily obtained by expressing the product J_-J_+ in

$$|c^+_{jm}|^2 = \langle j,m|(J_+)^\dagger J_+|j,m\rangle = \langle j,m|J_-J_+|j,m\rangle\,, \tag{6.43}$$

in terms of J^2 and J_z,

$$J_-J_+ = (J_x - iJ_y)(J_x+iJ_y) = J_x^2 + J_y^2 + i[J_x,J_y] = J^2 - J_z^2 - \hbar J_z\,, \tag{6.44}$$

such that

$$|c^+_{jm}|^2 = \langle j,m|J^2 - J_z^2 - \hbar J_Z|j,m\rangle = \hbar^2[j(j+1) - m(m+1)]\,, \tag{6.45}$$

where we have used that $\langle j,m|j,m\rangle = 1$. Combined with the condition (6.39) this expression determines c^+_{jm}, and c^-_{jm} up to a phase factor, which is generally chosen to make them real and positive, thereby defining the arbitrary phase in $|j,m\rangle$. With this choice, Eqs. (6.38) and (6.39) become

$$J_+|j,m\rangle = \hbar\sqrt{j(j+1)-m(m+1)}|j,m+1\rangle\,, \tag{6.46}$$

$$J_-|j,m\rangle = \hbar\sqrt{j(j+1)-m(m-1)}|j,m-1\rangle\,. \tag{6.47}$$

The matrices representing the different angular momentum components J_i for a given j can be easily obtained from Eq. (6.29) and these expressions.

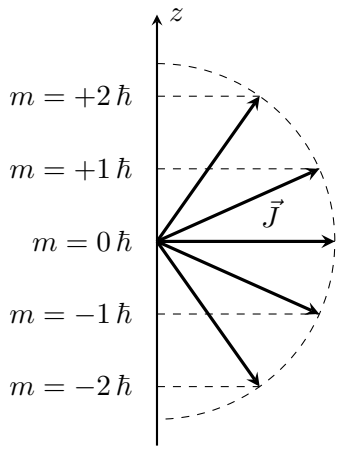
Note that since $J^2 - J_z^2 = J_x^2 + J_y^2$ must have positive definite eigenvalues, one must have

$$j(j+1) - m^2 \geq 0\,, \quad \longrightarrow \quad |m| \leq \sqrt{j(j+1)} \quad \longrightarrow \quad |m| \leq j+1 \tag{6.48}$$

The existence of a maximum and a minimum cutoff in (6.46) and (6.47) restricts the difference the difference $m_{\max} - m_{\min} = 2j$ to be an integer and with it the allowed quantum numbers

$$j = 0,\, 1/2,\, 1,\, 3/2,\, \ldots \qquad m = j,\, j-1\,, \ldots,\, -j+1, -j\,. \tag{6.49}$$

Angular momentum can be only integral or half-integral!



$l = 2 \qquad |\vec{J}| = \sqrt{2(2+1)}\hbar$

The quantization of angular momentum is a direct consequence of commutation relations, which in turn derive from the properties of rotations and the definition of J as the generator of rotations. For a given j, there exists therefore $2j+1$ basis states, namely

$$\begin{aligned}
&j = 0, && |0,0\rangle\,,\\
&j = 1/2, && |1/2,1/2\rangle, |1/2,-1/2\rangle\,,\\
&j = 1, && |1,1\rangle, |1,0\rangle, |1,-1\rangle\,,\\
&j = 3/2, && |3/2,3/2\rangle, |3/2,1/2\rangle, |3/2,-1/2\rangle, |3/2,-3/2\rangle\,,\\
&\text{etc.}\ldots
\end{aligned}$$

In this basis, the x and y components of the angular momentum,

$$J_x = \frac{1}{2}\left(J_+ + J_-\right), \qquad J_y = \frac{1}{2i}\left(J_+ - J_-\right)\,, \tag{6.50}$$

follow straightforwardly from Eqs. (6.46) and (6.47),

$$\begin{aligned}
\langle j, m-1\,|J_x|\,j,m\rangle &= \langle j,m\,|J_x|\,j,m-1\rangle = \frac{\hbar}{2}\sqrt{j(j+1)-m(m-1)}\,,\\
\langle j, m-1\,|J_y|\,j,m\rangle &= -\langle j,m\,|J_y|\,j,m-1\rangle = \frac{i\hbar}{2}\sqrt{j(j+1)-m(m-1)}\,.
\end{aligned} \tag{6.51}$$

All other matrix elements are equal to zero, and in particular the diagonal elements determining the average value of these operators in the corresponding state, namely

$$\langle J_x\rangle = \langle J_y\rangle = 0\,. \tag{6.52}$$

Example 32: Spin 3/2

Acting on the set $\varphi = (|3/2, 3/2\rangle, |3/2, 1/2\rangle, |3/2, -1/2\rangle, |3/2, -3/2\rangle)$ for $j = 3/2$, we get the 4×4 matrices

$$J_z = \hbar \begin{pmatrix} \frac{3}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & -\frac{3}{2} \end{pmatrix}, \tag{6.53}$$

$$J_+ = \hbar \begin{pmatrix} 0 & \sqrt{3} & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & \sqrt{3} \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad J_- = \hbar \begin{pmatrix} 0 & 0 & 0 & 0 \\ \sqrt{3} & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & \sqrt{3} & 0 \end{pmatrix}, \tag{6.54}$$

and from them the *irreducible representations*[a]

$$J_x = \frac{1}{2}(J_+ + J_-) = \frac{\hbar}{2} \begin{pmatrix} 0 & \sqrt{3} & 0 & 0 \\ \sqrt{3} & 0 & 2 & 0 \\ 0 & 2 & 0 & \sqrt{3} \\ 0 & 0 & \sqrt{3} & 0 \end{pmatrix},$$

$$J_y = -\frac{i}{2}(J_+ - J_-) = i\frac{\hbar}{2} \begin{pmatrix} 0 & -\sqrt{3} & 0 & 0 \\ \sqrt{3} & 0 & -2 & 0 \\ 0 & 2 & 0 & -\sqrt{3} \\ 0 & 0 & \sqrt{3} & 0 \end{pmatrix}, \tag{6.55}$$

$$J_z = \hbar \begin{pmatrix} \frac{3}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & -\frac{3}{2} \end{pmatrix}.$$

[a] Irreducible means that there is no basis where these matrices can be simultaneously reduced to smaller diagonal blocks.

Rotation symmetry and energy degeneracies

If a given Hamiltonian H is invariant under rotations, $\left[J^2, H\right] = [J_i, H] = 0$, $[J_\pm, H] = 0$,

$$H(J_\pm|njm\rangle) = J_\pm H|njm\rangle = E_n(J_\pm|njm\rangle) \longrightarrow H|njm\pm1\rangle = E_n|njm\pm1\rangle\,, \tag{6.56}$$

and the $(2j+1)$ states with $m = -j, \ldots, j$ the same energy,

$$H|njm\rangle = E_n|njm\rangle\,, \tag{6.57}$$

with n an index labelling it. This happens for instance in the hydrogen atom, where energies only depend on the principal quantum number n,

$$E = -\frac{1}{2n^2}\frac{m\left(Ze^2\right)^2}{\hbar^2}\,. \tag{6.58}$$

6.4 Orbital angular mometum

As this point you should be confused... Using only commutation relations coinciding exactly with those of the orbital angular momentum, we have just found that the quantum numbers j and m can take half-integer values. However, the treatment of the angular momentum in *Quantum Physics I* yielded only integer values for l and m. What are we missing? To answer this question, let us consider the form of the various angular momentum operators in a spherical coordinate system

$$x = r\sin\theta\cos\phi\,, \qquad y = r\sin\theta\sin\phi\,, \qquad z = r\cos\theta\,, \tag{6.59}$$

with $0 \leq \theta \leq \pi$ and $0 \leq \phi \leq 2\pi$. From Eq. (6.8) with $J = L = \hbar K$ we have

$$\langle r,\theta,\phi|\left(\mathbb{I} - \frac{i}{\hbar}\delta\phi L_z\right)|\alpha\rangle = \langle r,\theta,\phi|\left(\mathbb{I} + \frac{i}{\hbar}\delta\phi L_z\right)^+|\alpha\rangle = \langle r,\theta,\phi-\delta\phi|\alpha\rangle =$$
$$= \langle r,\theta,\phi|\alpha\rangle - \delta\phi\frac{\partial}{\partial\phi}\langle r,\theta,\phi|\alpha\rangle\,,$$

or equivalently

$$\boxed{L_z = -i\hbar\frac{\partial}{\partial\phi}\,.} \tag{6.60}$$

The corresponding expressions for the operators L_x, L_y and $L^2 = L_x^2 + L_y^2 + L_z^2$ are more easily obtained by transforming their Cartesian expressions to spherical coordinates. To this end, we make use of the chain rule

$$\frac{\partial}{\partial x_i} = \sum_j \frac{\partial\alpha_j}{\partial x_i}\frac{\partial}{\partial\alpha_j} = \sum_j \left(\frac{\partial x_i}{\partial\alpha_j}\right)^{-1}\frac{\partial}{\partial\alpha_j}\,, \tag{6.61}$$

with $(\partial x_i/\partial\alpha_j)^{-1}$ the inverse of the matrix $\partial x_i/\partial\alpha_j$ relating the total differential of a Cartesian coordinate dx_i to the spherical coordinates $\alpha_j = r,\theta,\phi$, namely

$$dx_i = \sum_j \frac{\partial x_i}{\partial\alpha_j}\,d\alpha_j\,, \tag{6.62}$$

with

$$\frac{\partial x_i}{\partial\alpha_j} = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial\theta} & \frac{\partial x}{\partial\phi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial\theta} & \frac{\partial y}{\partial\phi} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial\theta} & \frac{\partial z}{\partial\phi} \end{pmatrix} = \begin{pmatrix} \sin\theta\cos\phi & r\cos\theta\cos\phi & -r\sin\theta\sin\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi & r\sin\theta\cos\phi \\ \cos\theta & -r\sin\theta & 0 \end{pmatrix}. \tag{6.63}$$

Inverting this matrix, we get [5]

$$\begin{pmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial r} & \frac{\partial}{\partial\theta} & \frac{\partial}{\partial\phi} \end{pmatrix} \begin{pmatrix} \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \\ \frac{1}{r}\cos\theta\cos\phi & \frac{1}{r}\cos\theta\sin\phi & -\frac{1}{r}\sin\theta \\ -\frac{1}{r}\frac{\sin\phi}{\sin\theta} & \frac{1}{r}\frac{\cos\phi}{\sin\theta} & 0 \end{pmatrix}, \tag{6.64}$$

[5] Note that since the elements dx_i form a column vector, the components $\partial/\partial x_i$ form a row vector or covector.

or equivalently

$$\begin{aligned}\frac{\partial}{\partial x} &= \sin\theta\cos\phi\frac{\partial}{\partial r} + \frac{1}{r}\cos\theta\cos\phi\frac{\partial}{\partial\theta} - \frac{1}{r}\frac{\sin\phi}{\sin\theta}\frac{\partial}{\partial\phi}\,,\\ \frac{\partial}{\partial y} &= \sin\theta\sin\phi\frac{\partial}{\partial r} + \frac{1}{r}\cos\theta\sin\phi\frac{\partial}{\partial\theta} + \frac{1}{r}\frac{\cos\phi}{\sin\theta}\frac{\partial}{\partial\phi}\,,\\ \frac{\partial}{\partial z} &= \cos\theta\frac{\partial}{\partial r} - \frac{1}{r}\sin\theta\frac{\partial}{\partial\theta}\,.\end{aligned} \tag{6.65}$$

Using these expressions, we have

$$L_x = yp_z - zp_y = -i\hbar\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right) = i\hbar\left(\sin\phi\frac{\partial}{\partial\theta} + \frac{\cos\phi}{\tan\theta}\frac{\partial}{\partial\phi}\right)\,, \tag{6.66}$$

$$L_y = zp_x - xp_z = -i\hbar\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right) = i\hbar\left(-\cos\phi\frac{\partial}{\partial\theta} + \frac{\sin\phi}{\tan\theta}\frac{\partial}{\partial\phi}\right)\,, \tag{6.67}$$

$$L^2 = -\hbar^2\left[\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)\right]\,, \tag{6.68}$$

with the last expression coinciding, up to a multiplicative factor, with the Laplace operator ∇^2. Combining Eq. (6.60) and (6.68) with the result of multiplying the equivalent of (6.29) by the bra $\langle\theta,\phi|$,

$$\langle\theta,\phi|L^2|l,m\rangle = \hbar^2 l(l+1)\langle\theta,\phi|l,m\rangle\,, \qquad \langle\theta,\phi|L_z|l,m\rangle = \hbar m\langle\theta,\phi|l,m\rangle\,, \tag{6.69}$$

we arrive to the eigenvalue equations

$$L^2 Y_{l,m}(\theta,\phi) = -\hbar^2\left[\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)\right]Y_{l,m}(\theta,\phi) = \hbar^2 l(l+1)Y_{l,m}(\theta,\phi)\,, \tag{6.70}$$

$$L_z Y_{l,m}(\theta,\phi) = -i\hbar\frac{\partial}{\partial\phi}Y_{l,m}(\theta,\phi) = \hbar m Y_{l,m}(\theta,\phi)\,, \tag{6.71}$$

with

$$Y_{l,m}(\theta,\phi) = \langle\theta,\phi|l,m\rangle \tag{6.72}$$

the so-called spherical harmonics forming a basis in the corresponding two-dimensional angular space,

$$\sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell} Y_{l,m}(\theta,\phi)Y^*_{l,m}\left(\theta',\phi'\right) = \delta\left(\cos\theta - \cos\theta'\right)\delta\left(\phi-\phi'\right)\,, \qquad \text{(Closure)} \tag{6.73}$$

$$\int_0^{2\pi} d\phi\int_0^{\pi} Y^*_{l',m'}(\theta,\phi)Y_{l,m}(\theta,\phi)\sin\theta d\theta = \delta_{\ell'\ell}\delta_{m'm}\,, \qquad \text{(Orthonormality)}. \tag{6.74}$$

The solution of the above system of equations becomes easier upon realizing that Eq. (6.71) allows for a further separation of variables, $Y_{l,m}(\theta,\phi) = \Theta_{l,m}(\theta)\Phi_m(\phi)$,

$$-i\frac{\partial}{\partial\phi}\Phi_m(\phi) = m\Phi_m(\phi)\,, \tag{6.75}$$

admitting therefore a solution

$$\Phi_m(\phi) \propto e^{im\phi} \,. \tag{6.76}$$

This expression makes explicit that the lack of half-integer values for l in the hydrogen atom is due to additional restrictions beyond the commutation relations. In particular, in order to have the same value at ϕ and $\phi + 2\pi$, we must require $\exp(2\pi i m) = 1$, forcing m, and hence l, to be an integer.

To determine the remaining functions $\Theta_{l,m}(\theta)$, we made use of the *raising* and *lowering* operators in spherical coordinates,

$$L_+ = L_x + iL_y = \hbar\, e^{i\phi} \left(\frac{\partial}{\partial\theta} + \frac{i}{\tan\theta} \frac{\partial}{\partial\phi} \right) , \tag{6.77}$$

$$L_- = L_x - iL_y = -\hbar\, e^{-i\phi} \left(\frac{\partial}{\partial\theta} - \frac{i}{\tan\theta} \frac{\partial}{\partial\phi} \right) . \tag{6.78}$$

In particular, acting on the state with the highest projection $m = l$, $L_+ Y_{l,l} = 0$, we get

$$\left(\frac{\partial\, \Theta_{l,l}}{\partial\theta} - \frac{l\, \Theta_{l,l}}{\tan\theta} \right) e^{il\phi} = 0 \quad \longrightarrow \quad \frac{\partial}{\partial\theta} \ln \Theta_{l,l} = \frac{\partial}{\partial\theta} \ln(\sin\theta)^l \quad \longrightarrow \quad \Theta_{l,l} = c_l (\sin\theta)^l \,, \tag{6.79}$$

or equivalently

$$Y_{l,m}(\theta,\phi) = c_l (\sin\theta)^l e^{il\phi} \,, \tag{6.80}$$

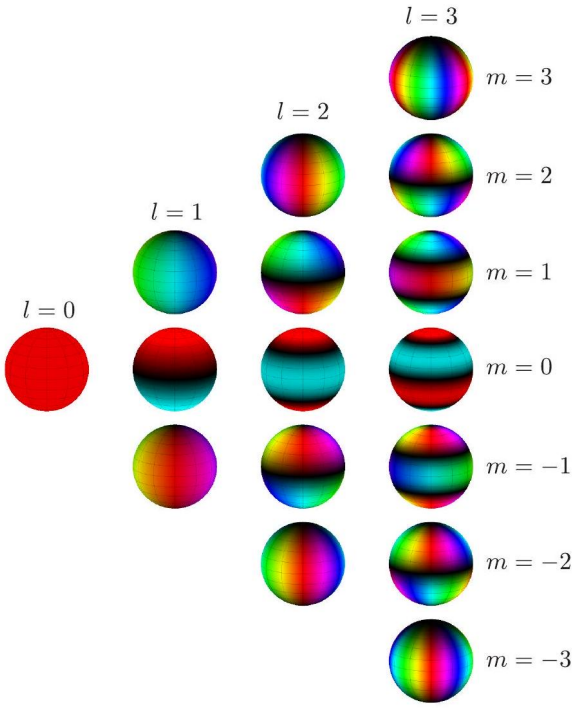
with c_l some integration constants to be fixed, up to a phase, by the normalization condition (6.74). One obtains

$$c_l = \frac{1}{\sqrt{2\pi}} \frac{(-1)^l}{2^l l!} \sqrt{\frac{(2l+1)!}{2}} \,. \tag{6.81}$$

The remaining states with $m = l-1, \ldots, -l+1, -l$ are obtained by applying (6.47), namely

$$Y_{l,m-1}(\theta,\phi) = \frac{L_- Y_{l,m}(\theta,\phi)}{\hbar\sqrt{l(l+1) - m(m-1)}} \,. \tag{6.82}$$

A list of the first few spherical harmonics can be found in Appendix C. Their graphical representation as color-phase density plots on spherical surfaces is displayed in the following figure. The z-axis points upward, the x-axis extends out of the page, and the y-axis points to the right. Black regions represent nodes or function values of zero.



Example 33: A few spherical harmonics

Let us determine the explicit form of the spherical harmonics for the cases $l = 0$ and $l = 1$.
For $l = m = 0$, we have $Y_{0,0}(\theta, \phi) = c_0$, with c_0 fixed, up to a phase, by the normalization condition (6.74),

$$\int_0^{2\pi}\int_0^{\pi} |c_0|^2 \sin\theta d\theta d\phi = 1 \quad \longrightarrow \quad |c_0|^2 = \frac{1}{4\pi} \quad \longrightarrow \quad Y_{0,0}(\theta, \phi) = \frac{1}{\sqrt{4\pi}}\,.$$

For $l = m = 1$, we rather have $Y_{1,1}(\theta, \phi) = c_1 \sin\theta e^{i\phi}$, with c_1 fixed, up to a phase, by the normalization condition (6.74),

$$\int_0^{2\pi}\int_0^{\pi} |c_1|^2 \sin^3\theta d\theta d\phi = 1 \quad \longrightarrow \quad |c_1|^2 = \frac{3}{8\pi} \quad \longrightarrow \quad Y_{1,1}(\theta, \phi) = -\sqrt{\frac{3}{8\pi}} \sin\theta e^{i\phi}\,.$$

Acting with L_- on this normalized state and taking into account Eqs. (6.78) and (6.82), we obtain

$$Y_{1,0}(\theta, \phi) = \frac{1}{\hbar\sqrt{2}} L_- Y_{1,1}(\theta, \phi) = \sqrt{\frac{3}{16\pi}}\, e^{-i\phi} \left(\cos\theta\, e^{i\phi} + \cos\theta e^{i\phi}\right) = \sqrt{\frac{3}{4\pi}} \cos\theta\,,$$

and

$$Y_{1,-1}(\theta, \phi) = \frac{1}{\hbar\sqrt{2}} L_- Y_{1,0}(\theta, \phi) = \sqrt{\frac{3}{8\pi}}\, e^{-i\phi} \sin\theta\,.$$

6.5 Addition of angular momenta

In systems involving several angular momenta interacting with each other, the conservation law of angular momentum applies only to the total angular momentum and not to the separate components, making the former a more relevant quantity.

Example 34: Two spinless particles interacting through a central potential

Consider for instance the Hamiltonian of two spinless particles interacting with each other through a function $V[(\vec{r}_1 - \vec{r}_2)^2]$ of the distance between them,

$$|\vec{r}_1 - \vec{r}_2| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}\,.$$

While the kinetic contributions commute with both angular momenta, this is not the case of the potential counterpart. For instance, for an arbitrary ket $|\varphi\rangle$, we have

$$\begin{aligned}
\langle\vec{r}|[L_{1z}, H]|\varphi\rangle &= \langle\vec{r}|x_1[p_{1y}, V] - y_1[p_{1x}, V]|\varphi\rangle \\
&= -i\hbar\left[x_1\left(\frac{\partial\,(V\varphi)}{\partial y_1} - V\frac{\partial\varphi}{\partial y_1}\right) - y_1\left(\frac{\partial\,(V\varphi)}{\partial x_1} - V\frac{\partial\varphi}{\partial x_1}\right)\right] \\
&= -i\hbar\left(x_1\frac{\partial V}{\partial y_1} - y_1\frac{\partial V}{\partial x_1}\right)\varphi \\
&= -\frac{i\hbar}{r}\frac{\partial V}{\partial r}[x_1\,(y_1 - y_2) - y_1\,(x_1 - x_2)]\varphi \\
&\neq 0\,,
\end{aligned}$$

and

$$\langle\vec{r}|[L_{2z}, V]|\varphi\rangle = -i\hbar\left(x_2\frac{\partial V}{\partial y_2} - y_2\frac{\partial V}{\partial x_2}\right)\varphi = +\frac{i\hbar}{r}\frac{\partial V}{\partial r}[x_1\,(y_1 - y_2) - y_1\,(x_1 - x_2)]\varphi \neq 0\,.$$

Similar results hold for all other components,

$$[L_{1,j}, V] \neq 0\,, \qquad [L_{2,j}, V] \neq 0\,,$$

meaning that the angular momenta of the individual particles are not good quantum numbers in this case. They are not conserved quantities. Note, on the other hand, that the total orbital angular momentum is conserved

$$\langle\vec{r}|[\vec{L}_1 + \vec{L}_2, V]|\varphi\rangle = 0\,.$$

How do we construct eigenstates of the total angular momentum of a composite system with two angular momenta $\vec{J}_1$ and $\vec{J}_2$? As explained in Section 5.1, the Hilbert space of the resulting system is given by the tensor product of the Hilbert spaces of the individual systems, namely $\mathcal{H} = \mathcal{H}^{j_1} \otimes \mathcal{H}^{j_2}$. In particular, since $\vec{J}_1$ and $\vec{J}_2$ act on different spaces, $[\vec{J}_1, \vec{J}_2] = 0$,

they cannot be naively added, being necessary to consider a tensor product

$$\vec{J} = \vec{J}_1 \otimes \mathbb{I} + \mathbb{I} \otimes \vec{J}_2 \equiv \vec{J}_1 + \vec{J}_2 \tag{6.83}$$

such that, in the direct product basis

$$\{|j_1, m_1; j_2, m_2\rangle \equiv |j_1, m_1\rangle \otimes |j_2, m_2\rangle\} \ , \tag{6.84}$$

the operators $\vec{J}_1$ and $\vec{J}_2$ act respectively on the vectors of the Hilbert spaces $\mathcal{H}^{j_1}$ and $\mathcal{H}^{j_2}$, namely

$$\vec{J}\,|j_1, m_1; j_2, m_2\rangle = (\vec{J}_1\,|j_1, m_1\rangle)\,|j_2, m_2\rangle + |j_1, m_1\rangle\,(\vec{J}_2\,|j_2, m_2\rangle)\,. \tag{6.85}$$

Uniqueness

Note that the way of summing angular momenta in Eq. (6.83) is unique. Indeed, it is a simple exercise to check that if we had introduce some non-zero complex constants α and β,

$$\vec{J} = a\,\vec{J}_1 \otimes \mathbb{I} + b\,\mathbb{I} \otimes \vec{J}_2\,,$$

we would have

$$[J_i, J_j] = i\hbar\epsilon_{ijk}\left(a^2\,J_{1k} \otimes \mathbb{I} + b^2\,\mathbb{I} \otimes J_{2k}\right) \ ,$$

forcing us to impose the conditions $a^2 = a$ and $b^2 = b$ in order to recover the algebra of an angular momentum if the operator in parenthesis is J_k. Since neither a nor b is zero (otherwise we would not summing anything), the only solution is $a = b = 1$.

Every irreducible representation of $\vec{J}$ is labeled by an eigenvalue of J^2,

$$\begin{aligned} J^2 &= J_1^2 \otimes \mathbb{I} + \mathbb{I} \otimes J_2^2 + (\vec{J}_1 \otimes \mathbb{I}) \cdot (\mathbb{I} \otimes \vec{J}_2) + (\mathbb{I} \otimes \vec{J}_2) \cdot (\vec{J}_1 \otimes \mathbb{I}) \\ &= J_1^2 \otimes \mathbb{I} + \mathbb{I} \otimes J_2^2 + 2 \sum_{i=x,y,z} J_{1i} \otimes J_{2i} \equiv J_1^2 + J_2^2 + 2\vec{J}_1 \cdot \vec{J}_2\,, \end{aligned} \tag{6.86}$$

with the corresponding states tagged by those of J_z,

$$J_z = J_{1z} \otimes \mathbb{I} + \mathbb{I} \otimes J_{2z} \equiv J_{1z} + J_{2z}\,. \tag{6.87}$$

Since the operators J_1^2, J_2^2, J^2, J_z commute with each other[6], we can make use of them to define a new (*coupled*) basis of eigenstates $\{|j_1, j_2, j, m\rangle\}$, which we will compactly denote as $|j, m\rangle_C$ in order to avoid extensive repetition of the labels j_1 and j_2. The relation to the original *uncoupled* basis $\{|j_1, m_1; j_2, m_2\rangle\}$ constructed out of $J_1^2, J_2^2, J_{1z}, J_{2z}$ is given by

$$|j, m\rangle_C = \sum_{m_1=-j_1}^{j_1} \sum_{m_2=-j_2}^{j_2} |j_1, m_1; j_2, m_2\rangle\,\langle j_1, m_1; j_2, m_2|j, m\rangle_C\,. \tag{6.88}$$

[6]Note, however, that $\left[J^2, J_{1z}\right] \neq 0$ and $\left[J^2, J_{2z}\right] \neq 0$

where we have made use of the completeness relation[7]

$$\sum_{m_1,m_2} |j_1, j_2; m_1, m_2\rangle\langle j_1, j_2; m_1, m_2| = \mathbb{I} \tag{6.89}$$

to introduce the so-called *Clebsch-Gordan coefficients*

$$\langle j_1, m_1; j_2, m_2|j, m\rangle_C \ , \tag{6.90}$$

which are only different from zero if

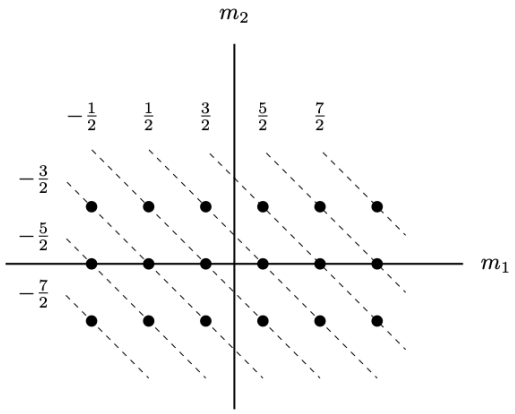
$$m = m_1 + m_2 \, , \tag{6.91}$$

since

$$0 = \langle j_1, m_1; j_2, m_2|J_z - J_{1z} - J_{2z}|j, m\rangle_C = \hbar(m - m_1 - m_2)\langle j_1, m_1; j_2, m_2|j, m\rangle_C \, . \tag{6.92}$$

[7]Note that we are restricting ourselves to those states with fixed j_1 and j_2, not summing therefore over them.

The condition can be represented graphically as



with each dot representing a vector $|j_1, m_1; j_2, m_2\rangle$ and the dashed lines selecting combination with $m = m_1 + m_2$.

Let us determine the possible states $|j_1, m_1; j_2, m_2\rangle$ corresponding to the total J_z quantum number $m = m_1 + m_2$. Assuming $j_1 \geq j_2$, without loss of generality, we have

$m = m_1 + m_2$	States	Number of them
$j_1 + j_2$	$\|j_1, j_1\rangle \otimes \|j_2, j_2\rangle$	1
$j_1 + j_2 - 1$	$\|j_1, j_1 - 1\rangle \otimes \|j_2, j_2\rangle$ $\|j_1, j_1\rangle \otimes \|j_2, j_2 - 1\rangle$	2
$j_1 + j_2 - 2$	$\|j_1, j_1 - 2\rangle \otimes \|j_2, j_2\rangle$ $\|j_1, j_1 - 1\rangle \otimes \|j_2, j_2 - 1\rangle$ $\|j_1, j_1\rangle \otimes \|j_2, j_2 - 2\rangle$	3
$\vdots$	$\vdots$	$\vdots$
$j_1 - j_2$	$\|j_1, j_1 - 2j_2\rangle \otimes \|j_2, j_2\rangle$ $\dots$ $\|j_1, j_1\rangle \otimes \|j_2, -j_2\rangle$	$2j_2 + 1$
$j_1 - j_2 - 1$	$\|j_1, j_1 - 2j_2 - 1\rangle \otimes \|j_2, j_2\rangle$ $\dots$ $\|j_1, j_1 - 1\rangle \otimes \|j_2, -j_2\rangle$	$2j_2 + 1$
$j_1 - j_2 - 2$	$\dots$	$2j_2 + 1$
$\vdots$	$\vdots$	$\vdots$
$j_2 - j_1$	$\|j_1, -j_1\rangle \otimes \|j_2, j_2\rangle$ $\dots$ $\|j_1, 2j_2 - j_1\rangle \otimes \|j_2, -j_2\rangle$	$2j_2 + 1$
$j_2 - j_1 - 1$	$\|j_1, -j_1\rangle \otimes \|j_2, j_2 - 1\rangle$ $\dots$ $\|j_1, 2j_2 - j_1 - 1\rangle \otimes \|j_2, -j_2\rangle$	$2j_2$
$\vdots$	$\vdots$	$\vdots$
$-j_1 - j_2$	$\|j_1, -j_1\rangle \otimes \|j_2, -j_2\rangle$	1

From this, we can determine the possible values of j by demanding that the set of states $|j_1, m_1; j_2, m_2\rangle$ yields complete multiplets $|j, m\rangle_C$, where m takes the $2j+1$ values $j \geq m \geq -m$ for each j.

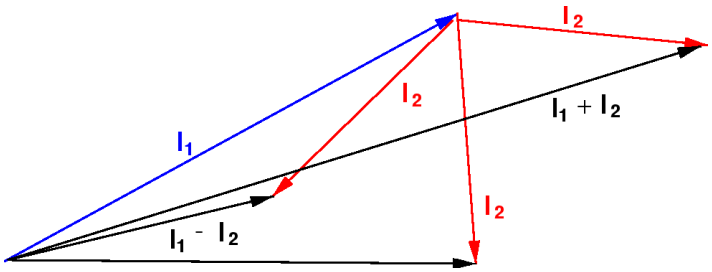
Let us start by considering the maximum value of j, corresponding to the maximum value of m, namely $j_1 + j_2$. This means that we have a total of $2(j_1 + j_2) + 1$ states $|j, m\rangle_C$ with $j = j_1 + j_2$ and $j \geq m \geq -j$.

Now, let's consider the value $m = j_1 + j_2 - 1$. For this, there are 2 states. A linear combination of these states belongs to the multiplet corresponding to $j = j_1 + j_2$. The other must correspond to the value $(j_1 + j_2 - 1)$ of j. Otherwise, this would contradict the fact that all $|j_1, m_1; j_2, m_2\rangle$ states can be grouped into complete multiplets $|j, m, j_1, j_2\rangle$.

The above reasoning can be continued as long as decreasing m by 1 increases the number of states with the given value of m also by 1. At it should be clear from the above table, this happens till the value $m = j_1 - j_2$. From there on, the number of states in the table does not increase further. That is, for $m = j_1 - j_2 - 1$, all the states are already included in multiplets with $j \geq j_1 - j_2$. Therefore, no lower values of j are possible. This implies

$$j_1 + j_2 \geq j \geq |j_1 - j_2|\,, \tag{6.93}$$

accounting in total to $2j_2 + 1$ different values ($j_1 \geq j_2$). This result is known as the triangle condition, as the three numbers j, j_1, j_2 must be such, that they could form a triangle.



Note that if both j_1 and j_2 are simultaneously half-integers or integers, the possible j's are integers, while if only one of them is half-integer, the j-values become half-integral.

A simpler proof

I like the above demonstration because it illustrates that the states $|j_1, m_1; j_2, m_2\rangle$ can contribute to different values of j, as it should be indeed expected since $[\vec{J}^2, J_{1z}] \neq 0$, $[\vec{J}^2, J_{1z}] \neq 0$. A less insightful but fully algebraic approach it to combine the already known relations

$$-j \leq m \leq j\,, \qquad -j_1 \leq m_1 \leq j_1\,, \qquad -j_2 \leq m_2 \leq j_2\,.$$

In particular, setting $m = j$ and $j_1 = m_1$ we have

$$m_2 = m - m_1 = j - j_1 \;\rightarrow\; -j_2 \leq j - j_1 \leq j_2 \;\rightarrow\; j_1 - j_2 \leq j \leq j_1 + j_2\,.$$

Similarly, setting $m = j$ and $m_2 = j_2$ we end up with

$$j_2 - j_1 \leq j \leq j_1 + j_2\,,$$

which combined with the previous case boils down to Eq. (6.93).

As an extra consistency check of the above results, let us determine the total number of states $|j,m\rangle_C$ and compare it with the number of states in the original basis $|j_1,m_1;j_2,m_2\rangle$, namely $N=(2j_1+1)(2j_2+1)$, as $j_1\geq m_1\geq -j_1$ and $j_2\geq m_2\geq -j_2$. Taking into account that for each value of j there are $(2j+1)$ states $|j,m\rangle_C$, we have $(j_1\geq j_2)$

$$\begin{aligned} N &= \sum_{j=j_1-j_2}^{j_1+j_2}(2j+1)=\sum_{k=0}^{2j_2}\left[2\left(k+j_1-j_2\right)+1\right]=2\sum_{k=0}^{2j_2}k+\left[2\left(j_1-j_2\right)+1\right](2j_2+1) \\ &= 2j_2\left(2j_2+1\right)+\left[2\left(j_1-j_2\right)+1\right](2j_2+1)=(2j_1+1)(2j_2+1), \end{aligned} \tag{6.94}$$

with $k\equiv j-(j_1-j_2)$. Voilà!

Example 35: Spin-orbit interactions

Consider a Hamiltonian

$$H=\beta\,\vec{L}\cdot\vec{S}\,, \tag{6.95}$$

with some constant β. The presence of this interaction precludes the use of the uncoupled basis $\{|l,m_l;s,m_s\rangle\}$ since $\vec{L}\cdot\vec{S}$ does not commute with L_z or S_z,[a]

$$\begin{aligned} \left[\vec{L}\cdot\vec{S},L_z\right] &= \left[L_x,L_z\right]S_x+\left[L_y,L_z\right]S_y=-i\hbar L_yS_x+i\hbar L_xS_y\,, \\ \left[\vec{L}\cdot\vec{S},S_z\right] &= L_x\left[S_x,S_z\right]+L_y\left[S_y,S_z\right]=-i\hbar L_xS_y+i\hbar L_yS_x\,. \end{aligned}$$

In order to find a new basis that diagonalizes this interaction, let us notice that

$$J^2=L^2+S^2+2\vec{L}\cdot\vec{S}\quad\longrightarrow\quad\vec{L}\cdot\vec{S}=\frac{1}{2}\left(J^2-L^2-S^2\right) \tag{6.96}$$

with $\vec{J}=\vec{L}+\vec{S}$ the total angular momentum. The operators $\left\{J^2,J_z,L^2,S^2\right\}$ all commute with each other. For simultaneous eigenstates $|j,m\rangle_C$, we have

$$\begin{aligned} J^2\,|j,m\rangle_C &= \hbar^2 j(j+1)\,|j,m\rangle_C\,, \\ J_z\,|j,m\rangle_C &= \hbar m\,|j,m\rangle_C\,, \\ L^2\,|j,m\rangle_C &= \hbar^2 l(l+1)\,|j,m\rangle_C\,, \\ S^2\,|j,m\rangle_C &= \hbar^2 s(s+1)\,|j,m\rangle_C\,, \end{aligned}$$

and therefore

$$\vec{L}\cdot\vec{S}\,|j,m\rangle_C=\frac{\hbar^2}{2}[j(j+1)-l(l+1)-s(s+1)]\,|j,m\rangle_C\,, \tag{6.97}$$

with j running from $|l-s|$ to $l+s$, and for each j,m runs from $-j$ to j. Note that the result is m-independent. This is a nontrivial consequence of $\vec{L}\cdot\vec{S}$ being a scalar under $\vec{J}$.

For one-electron atom (such as the hydrogen atom) with $s = 1/2$, there are two possible values of j for $l \geq 1, j = l \pm \frac{1}{2}$, and

$$j(j+1) - l(l+1) - s(s+1) = \left(l \pm \frac{1}{2}\right)\left(l + 1 \pm \frac{1}{2}\right) - l(l+1) - \frac{3}{4} = \begin{cases} l\,, \\ -(l+1)\,. \end{cases} \tag{6.98}$$

[a] Note that $\vec{L}$ and $\vec{S}$ are completely independent degrees of freedom, i.e. $[\vec{L}, \vec{S}] = 0$.

Example 36: Mesons and baryons

Quarks are elementary particles of spin 1/2. Due to the peculiarities of the strong interaction, these fundamental constituents of matter are never found in isolation but rather confined within hadrons. When two quarks, or more accurately a quark and an antiquark, combine, they form a meson, such as the pion or the kaon. When three quarks combine, they form a baryon, such as the proton or the neutron. Knowing the spin of the quarks, we can determine the spin of mesons and baryons by taking into account the angular momentum sum rule in Eq. (6.93).[a] Assuming the quarks to be in the ground state (so their orbital angular momentum is zero), we have:

- Mesons:

$$s_1 = 1/2\,, \quad s_2 = 1/2 \quad \longrightarrow \quad s_1 + s_2 \geq s_{12} \geq |s_1 - s_2| \quad \longrightarrow \quad s_{12} = 1, 0\,,$$

  in agreement with the observed spins for the lightest mesons (π 's and K 's have spin 0, ρ 's and ω 's have spin 1).

- Baryons:

$$s_1 = 1/2\,, \quad s_2 = 1/2 \quad \longrightarrow \quad s_1 + s_2 \geq s_{12} \geq |s_1 - s_2| \quad \longrightarrow \quad s_{12} = 1, 0\,,$$

$$s_{12} = 1\,, \quad s_3 = 1/2 \quad \longrightarrow \quad s_{12} + s_3 \geq s_{123} \geq |s_{12} - s_3| \quad \longrightarrow \quad s_{123} = 3/2, 1/2\,,$$

$$s_{12} = 0\,, \quad s_3 = 1/2 \quad \longrightarrow \quad s_{12} + s_3 \geq s_{123} \geq |s_{12} - s_3| \quad \longrightarrow \quad s_{123} = 1/2\,,$$

  in agreement with the observed spins for the lightest baryons (proton, neutron, etc... do carry spin 1/2 and $\Delta, \Omega^-, \ldots$ carry spin 3/2).

[a] This explanation is intentionally adapted to the level of this 3rd year course. The distribution of quarks and gluons inside particles like the proton is described by parton distribution functions (PDFs), which depend on the momentum fraction carried by each parton (quark or gluon) and the resolution scale of the probe (e.g., in deep inelastic scattering experiments). Also, the spin of the proton is not simply the sum of the spins of the valence quarks; contributions also come from the spin and orbital angular momentum of the sea quarks and gluons.

6.6 The Clebsch-Gordan coefficients

Given m_1 and m_2, the matrix elements (6.90) allow determining the quantum numbers j and m the whole system may possess and vice versa. Finding them is completely straight-forward but rather tedious: [8]

1. Start with the upper values of m and j,

$$m^{\rm up} = m_1^{\rm up} + m_2^{\rm up} = j_1 + j_2 \qquad \longrightarrow \qquad j^{\rm up} = j_1 + j_2 \,, \tag{6.99}$$

Since $m_1 + m_2 = m$, this singles out a unique state in the sum (6.88) that, as any other must be normalized to unity, namely

$$|j_1 + j_2, j_1 + j_2\rangle_C = |j_1, j_1; j_2, j_2\rangle \, \langle j_1, j_1; j_2, j_2 | j_1 + j_2, j_1 + j_2\rangle_C = |j_1, j_1; j_2, j_2\rangle \ , \tag{6.100}$$

where we have adopted a phase convention

$$\langle j_1, j_1; j_2, j_2 | j_1 + j_2, j_1 + j_2\rangle_C = 1 \,. \tag{6.101}$$

This fixes our first matrix element in (6.90).

2. Apply recursively the lowering operator $J_- = J_{1-} + J_{2-}$ and Eq. (6.47) to systematically determine all others terms involving the same $j = j_1 + j_2$ and a reduced m, decreasing the latter quantum number step by step all the way down till its minimum value $-(j_1 + j_2)$, where the whole thing stops. This generates the $2\,(j_1 + j_2) + 1$ states forming the irreducible representation $j = j_1 + j_2$, For instance, the Clebsch-Gordan coefficients

$$\langle j_1, j_1 - 1; j_2, j_2 | j_1 + j_2, j_1 + j_2 - 1\rangle_C \qquad \text{and} \qquad \langle j_1, j_1; j_2, j_2 - 1 | j_1 + j_2, j_1 + j_2 - 1\rangle_C \tag{6.102}$$

can be read off from

$$\begin{aligned} |j^{\rm up}, m^{\rm up} - 1\rangle &= \frac{J_- \, |j^{\rm up}, m^{\rm up}\rangle}{\sqrt{j^{\rm up}(j^{\rm up} + 1) - m^{\rm up}\,(m^{\rm up} - 1)}} = \frac{(J_{1-} + J_{2-})\, |j_1, j_1; j_2, j_2\rangle}{\sqrt{2 j^{\rm up}}} \\ &= \sqrt{\frac{j_1}{j^{\rm up}}}\, |j_1, j_1 - 1; j_2, j_2\rangle + \sqrt{\frac{j_2}{j^{\rm up}}}\, |j_1, j_1; j_2, j_2 - 1\rangle \ , \end{aligned} \tag{6.103}$$

namely

$$\langle j_1, j_1 - 1; j_2, j_2 | j_1 + j_2, j_1 + j_2 - 1\rangle_C = \sqrt{\frac{j_1}{j_1 + j_2}} \,, \tag{6.104}$$

$$\langle j_1, j_1; j_2, j_2 - 1 | j_1 + j_2, j_1 + j_2 - 1\rangle_C = \sqrt{\frac{j_2}{j_1 + j_2}} \,. \tag{6.105}$$

3. Repeat the iterative process again, starting this time with the state $|j_1 + j_2 - 1, j_1 + j_2 - 1\rangle_C$, which can be fixed up to a phase factor by requiring it to be properly normalized and orthogonal to the state $|j_1 + j_2, j_1 + j_2 - 1\rangle$ you previously determined.

[8] Note also that the notation in Eq. (6.90) is certainly not unique. Indeed, a variety of relabellings and permutations of its arguments can be found in the literature, together with alternative notations such as $C(j_1, j_2, m_1, m_2; j, m)$. It is therefore very important that you understand the overall logic of the computation, avoiding cherry-picking formulas here and there!

4. Of course, one could multiply the states by any arbitrary phase and the coefficients would work as well. Here, we take the Clebsch-Gordan coefficients to be real.

Example 37: Clebsch-Gordan coefficients for two spin 1/2 particles

To illustrate the above procedure, let us consider our well known system of two spin 1/2 particles. This exercise is useful for the computation of the hyperfine splitting in the hydrogen atom, where the relevant spins are those of the proton and the electron. Identifying $J_i \to S_i$ and $j_1 = s_1 = 1/2$, $j_2 = s_2 = \frac{1}{2}$, we can construct $N = (2s_1+1)(2s_2+1) = 4$ states in the common basis of the operators $S_1^2, S_2^2, S_{1z}, S_{2z}$, namely

$$|1/2,1/2;1/2,1/2\rangle \equiv |++\rangle\,, \qquad |1/2,1/2;1/2,-1/2\rangle \equiv |+-\rangle\,,$$
$$|1/2,-1/2;1/2,1/2\rangle \equiv |-+\rangle\,, \qquad |1/2,-1/2;1/2,-1/2\rangle \equiv |--\rangle\,.$$

In agreement with the composition law (6.93), the total angular momentum of the system

$$\vec{S} = \vec{S}_1 + \vec{S}_2$$

can take only two values $s_1+s_2 = 1$ and $s_1 - s_2 = 0$. To calculate the Clebsch-Gordan coefficients relating the coupled basis $|s,m\rangle_C$ to the uncoupled one $|s_1,m_1;s_2,m_2\rangle$, we consider the coupled basis state with higher values of s and m, which can be only obtained for $m_1 = m_2 = 1/2$, i.e.

$$|1,1\rangle_C = |++\rangle\,.$$

Using the ladder operator S_- in both sides of this equation, we obtain

$$S_-|1,1\rangle_C = \hbar\sqrt{s(s+1)-m(m-1)}|1,0\rangle_C = \hbar\sqrt{2}|1,0\rangle_C\,,$$
$$S_-|1,1\rangle_C = (S_{1-}+S_{2-})|++\rangle = \hbar\sqrt{\frac{1}{2}\left(\frac{1}{2}+1\right)-\frac{1}{2}\left(\frac{1}{2}-1\right)}(|-+\rangle+|+-\rangle)$$
$$= \hbar(|+-\rangle+|-+\rangle),$$

or equivalently

$$|1,0\rangle_C = \frac{1}{\sqrt{2}}(|+-\rangle+|-+\rangle)\,.$$

Applying S_- again to this state, we find

$$|1,-1\rangle_C = |--\rangle\,.$$

This completes the triplet of states forming the irreducible representation $s = 1$, namely

$$|1,1\rangle_C = |++\rangle\,, \quad |1,0\rangle_C = \frac{1}{\sqrt{2}}(|+-\rangle+|-+\rangle)\,, \quad |1,-1\rangle_C = |--\rangle\,. \tag{6.106}$$

The remaining state with $s = m = 0$ state can be determined by requiring it to be orthogonal to the state $|1,0\rangle_C$ previously determined. This provides the singlet state

$$|0,0\rangle = \frac{1}{\sqrt{2}}(|+-\rangle - |-+\rangle)\,. \tag{6.107}$$

The above results are often expressed by saying the product space of two spins 1/2 is the direct sum of a spin one space and a spin zero space,

$$\frac{1}{2} \otimes \frac{1}{2} = 1 \oplus 0\,. \tag{6.108}$$

As we showed in Section 5.4.2, the spin-one representation consists of bosonic (symmetric) states, while the spin-zero representation is a fermionic (anti-symmetric) state. The total angular momentum operators can be seen to commute with the action of permutations on the triplet and singlet states. For instance, for J_z and the triplet state, we have, by simple inspection, that

$$\begin{aligned} J_z|1,1\rangle &= \hbar|1,1\rangle\,, & \pi|1,1\rangle &= |1,1\rangle\,,\\ J_z|1,0\rangle &= 0\,, & \pi|1,0\rangle &= |1,0\rangle\,,\\ J_z|1,-1\rangle &= -\hbar|1,-1\rangle\,, & \pi|1,-1\rangle &= |1,-1\rangle\,, \end{aligned}$$

and therefore $J_z\,\pi|1,m\rangle = \pi J_z|1,m\rangle$. A similar analysis can be done for the other momentum components and the singlet state, confirming that the total angular momentum operators commute generically with the permutation operator, This means that the bosonic and fermionic subspaces will always transform among themselves under rotations.

It is rather evident that developing this algorithm becomes quite laborious for states with angular momentum beyond the lower values. Luckily for us, the most common Clebsch-Gordan coefficients have been determined with computer programs and tabulated once and for all. There most conventional tabulations can be found in various places including the Particle Data Group site and the end of this section.[9] This tables are, however, rather confusing and require some explanation. There is one table of the form

$j_1 \otimes j_2$		j	j	$\cdots$
		m	m	$\cdots$
m_1	m_2	CGC		
m_1	m_2			
$\vdots$	$\vdots$			

for each j_1, j_2 pair, with the possible values of j and m indicated at the top of the table and those of m_1 and m_2 on the left. The associated Clebsch-Gordan coefficients (CGC) appear in the intersection cell, in some compact but tricky notation: the displayed numbers are the *coefficients squared times their sign*.

[9] Note that, as compared to this text, j and m are capitalized in the table.

Example 38: Use of CGC tables

For instance, for $j_1 = 1, j_2 = 1/2$ and $m = m_1 + m_2 = 1/2$, we have

$j_1 \otimes j_2$		$3/2$	$1/2$
		$+1/2$	$+1/2$
$+1$	$-1/2$	$1/3$	$2/3$
0	$+1/2$	$2/3$	$-1/3$

Taking into account the underlined comment above (and in particular, that $-1/3$ really means $-\sqrt{1/3}$), this table tell us that

$$|3/2, 1/2\rangle_C = \sqrt{\frac{1}{3}}|1, 1; 1/2, -1/2\rangle + \sqrt{\frac{2}{3}}|1, 0; 1/2, 1/2\rangle \,,$$

and

$$|1/2, 1/2\rangle_C = \sqrt{\frac{2}{3}}|1, 1; 1/2, -1/2\rangle - \sqrt{\frac{1}{3}}|1, 0; 1/2, 1/2\rangle \,.$$

Note that although the tabulation of the Particle Data Group does not contain the related case $j_1 = 1/2, j_2 = 1$, this can be always be obtained through the relation

$$\langle j_2, m_2; j_1, m_1|j, m\rangle_C = (-1)^{j-j_1-j_2} \langle j_1, m_1; j_2, m_2|j, m\rangle_C \,. \tag{6.109}$$

which we provide without proof.

Example 39: A spin-spin interaction

Consider a system of two particles of spins $s_1 = 3/2$ and $s_2 = 1/2$ described by the effective Hamiltonian

$$H = \frac{\omega}{\hbar}\, \vec{S}_1 \cdot \vec{S}_2 \,,$$

with $\omega > 0$ a constant and $\vec{S}_1$ and $\vec{S}_2$ the spin operators of the two particles. If the system is initially prepared in an eigenstate $|s_1, m_{s_1}; s_2, m_{s_2}\rangle = |3/2, 1/2; 1/2, 1/2\rangle$ of S_1^2, S_{1z}, S_2^2, and S_{2z}, which is the probability of finding it in the state $|s_1', m_{s_1}'; s_2', m_{s_2}'\rangle = |3/2, 3/2; 1/2, -1/2\rangle$ at later times?
Since $\vec{S}_1 \cdot \vec{S}_2$ does not commute with the z components of $\vec{S}_1$ and $\vec{S}_2$, the set of operators $\{S_1^2, S_2^2, S_{1z}, S_{2z}\}$ do not form a CSOC. On the other hand, the set of operators $\left\{S^2, S_1^2, S_2^2\right\}$ with $\vec{S} = \vec{S}_1 + \vec{S}_2$ the total spin of the system, commute with each other, thus providing an appropriate *coupled basis* $|s\, m_s\rangle_C$ for diagonalizing H, with s and m_s the eigenvalues of the total spin operator and its third component. In this basis, we have indeed

$$_C\langle s, m_s|H|s, m_s\rangle_C = \frac{\omega}{2\hbar}{}_C\langle s, m_s| \left(S^2 - S_1^2 - S_2^2\right)|s, m_s\rangle_C = \frac{\hbar\omega}{2}\left[s(s+1) - \frac{9}{2}\right] \,,$$

which, with the allowed values of the total spin $s = s_1 \pm s_2 = 1, 2$, leads to the energy

eigenvalues

$$E_1 = -5\omega\hbar/4\,, \qquad E_2 = 3\omega\hbar/4\,.$$

The eigenstates of S^2 and S_z will be also stationary states. These can be expressed in terms of the *uncoupled basis* $|s_1, m_{s_1}; s_2, m_{s_2}\rangle$ through some Clebsch-Gordan coefficients a, b, c and d, namely

$$|11\rangle_C = a\left|\frac{3}{2},\frac{3}{2};\frac{1}{2},-\frac{1}{2}\right\rangle + b\left|\frac{3}{2},\frac{1}{2};\frac{1}{2},\frac{1}{2}\right\rangle\,, \quad |21\rangle_C = c\left|\frac{3}{2},\frac{1}{2};\frac{1}{2},\frac{1}{2}\right\rangle + d\left|\frac{3}{2},\frac{3}{2};\frac{1}{2},-\frac{1}{2}\right\rangle\,.$$

These coefficients can be either directly read off from the table of Clebsch-Gordan coefficients or easily determined from the action of the ladder operators. Considering for instance the action of the ladder operator $S_+ = S_{1,+} + S_{2,+}$ in the lowest energy state $|11\rangle$, we obtain

$$\begin{aligned} S_+|11\rangle_C = 0 &= aS_{2,+}\left|\frac{3}{2},\frac{3}{2};\frac{1}{2},-\frac{1}{2}\right\rangle + bS_{1,+}\left|\frac{3}{2},\frac{1}{2};\frac{1}{2},\frac{1}{2}\right\rangle \\ &= a\hbar\left|\frac{3}{2},\frac{3}{2};\frac{1}{2},\frac{1}{2}\right\rangle + b\hbar\sqrt{3}\left|\frac{3}{2},\frac{3}{2};\frac{1}{2},\frac{1}{2}\right\rangle\,, \end{aligned}$$

which implies $a + \sqrt{3}b = 1$ and, therefore, the (normalized) coefficients

$$a = \frac{\sqrt{3}}{2}\,, \qquad b = -\frac{1}{2}\,.$$

Similarly, acting with the ladder operator in the highest energy state $|21\rangle$, we get

$$S_+|21\rangle_C = 2\hbar|22\rangle_C = 2\hbar\left|\frac{3}{2},\frac{3}{2};\frac{1}{2},\frac{1}{2}\right\rangle = c\hbar\sqrt{3}\left|\frac{3}{2},\frac{3}{2};\frac{1}{2},\frac{1}{2}\right\rangle + d\hbar\left|\frac{3}{2},\frac{3}{2};\frac{1}{2},\frac{1}{2}\right\rangle\,,$$

and, therefore, $\sqrt{3}c + d = 2$ or

$$d = \frac{1}{2}\,, \qquad c = \frac{\sqrt{3}}{2}\,.$$

Collecting these results, we have

$$\begin{aligned} |11\rangle_C &= \frac{\sqrt{3}}{2}\left|\frac{3}{2},\frac{3}{2};\frac{1}{2},-\frac{1}{2}\right\rangle - \frac{1}{2}\left|\frac{3}{2},\frac{1}{2};\frac{1}{2},\frac{1}{2}\right\rangle\,, \\ |21\rangle_C &= \frac{1}{2}\left|\frac{3}{2},\frac{3}{2};\frac{1}{2},-\frac{1}{2}\right\rangle + \frac{\sqrt{3}}{2}\left|\frac{3}{2},\frac{1}{2};\frac{1}{2},\frac{1}{2}\right\rangle\,, \end{aligned}$$

with inverse relations

$$\left|\frac{3}{2},\frac{1}{2};\frac{1}{2},\frac{1}{2}\right\rangle = \frac{\sqrt{3}}{2}|21\rangle_C - \frac{1}{2}|11\rangle_C\,, \qquad \left|\frac{3}{2},\frac{3}{2};\frac{1}{2},-\frac{1}{2}\right\rangle = \frac{1}{2}|21\rangle_C + \frac{\sqrt{3}}{2}|11\rangle_C\,.$$

The evolution of the initial state $|\varphi, 0\rangle = |3/2, 1/2; 1/2, 1/2\rangle$ is then given by

$$|\varphi(t)\rangle = \frac{1}{2}\left(\sqrt{3}e^{-iE_2t/\hbar}|21\rangle_C - e^{-iE_1t/\hbar}|11\rangle_C\right)\,.$$

The requested probability of finding the system in the state $\left|\frac{3}{2},\frac{3}{2};\frac{1}{2},-\frac{1}{2}\right\rangle$ at time $t > 0$ is therefore

$$\mathcal{P} = \left|\left\langle\frac{3}{2}\frac{3}{2};\frac{1}{2},-\frac{1}{2}\middle|\varphi(t)\right\rangle\right|^2 = \frac{3}{8}\left(1 - \cos\frac{(E_2 - E_1)\,t}{\hbar}\right) = \frac{3}{4}\sin^2\frac{(E_2 - E_1)\,t}{2\hbar} = \frac{3}{4}\sin^2\omega t\,.$$

34. CLEBSCH-GORDAN COEFFICIENTS, SPHERICAL HARMONICS, AND *d* FUNCTIONS

Note: A square-root sign is to be understood over *every* coefficient, *e.g.*, for $-8/15$ read $-\sqrt{8/15}$.

Notation:

		J J ...
		M M ...
m_1	m_2	Coefficients
m_1	m_2	
.	.	

1/2×1/2

m_1	m_2	J=1, M=+1
+1/2	+1/2	1

m_1	m_2	J=1, M=0	J=0, M=0
+1/2	−1/2	1/2	1/2
−1/2	+1/2	1/2	−1/2

m_1	m_2	J=1, M=−1
−1/2	−1/2	1

1×1/2

m_1	m_2	J=3/2, M=+3/2
+1	+1/2	1

m_1	m_2	J=3/2, M=+1/2	J=1/2, M=+1/2
+1	−1/2	1/3	2/3
0	+1/2	2/3	−1/3

m_1	m_2	J=3/2, M=−1/2	J=1/2, M=−1/2
0	−1/2	2/3	1/3
−1	+1/2	1/3	−2/3

m_1	m_2	J=3/2, M=−3/2
−1	−1/2	1

2×1

m_1	m_2	J=3, M=+3
+2	+1	1

m_1	m_2	J=3, M=+2	J=2, M=+2
+2	0	1/3	2/3
+1	+1	2/3	−1/3

m_1	m_2	J=3, M=+1	J=2, M=+1	J=1, M=+1
+2	−1	1/15	1/3	3/5
+1	0	8/15	1/6	−3/10
0	+1	2/5	−1/2	1/10

m_1	m_2	J=3, M=0	J=2, M=0	J=1, M=0
+1	−1	1/5	1/2	3/10
0	0	3/5	0	−2/5
−1	+1	1/5	−1/2	3/10

m_1	m_2	J=3, M=−1	J=2, M=−1	J=1, M=−1
0	−1	2/5	1/2	1/10
−1	0	8/15	−1/6	−3/10
−2	+1	1/15	−1/3	3/5

m_1	m_2	J=3, M=−2	J=2, M=−2
−1	−1	2/3	1/3
−2	0	1/3	−2/3

m_1	m_2	J=3, M=−3
−2	−1	1

1×1

m_1	m_2	J=2, M=+2
+1	+1	1

m_1	m_2	J=2, M=+1	J=1, M=+1
+1	0	1/2	1/2
0	+1	1/2	−1/2

m_1	m_2	J=2, M=0	J=1, M=0	J=0, M=0
+1	−1	1/6	1/2	1/3
0	0	2/3	0	−1/3
−1	+1	1/6	−1/2	1/3

m_1	m_2	J=2, M=−1	J=1, M=−1
0	−1	1/2	1/2
−1	0	1/2	−1/2

m_1	m_2	J=2, M=−2
−1	−1	1

$$Y_1^0 = \sqrt{\frac{3}{4\pi}}\cos\theta$$

$$Y_1^1 = -\sqrt{\frac{3}{8\pi}}\sin\theta\, e^{i\phi}$$

$$Y_2^0 = \sqrt{\frac{5}{4\pi}}\Big(\frac{3}{2}\cos^2\theta - \frac{1}{2}\Big)$$

$$Y_2^1 = -\sqrt{\frac{15}{8\pi}}\sin\theta\cos\theta\, e^{i\phi}$$

$$Y_2^2 = \frac{1}{4}\sqrt{\frac{15}{2\pi}}\sin^2\theta\, e^{2i\phi}$$

$$Y_\ell^{-m} = (-1)^m Y_\ell^{m*}$$

$$d^{\,\ell}_{m,0} = \sqrt{\frac{4\pi}{2\ell+1}}\, Y_\ell^m\, e^{-im\phi}$$

2×1/2

m_1	m_2	J=5/2, M=+5/2
+2	+1/2	1

m_1	m_2	J=5/2, M=+3/2	J=3/2, M=+3/2
+2	−1/2	1/5	4/5
+1	+1/2	4/5	−1/5

m_1	m_2	J=5/2, M=+1/2	J=3/2, M=+1/2
+1	−1/2	2/5	3/5
0	+1/2	3/5	−2/5

m_1	m_2	J=5/2, M=−1/2	J=3/2, M=−1/2
0	−1/2	3/5	2/5
−1	+1/2	2/5	−3/5

m_1	m_2	J=5/2, M=−3/2	J=3/2, M=−3/2
−1	−1/2	4/5	1/5
−2	+1/2	1/5	−4/5

m_1	m_2	J=5/2, M=−5/2
−2	−1/2	1

3/2×1/2

m_1	m_2	J=2, M=+2
+3/2	+1/2	1

m_1	m_2	J=2, M=+1	J=1, M=+1
+3/2	−1/2	1/4	3/4
+1/2	+1/2	3/4	−1/4

m_1	m_2	J=2, M=0	J=1, M=0
+1/2	−1/2	1/2	1/2
−1/2	+1/2	1/2	−1/2

m_1	m_2	J=2, M=−1	J=1, M=−1
−1/2	−1/2	3/4	1/4
−3/2	+1/2	1/4	−3/4

m_1	m_2	J=2, M=−2
−3/2	−1/2	1

3/2×1

m_1	m_2	J=5/2, M=+5/2
+3/2	+1	1

m_1	m_2	J=5/2, M=+3/2	J=3/2, M=+3/2
+3/2	0	2/5	3/5
+1/2	+1	3/5	−2/5

m_1	m_2	J=5/2, M=+1/2	J=3/2, M=+1/2	J=1/2, M=+1/2
+3/2	−1	1/10	2/5	1/2
+1/2	0	3/5	1/15	−1/3
−1/2	+1	3/10	−8/15	1/6

m_1	m_2	J=5/2, M=−1/2	J=3/2, M=−1/2	J=1/2, M=−1/2
+1/2	−1	3/10	8/15	1/6
−1/2	0	3/5	−1/15	−1/3
−3/2	+1	1/10	−2/5	1/2

m_1	m_2	J=5/2, M=−3/2	J=3/2, M=−3/2
−1/2	−1	3/5	2/5
−3/2	0	2/5	−3/5

m_1	m_2	J=5/2, M=−5/2
−3/2	−1	1

$$\langle j_1 j_2 m_1 m_2 | j_1 j_2 J M\rangle = (-1)^{J-j_1-j_2}\langle j_2 j_1 m_2 m_1 | j_2 j_1 J M\rangle$$

$$d^{\,j}_{m',m} = (-1)^{m-m'} d^{\,j}_{m,m'} = d^{\,j}_{-m,-m'}$$

3/2×3/2

m_1	m_2	J=3, M=+3
+3/2	+3/2	1

m_1	m_2	J=3, M=+2	J=2, M=+2
+3/2	+1/2	1/2	1/2
+1/2	+3/2	1/2	−1/2

m_1	m_2	J=3, M=+1	J=2, M=+1	J=1, M=+1
+3/2	−1/2	1/5	1/2	3/10
+1/2	+1/2	3/5	0	−2/5
−1/2	+3/2	1/5	−1/2	3/10

m_1	m_2	J=3, M=0	J=2, M=0	J=1, M=0	J=0, M=0
+3/2	−3/2	1/20	1/4	9/20	1/4
+1/2	−1/2	9/20	1/4	−1/20	−1/4
−1/2	+1/2	9/20	−1/4	−1/20	1/4
−3/2	+3/2	1/20	−1/4	9/20	−1/4

m_1	m_2	J=3, M=−1	J=2, M=−1	J=1, M=−1
+1/2	−3/2	1/5	1/2	3/10
−1/2	−1/2	3/5	0	−2/5
−3/2	+1/2	1/5	−1/2	3/10

m_1	m_2	J=3, M=−2	J=2, M=−2
−1/2	−3/2	1/2	1/2
−3/2	−1/2	1/2	−1/2

m_1	m_2	J=3, M=−3
−3/2	−3/2	1

$$d^{\,1}_{0,0} = \cos\theta \qquad d^{\,1/2}_{1/2,1/2} = \cos\frac{\theta}{2} \qquad d^{\,1}_{1,1} = \frac{1+\cos\theta}{2}$$

$$d^{\,1/2}_{1/2,-1/2} = -\sin\frac{\theta}{2} \qquad d^{\,1}_{1,0} = -\frac{\sin\theta}{\sqrt{2}}$$

$$d^{\,1}_{1,-1} = \frac{1-\cos\theta}{2}$$

2×3/2

m_1	m_2	J=7/2, M=+7/2
+2	+3/2	1

m_1	m_2	J=7/2, M=+5/2	J=5/2, M=+5/2
+2	+1/2	3/7	4/7
+1	+3/2	4/7	−3/7

m_1	m_2	J=7/2, M=+3/2	J=5/2, M=+3/2	J=3/2, M=+3/2
+2	−1/2	1/7	16/35	2/5
+1	+1/2	4/7	1/35	−2/5
0	+3/2	2/7	−18/35	1/5

m_1	m_2	J=7/2, M=+1/2	J=5/2, M=+1/2	J=3/2, M=+1/2	J=1/2, M=+1/2
+2	−3/2	1/35	6/35	2/5	2/5
+1	−1/2	12/35	5/14	0	−3/10
0	+1/2	18/35	−3/35	−1/5	1/5
−1	+3/2	4/35	−27/70	2/5	−1/10

m_1	m_2	J=7/2, M=−1/2	J=5/2, M=−1/2	J=3/2, M=−1/2	J=1/2, M=−1/2
+1	−3/2	4/35	27/70	2/5	1/10
0	−1/2	18/35	3/35	−1/5	−1/5
−1	+1/2	12/35	−5/14	0	3/10
−2	+3/2	1/35	−6/35	2/5	−2/5

m_1	m_2	J=7/2, M=−3/2	J=5/2, M=−3/2	J=3/2, M=−3/2
0	−3/2	2/7	18/35	1/5
−1	−1/2	4/7	−1/35	−2/5
−2	+1/2	1/7	−16/35	2/5

m_1	m_2	J=7/2, M=−5/2	J=5/2, M=−5/2
−1	−3/2	4/7	3/7
−2	−1/2	3/7	−4/7

m_1	m_2	J=7/2, M=−7/2
−2	−3/2	1

2×2

m_1	m_2	J=4, M=+4
+2	+2	1

m_1	m_2	J=4, M=+3	J=3, M=+3
+2	+1	1/2	1/2
+1	+2	1/2	−1/2

m_1	m_2	J=4, M=+2	J=3, M=+2	J=2, M=+2
+2	0	3/14	1/2	2/7
+1	+1	4/7	0	−3/7
0	+2	3/14	−1/2	2/7

m_1	m_2	J=4, M=+1	J=3, M=+1	J=2, M=+1	J=1, M=+1
+2	−1	1/14	3/10	3/7	1/5
+1	0	3/7	1/5	−1/14	−3/10
0	+1	3/7	−1/5	−1/14	3/10
−1	+2	1/14	−3/10	3/7	−1/5

m_1	m_2	J=4, M=0	J=3, M=0	J=2, M=0	J=1, M=0	J=0, M=0
+2	−2	1/70	1/10	2/7	2/5	1/5
+1	−1	8/35	2/5	1/14	−1/10	−1/5
0	0	18/35	0	−2/7	0	1/5
−1	+1	8/35	−2/5	1/14	1/10	−1/5
−2	+2	1/70	−1/10	2/7	−2/5	1/5

m_1	m_2	J=4, M=−1	J=3, M=−1	J=2, M=−1	J=1, M=−1
+1	−2	1/14	3/10	3/7	1/5
0	−1	3/7	1/5	−1/14	−3/10
−1	0	3/7	−1/5	−1/14	3/10
−2	+1	1/14	−3/10	3/7	−1/5

m_1	m_2	J=4, M=−2	J=3, M=−2	J=2, M=−2
0	−2	3/14	1/2	2/7
−1	−1	4/7	0	−3/7
−2	0	3/14	−1/2	2/7

m_1	m_2	J=4, M=−3	J=3, M=−3
−1	−2	1/2	1/2
−2	−1	1/2	−1/2

m_1	m_2	J=4, M=−4
−2	−2	1

$$d^{\,3/2}_{3/2,3/2} = \frac{1+\cos\theta}{2}\cos\frac{\theta}{2}$$

$$d^{\,3/2}_{3/2,1/2} = -\sqrt{3}\frac{1+\cos\theta}{2}\sin\frac{\theta}{2} \qquad d^{\,2}_{2,2} = \Big(\frac{1+\cos\theta}{2}\Big)^2$$

$$d^{\,3/2}_{3/2,-1/2} = \sqrt{3}\frac{1-\cos\theta}{2}\cos\frac{\theta}{2} \qquad d^{\,2}_{2,1} = -\frac{1+\cos\theta}{2}\sin\theta$$

$$d^{\,3/2}_{3/2,-3/2} = -\frac{1-\cos\theta}{2}\sin\frac{\theta}{2} \qquad d^{\,2}_{2,0} = \frac{\sqrt{6}}{4}\sin^2\theta \qquad d^{\,2}_{1,1} = \frac{1+\cos\theta}{2}(2\cos\theta-1)$$

$$d^{\,3/2}_{1/2,1/2} = \frac{3\cos\theta-1}{2}\cos\frac{\theta}{2} \qquad d^{\,2}_{2,-1} = -\frac{1-\cos\theta}{2}\sin\theta \qquad d^{\,2}_{1,0} = -\sqrt{\frac{3}{2}}\sin\theta\,\cos\theta$$

$$d^{\,3/2}_{1/2,-1/2} = -\frac{3\cos\theta+1}{2}\sin\frac{\theta}{2} \qquad d^{\,2}_{2,-2} = \Big(\frac{1-\cos\theta}{2}\Big)^2 \qquad d^{\,2}_{1,-1} = \frac{1-\cos\theta}{2}(2\cos\theta+1) \qquad d^{\,2}_{0,0} = \Big(\frac{3}{2}\cos^2\theta - \frac{1}{2}\Big)$$

Figure 34.1: The sign convention is that of Wigner (*Group Theory*, Academic Press, New York, 1959), also used by Condon and Shortley (*The Theory of Atomic Spectra*, Cambridge Univ. Press, New York, 1953), Rose (*Elementary Theory of Angular Momentum*, Wiley, New York, 1957), and Cohen (*Tables of the Clebsch-Gordan Coefficients*, North American Rockwell Science Center, Thousand Oaks, Calif., 1974). The coefficients here have been calculated using computer programs written independently by Cohen and at LBNL.

Particle	t	t_z	j	q	Quark content	Mass [MeV]
p	$\frac{1}{2}$	$+\frac{1}{2}$	$\frac{1}{2}$	$+1$	uud	$938.27208816 \pm 0.00000029$
n	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	0	udd	$939.56542052 \pm 0.00000054$
Δ^{++}	$\frac{3}{2}$	$+\frac{3}{2}$	$\frac{3}{2}$	$+2$	uuu	1232 ± 2
Δ^{+}	$\frac{3}{2}$	$+\frac{1}{2}$	$\frac{3}{2}$	$+1$	uud	1232 ± 2
Δ^{0}	$\frac{3}{2}$	$-\frac{1}{2}$	$\frac{3}{2}$	0	udd	1232 ± 2
Δ^{-}	$\frac{3}{2}$	$-\frac{3}{2}$	$\frac{3}{2}$	-1	ddd	1232 ± 2
π^{+}	1	$+1$	0	$+1$	$u\bar{d}$	139.57039 ± 0.00018
π^{0}	1	0	0	0	$\frac{1}{\sqrt{2}}(u\bar{u} - d\bar{d})$	134.9768 ± 0.0005
π^{-}	1	-1	0	-1	$d\bar{u}$	139.57039 ± 0.00018
ρ^{+}	1	$+1$	1	$+1$	$u\bar{d}$	775.26 ± 0.25
ρ^{0}	1	0	1	0	$\frac{1}{\sqrt{2}}(u\bar{u} - d\bar{d})$	775.26 ± 0.25
ρ^{-}	1	-1	1	-1	$d\bar{u}$	775.26 ± 0.25

Table 6.1: Properties of selected hadrons: isospin t, projection t_z, spin j, electric charge q, quark content, and mass (in MeV). Values from the Particle Data Group.

6.7 Isospin: protons and neutrons in disguise

The formalism developed for the addition of angular momenta—whether spin, orbital, or total—applies far beyond these specific degrees of freedom. A notable example is *isospin*, which arises in nuclear and particle physics as an approximate symmetry of the strong interaction.

The concept of isospin (or *isotopic spin*) was introduced by Werner Heisenberg in 1932, shortly after the discovery of the neutron. At the time, it was observed that the proton and neutron have nearly identical masses and interact symmetrically under the strong nuclear force, despite their differing electric charges. To explain this, Heisenberg proposed that the two nucleons could be viewed as two states of a single particle type—the *nucleon*—distinguished by a new internal quantum number that behaves like angular momentum. This quantum number was later dubbed *isospin* due to its formal similarity to spin.

Isospin is treated mathematically using the SU(2) symmetry group, just like ordinary spin. The proton and neutron are organized into an isospin doublet

$$\begin{pmatrix} |p\rangle \\ |n\rangle \end{pmatrix} = \begin{pmatrix} |t = 1/2, t_z = 1/2\rangle \\ |t = 1/2, t_z = -1/2\rangle \end{pmatrix}. \tag{6.110}$$

The isospin operators $\vec{T} = (T_x, T_y, T_z)$ satisfy the SU(2) commutation relations $[T_i, T_j] = i\hbar\epsilon_{ijk}T_k$ and the total isospin operator T^2 has eigenvalues $t(t+1)\hbar^2$. As for spin, ladder operators $T_\pm = T_x \pm iT_y$ allow one to move between states within a multiplet. This formalism allows us to describe nuclear and hadronic states in terms of isospin multiplets, simplifying the classification of particles and predicting symmetries in reactions governed by the strong force. For instance, the pions form an isospin triplet

$$|\pi^+\rangle = |t = 1, t_z = +1\rangle \ , \quad |\pi^0\rangle = |t = 1, t_z = 0\rangle \ , \quad |\pi^-\rangle = |t = 1, t_z = -1\rangle \ . \tag{6.111}$$

The above groupings reflect the fact that strong interactions are approximately invariant under isospin transformations, i.e. they depend on the total isospin, not on the charge state.

Consequently, many hadronic processes conserve both T and T_z, yielding selection rules and degeneracies analogous to angular momentum conservation. As isospin follows the SU(2) structure of spin, we can also apply the machinery of angular momentum addition. For two particles with isospins t_1 and t_2, the total isospin t ranges from $|t_1 - t_2|$ to $t_1 + t_2$ in integer steps. The states of definite total isospin are constructed using *Clebsch–Gordan coefficients*, just as in spin systems.

Example 40: Isospin decomposition in $\Delta \to \pi N$ decay

The Δ baryons form an isospin quartet with total isospin $t = 3/2$. Each member of this multiplet can decay strongly into a pion and a nucleon. Since pions form an isospin triplet with $t = 1$, and nucleons (proton and neutron) form a doublet with $t = 1/2$, we are dealing with a system in which the total isospin in the final state can be analyzed by coupling these two representations.

Let us consider for instance the decay of the Δ^+, which has isospin projection $t_z = 1/2$. The allowed final states must also have $t_z = 1/2$, which can occur in two combinations: either a π^0 (with $t_z = 0$) and a proton (with $t_z = 1/2$), or a π^+ (with $t_z = +1$) and a neutron (with $t_z = -1/2$). Both of these channels conserve isospin projection and are therefore allowed.

Note that the pion-nucleon system can in general have total isospin $t = 1/2$ or $t = 3/2$. However, since the initial state is a pure $t = 3/2$ state, isospin conservation implies that only the $t = 3/2$ part of the pion-nucleon system contributes to the decay amplitude. This justifies projecting the final state onto the subspace with total isospin $t = 3/2$.

To express the isospin state of the Δ^+ in terms of the above two-particle states, we use the table of Clebsch–Gordan coefficients. The state $|\Delta^+\rangle$, with total isospin $t = 3/2$ and $t_z = 1/2$, can be written as a linear combination of pion-nucleon states with total isospin $t = 3/2$ and the same projection,

$$|\Delta^+\rangle = \sqrt{\tfrac{1}{3}}\,|\pi^0 p\rangle + \sqrt{\tfrac{2}{3}}\,|\pi^+ n\rangle\,.$$

The probabilities of decay (decay widths) are proportional to the squared modulus of the amplitudes,

$$\frac{\Gamma(\Delta^+ \to \pi^+ n)}{\Gamma(\Delta^+ \to \pi^0 p)} = \frac{2/3}{1/3} = 2\,.$$

This result is consistent with experimental data, provided that the kinematic (phase space) factors for both channels are approximately equal.

The same reasoning applies to the Δ^0, which has $t_z = -1/2$. The two allowed decay channels in this case are $\pi^0 n$ and $\pi^- p$, and the decomposition reads

$$|\Delta^0\rangle = \sqrt{\tfrac{1}{3}}\,|\pi^0 n\rangle + \sqrt{\tfrac{2}{3}}\,|\pi^- p\rangle\,.$$

Again, the predicted ratio of decay widths is 2, which matches what is observed experimentally.

The development of the quark model gave isospin a more fundamental origin. The proton and neutron are made of $|p\rangle = |uud\rangle$ and $|n\rangle = |udd\rangle$, with the up and down quarks forming themselves an isospin doublet

$$|u\rangle = |t = 1/2, t_z = 1/2\rangle \ , \qquad |d\rangle = |t = 1/2, t_z = -1/2\rangle \ . \tag{6.112}$$

Isospin symmetry thus arises from the approximate equality of $m_u \approx m_d$ and the flavor-blind nature of QCD interactions. It is embedded in the broader SU(3) flavor symmetry, which also includes the strange quark, but remains particularly accurate and useful due to the small mass difference between the up and down quarks. Small mass splittings like $m_n - m_p \approx 1.29\,\text{MeV}$ and $m_{\pi^\pm} - m_{\pi^0} \approx 4.6\,\text{MeV}$ reflect the breaking of isospin symmetry due to electromagnetic interactions and quark mass differences. Nevertheless, many features of nuclear and particle physics can still be understood remarkably well using isospin conservation and its associated selection rules. The excellent agreement of the predicted and observed branching ratios in Δ decays serves as a concrete example of the utility of isospin symmetry in hadronic physics.

6.8 Some worked-out exercises

L_z measurements

Consider a system described by the wave function $\varphi(\vec{r}) = f(r)(x + iy + z)$ with $f(r)$ a radial function. What are the possible outcomes if we perform a measurement of the operator L_z,? And the corresponding probabilities?

Taking into account that

$$Y_{00} = \sqrt{\frac{1}{4\pi}}\,, \qquad Y_{1,\pm 1} = \mp\sqrt{\frac{3}{8\pi}}\sin\theta e^{\pm i\phi}\,, \qquad Y_{10} = \sqrt{\frac{3}{4\pi}}\cos\theta\,,$$

the given wave function can be written as

$$\varphi(r) = rf(r)\sqrt{\frac{4\pi}{3}}\left[-\sqrt{2}Y_{11} + Y_{10}\right]\,.$$

The measurement of L_z can yield $m = 1$ or $m = 0$, with probabilities

$$P(m=0) = \frac{|\langle 10|\varphi\rangle|^2}{\langle\varphi|\varphi\rangle} = \frac{\int |Y_{1,0}|^2\, d\Omega}{\int \left|Y_{10} - \sqrt{2}Y_{11}\right|^2 d\Omega} = \frac{1}{\int \left[|Y_{10}|^2 + 2\,|Y_{11}|^2\right] d\Omega} = \frac{1}{3}\,,$$

$$P(m=1) = \frac{|\langle 11|\varphi\rangle|^2}{\langle\varphi|\varphi\rangle} = \frac{2\int |Y_{1,1}|^2\, d\Omega}{\int \left|Y_{10} - \sqrt{2}Y_{11}\right|^2 d\Omega} = \frac{2}{\int \left[|Y_{10}|^2 + 2\,|Y_{11}|^2\right] d\Omega} = \frac{2}{3}\,.$$

Spin-dependent interactions of two identical particles in a 1D harmonic trap

Consider two indistinguishable particles of spin 1/2 and mass m constrained to move along a line and interacting through the Hamiltonian

$$H = \frac{1}{2m}\left(p_1^2 + p_2^2\right) + \frac{1}{2}m\omega^2\left(x_1^2 + x_2^2\right) - \frac{\omega}{2\hbar}\left(S^2 + \hbar S_y\right)\,,$$

with $\vec{S} = \vec{S}_1 + \vec{S}_2$ the total spin of the system and $\vec{S}_1, \vec{S}_2$ the spin operators of the two particles.
Calculate the energy and degeneracy of the two lowest energy levels. Express the result in terms of the eigenfunctions of the one-dimensional harmonic oscillator and the eigenfunctions $|s\, s_z\rangle$ of S^2 and S_z.

a) Spatial part: we are dealing with two one-dimensional harmonic oscillators. The corresponding eigenvalues and eigenfunctions for the first energy levels are then given by

$$E_{n_1 n_2} = \hbar\omega\left(n_1 + n_2 + 1\right)\,,$$

$$n_1=0,\, n_2=0\,, \quad E=\hbar\omega\,, \quad \text{non degenerate} \qquad |0\rangle_1|0\rangle_2\,,$$

$$n_1=1,\, n_2=0 \text{ or } n_1=0,\, n_2=1 \quad E=2\hbar\omega\,, \quad 2 \text{ times degenerate} \quad \begin{cases} |\varphi_s\rangle = \frac{1}{\sqrt{2}}\left(|1\rangle_1|0\rangle_2 + |0\rangle_1|1\rangle_2\right)\,, \\ |\varphi_a\rangle = \frac{1}{\sqrt{2}}\left(|1\rangle_1|0\rangle_2 - |0\rangle_1|1\rangle_2\right)\,, \end{cases}$$

$$\text{similar for } n_1,n_2,n_3=0,1 \quad E=3\hbar\omega\,, \quad 3 \text{ times degenerate} \quad \begin{cases} \frac{1}{\sqrt{2}}\left(|2\rangle_1|0\rangle_2 + |0\rangle_1|2\rangle_2\right)\,, \\ \frac{1}{\sqrt{2}}\left(|2\rangle_1|0\rangle_2 - |0\rangle_1|2\rangle_2\right)\,, \\ |1\rangle_1|1\rangle_2\,. \end{cases}$$

Spin part: The total spin can only be $s=0,1$. Denoting the spin part states as $|s\,s_y\rangle_y$, we have

$$E_s = -\frac{\hbar\omega}{2}(s(s+1)+s_y)$$

$$s=0\,, \quad E_s=0\,, \quad |00\rangle_y$$

$$s=1 \quad E_s=-\tfrac{\hbar\omega}{2}\,(2+s_y) = \begin{cases} -\frac{3}{2}\hbar\omega & s_y=1\,, \quad |11\rangle_y \\ -\hbar\omega & s_y=0\,, \quad |10\rangle_y \\ -\frac{1}{2}\hbar\omega & s_y=-1 \quad |1-1\rangle_y \end{cases}$$

According to the Pauli exclusion the total wave function of the system must be totally antisymmetric. Therefore, one must combine even spatial parts with (odd) spin states with $s=0$ and odd spatial parts with (even) spin states with $s=1$. Denoting the spin parts as $|s\,s_y\rangle$, the two lowest states of H are therefore

$$E=\frac{\hbar\omega}{2}\,, \qquad |\varphi_a\rangle|11\rangle_y \quad \text{non degenerate}$$

$$E=\hbar\omega \quad \begin{cases} |0\rangle_1|0\rangle_2|00\rangle_y \\ |\varphi_a\rangle|10\rangle_y \end{cases} \quad 2 \text{ times degenerate}$$

The next state has $E=\frac{3}{2}\hbar\omega$.
As requested, let us now express $|11\rangle_y, |10\rangle_y, |00\rangle_y$ in the basis $|s\,s_Z\rangle_z$. For spin 0, we have

$$|00\rangle_y = |00\rangle_z\,.$$

To calculate the expression for the other two states, we compute the matrix S_y for $s=1$ in the basis $|s\,s_z\rangle$. Using the definition of ladder operators, we have

$$\begin{array}{l} S_+ = S_x + iS_y, \\ S_- = S_x - iS_y, \end{array} \quad \longrightarrow \quad S_y = \frac{1}{2i}\left(S_+ - S_-\right)\,.$$

Taking into account the correspondence $|11\rangle_z=(1,0,0), |10\rangle_z=(0,1,0), |1-1\rangle_z=(0,0,1)]$, this gives

$$S_+ = \hbar\begin{pmatrix} 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2} \\ 0 & 0 & 0 \end{pmatrix}\,, \qquad S_- = \hbar\begin{pmatrix} 0 & 0 & 0 \\ \sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{pmatrix}\,, \qquad S_y = \frac{\hbar}{2i}\begin{pmatrix} 0 & \sqrt{2} & 0 \\ -\sqrt{2} & 0 & \sqrt{2} \\ 0 & -\sqrt{2} & 0 \end{pmatrix}\,.$$

Let's calculate $|11\rangle_y$:

$$\frac{\hbar}{2i}\begin{pmatrix} 0 & \sqrt{2} & 0 \\ -\sqrt{2} & 0 & \sqrt{2} \\ 0 & -\sqrt{2} & 0 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \hbar\begin{pmatrix} a \\ b \\ c \end{pmatrix} \longrightarrow \begin{cases} \frac{1}{2i}\sqrt{2}b = a, \\ \frac{1}{2i}(-\sqrt{2}a+\sqrt{2}c) = b, \\ \frac{1}{2i}(-\sqrt{2}b) = c, \end{cases} \longrightarrow \begin{cases} a = -\frac{i}{\sqrt{2}}b, \\ c = \frac{i}{\sqrt{2}}b. \end{cases}$$

$$|11\rangle_y = \left(-\frac{ib}{\sqrt{2}}, b, \frac{ib}{\sqrt{2}}\right) \quad \longrightarrow \quad |v|^2 = \frac{|b|^2}{2} + |b|^2 + \frac{|b|^2}{2} = 2|b|^2 \quad \longrightarrow \quad |b| = \frac{1}{\sqrt{2}} \quad b = \frac{i}{\sqrt{2}}\,,$$

with the phase chosen arbitrarily. Therefore,

$$|11\rangle_y = \frac{1}{2}|11\rangle_z + \frac{i}{\sqrt{2}}|10\rangle_z - \frac{1}{2}|1-1\rangle_z$$

Let us finally compute $|10\rangle_y$

$$\frac{\hbar}{2i}\begin{pmatrix} 0 & \sqrt{2} & 0 \\ -\sqrt{2} & 0 & \sqrt{2} \\ 0 & -\sqrt{2} & 0 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = 0 \longrightarrow \begin{cases} \frac{1}{2i}\sqrt{2}b = 0 \\ \frac{1}{2i}(-\sqrt{2}a + \sqrt{2}c) = 0 \\ \frac{1}{2i}\sqrt{2}(-b) = 0 \end{cases} \longrightarrow \quad b = 0\,, a = c\,.$$

$$|10\rangle_y = (a, 0, a) \quad \longrightarrow \quad |v|^2 = |a|^2 + |a|^2 = 2|a|^2 \quad \longrightarrow \quad a = \frac{1}{\sqrt{2}}\,.$$

Therefore,

$$|10\rangle_y = \frac{1}{\sqrt{2}}\left(|11\rangle_z + |1-1\rangle_z\right)\,.$$

Total spin of a diatomic molecule: symmetry constraints

Consider a diatomic molecule formed by two spin-1 atoms in a non-relativistic system, such as two deuterium atoms forming a deuterium molecule. The two deuterons are in an s-wave state relative to each other (no orbital angular momentum).

a) Determine the possible values of the total spin quantum number j for the deuterium molecule if the deuterons are treated as distinguishable particles. Write down the corresponding spin-wave functions in terms of the individual spin states of the two deuterons.

b) Reanalyze the situation when the deuterons are treated as identical particles.

a) When the deuterons are distinguishable, the possible total spin states j are 2, 1, and 0. To determine the states of total angular momentum for a system of two spin-1 particles, we refer to the table of Clebsch-Gordan coefficients. Accounting for the square root and sign conventions in the PDG table, we have

$$|2,2\rangle = |1,1\rangle\,,$$

$$|2,1\rangle = \frac{1}{\sqrt{2}}\left(|1,0\rangle + |0,1\rangle\right), \qquad |1,1\rangle = \frac{1}{\sqrt{2}}\left(|1,0\rangle - |0,1\rangle\right)\,,$$

$$|2,0\rangle = \frac{1}{\sqrt{6}}\left(|1,-1\rangle + |-1,1\rangle + 2|0,0\rangle\right)\,, \qquad |1,0\rangle = \frac{1}{\sqrt{2}}\left(|1,-1\rangle - |-1,1\rangle\right)\,,$$

$$|2,-1\rangle = \frac{1}{\sqrt{2}}\left(|0,-1\rangle + |-1,0\rangle\right)\,, \qquad |1,-1\rangle = \frac{1}{\sqrt{2}}\left(|0,-1\rangle - |-1,0\rangle\right)\,,$$

$$|2,-2\rangle = |-1,-1\rangle. \qquad |0,0\rangle = \frac{1}{\sqrt{3}}\left(|1,-1\rangle - |0,0\rangle + |-1,1\rangle\right)\,.$$

b) According to the spin-statistics theorem, since the two identical spin-1 deuterons are bosons, their total wave function must be symmetric under the exchange of particles. This symmetry requirement implies that any state that is antisymmetric under particle exchange cannot exist in the physical system.
Since all states with $j = 1$ in a) are antisymmetric under particle exchange (for instance, we have $\frac{1}{\sqrt{2}}(|1,0\rangle - |0,1\rangle) = -\frac{1}{\sqrt{2}}(|0,1\rangle - |1,0\rangle)$), they do not survive when the particles are identical. In other words, for two identical spin-1 particles, the possible values of the total angular momentum j are restricted to 2 and 0.

Orbital and spin dynamics of a particle on a sphere

Consider a particle of mass m and spin 1/2 restricted to move on the surface of a unit radius sphere and in a state specified by the normalized *spinor*

$$|\varphi_0\rangle = \frac{1}{\sqrt{40\pi}} \begin{pmatrix} 3\sin\theta\, e^{-i\varphi} \\ 2\, e^{i\alpha} \end{pmatrix} ,$$

with α a constant parameter.

a) What values of L^2 can be obtained if we perform a measurement on $|\varphi_0\rangle$? What are the corresponding probabilities?

b) If a measurement of the total angular momentum J^2 at time $t = 0$ yields the value $3/4\hbar^2$, what is the normalized state $|\varphi_1\rangle$ of the system immediately after the measurement?

c) Determine the evolution of $|\varphi_1\rangle$ under a Hamiltonian

$$H = \frac{\omega}{2\hbar} L^2 + \omega\,(L_z + 2S_z) ,$$

with ω a constant. If a measurement of the third spin component is performed at time $t > 0$, which is the probability of finding the particle with spin up?

a) In the notation $|\,l\, l_z\, s_z\rangle$ and $Y_{l,m}$, we have

$$|\varphi_0\rangle = \frac{1}{\sqrt{40\pi}} \left(\sqrt{24\pi} Y_{1,-1} |1/2\rangle + e^{i\alpha} \sqrt{16\pi} Y_{0,0} |-1/2\rangle \right) = \sqrt{\frac{3}{5}} |\,1\; -1\, 1/2\rangle + e^{i\alpha} \sqrt{\frac{2}{5}} |\,0\, 0\; -1/2\rangle .$$

Therefore, the possible outcomes are states with $l = 1, 0$. The corresponding values of L^2 are, respectively, $l(l+1) = 2\hbar^2$ and 0, with probabilities

$$\text{Prob}\left(L^2 = 2\hbar^2\right) = \frac{3}{5} , \qquad \text{Prob}\left(L^2 = 0\right) = \frac{2}{5} .$$

b) In the coupled basis $|\,j\, m\,\rangle_{\rm C}$, the state $|\varphi_0\rangle$ takes the form (cf. CG table for the first term in $|\varphi_0\rangle$, no tables needed for the second!)

$$|\varphi_0\rangle = \sqrt{\frac{3}{5}} \left[\sqrt{\frac{1}{3}} \left| 1\, \frac{1}{2}\, \frac{3}{2}\; -\frac{1}{2} \right\rangle_{\rm C} - \sqrt{\frac{2}{3}} \left| 1\, \frac{1}{2}\, \frac{1}{2}\; -\frac{1}{2} \right\rangle_{\rm C} \right] + e^{i\alpha} \sqrt{\frac{2}{5}} \left| 0\, \frac{1}{2}\, \frac{1}{2}\; -\frac{1}{2} \right\rangle_{\rm C} .$$

Since $j(j+1) = 3/4$, the measurement select the part of the state with $j = 1/2$, namely

$$-\sqrt{\frac{2}{5}} \left| 1\, \frac{1}{2}\, \frac{1}{2}\; -\frac{1}{2} \right\rangle_{\rm C} + e^{i\alpha} \sqrt{\frac{2}{5}} \left| 0\, \frac{1}{2}\, \frac{1}{2}\; -\frac{1}{2} \right\rangle_{\rm C} .$$

Normalizing this quantity, we obtain

$$|\varphi_1\rangle = \frac{1}{\sqrt{2}} \left| 1\, \frac{1}{2}\, \frac{1}{2}\; -\frac{1}{2} \right\rangle_{\rm C} - e^{i\alpha} \frac{1}{\sqrt{2}} \left| 0\, \frac{1}{2}\, \frac{1}{2}\; -\frac{1}{2} \right\rangle_{\rm C} ,$$

c) Going back to the uncoupled basis $|\,l\,l_z\,s_z\rangle$, the state $|\varphi_1\rangle$ becomes

$$\begin{aligned}|\varphi_1\rangle =&\frac{1}{\sqrt{2}}\left[\sqrt{\frac{1}{3}}\left|1\,0\,-\frac{1}{2}\right\rangle-\sqrt{\frac{2}{3}}\left|1\,-1\,\frac{1}{2}\right\rangle\right]-\frac{e^{i\alpha}}{\sqrt{2}}\left|0\,0\,-\frac{1}{2}\right\rangle\\ =&\frac{1}{\sqrt{6}}\left|1\,0\,-\frac{1}{2}\right\rangle-\frac{1}{\sqrt{3}}\left|1\,-1\,\frac{1}{2}\right\rangle-\frac{e^{i\alpha}}{\sqrt{2}}\left|0\,0\,-\frac{1}{2}\right\rangle\,.\end{aligned}$$

Taking into account that

$$H|\,l\,l_z,s_z\rangle=\left[\frac{\omega}{2\hbar}L^2+\omega\,(L_z+2S_z)\right]|\,l\,l_z,s_z\rangle=\hbar\omega\left[\frac{l(l+1)}{2}+l_z+2s_z\right]|\,l\,l_z,s_z\rangle\,,$$

the energies corresponding to each of the three states in the right hand side of this expression are given respectively by

$$E=\hbar\omega(1+0-1)=0\,,\qquad E=\hbar\omega(1-1+1)=\hbar\omega\,,\qquad E=\hbar\omega(0+0-1)=-\hbar\omega\,.$$

Therefore, we have

$$|\varphi_1(t)\rangle=\frac{1}{\sqrt{6}}\left|1\,0\,-\frac{1}{2}\right\rangle-\frac{e^{-i\omega t}}{\sqrt{3}}\left|1\,-\frac{1}{2}\,\frac{1}{2}\right\rangle-\frac{e^{i\alpha+i\omega t}}{\sqrt{2}}\left|0\,0\,-\frac{1}{2}\right\rangle=|\varphi_A\rangle\left|\frac{1}{2}\right\rangle+|\varphi_B\rangle\left|-\frac{1}{2}\right\rangle\,,$$

with

$$|\varphi_A\rangle\equiv-\frac{e^{-i\omega t}}{\sqrt{3}}\left|1-\frac{1}{2}\right\rangle\,,\qquad\qquad|\varphi_B\rangle=\frac{1}{\sqrt{6}}|10\rangle-\frac{1}{\sqrt{2}}e^{i\alpha+i\omega t}|00\rangle\,.$$

The probability of finding the particle with spin up at time t is $1/3$.

A hydrogen atom in a constant magnetic field

Consider the reduced Hamiltonian of a hydrogen atom in the presence of a constant magnetic field along the z-direction,

$$H=-\frac{\hbar^2}{2m}\nabla^2+\frac{e^2}{r}+\omega S_z\,,$$

with m the reduced mass of the system, $\omega=eB/(2m)$ the Larmor frequency, B the magnetic field magnitude and S_z the third component of the electron's spin. The system is further characterized by the orbital angular momentum $\vec{L}$ and the total angular momentum $\vec{J}=\vec{L}+\vec{S}$. The initial state of the system has the following properties:

- Its radial quantum number is $n=3$.
- It is an eigenstate of J^2 with eigenvalue $15\hbar^2/4$.
- It is an eigenstate of J_z with eigenvalue $3\hbar/2$.
- A measurement of L^2 yields $6\hbar^2$ with probability $1/2$.

a) Which other values can be obtained when measuring L^2 in this state? With what probabilities?
b) Write down the vector representing the initial state of the system in the coupled and uncoupled bases.

c) Determine the possible results of an energy measurement and the associated probabilities.
d) Calculate the probability of obtaining the value $35\hbar^2/4$ when measuring J^2 after a time t.

a) Since $j(j+1)\hbar^2 = 15\hbar^2/4$ and $m_j\hbar = 3\hbar/2$, the quantum numbers of the initial state are $j = 3/2$ and $m_j = 3/2$. The possible values of l are integers in the range $|j-s| \leq l \leq j+s$ with $s = 1/2$ the spin of the electron, i.e.,

$$l = \frac{3}{2} - \frac{1}{2} = 1 \qquad \text{and} \qquad l = \frac{3}{2} + \frac{1}{2} = 2\,.$$

Since the probability of the state having $l = 2$ (corresponding to $l(l+1)\hbar^2 = 6\hbar^2$) is 50 %, the probability of it having the other possible value $l = 1$ (corresponding to the eigenvalue $2\hbar^2$) is the remaining 50 %.

b) Taking into account the previous results, the state vector in the coupled basis can be written as

$$|\varphi\rangle = \sqrt{\frac{1}{2}}\left|3,2,\frac{1}{2},\frac{3}{2},\frac{3}{2}\right\rangle_C + e^{i\alpha}\sqrt{\frac{1}{2}}\left|3,1,\frac{1}{2},\frac{3}{2},\frac{3}{2}\right\rangle_C\,,$$

with $e^{i\alpha}$ an arbitrary phase. Decomposing each term in this expression into the uncoupled basis using the table of Clebsch-Gordan coefficients,

$$\left|3,2,\frac{1}{2},\frac{3}{2},\frac{3}{2}\right\rangle_C = \sqrt{\frac{4}{5}}\left|3,2,2,\frac{1}{2},-\frac{1}{2}\right\rangle - \sqrt{\frac{1}{5}}\left|3,2,1,\frac{1}{2},\frac{1}{2}\right\rangle\,,$$
$$\left|3,1,\frac{1}{2},\frac{3}{2},\frac{3}{2}\right\rangle_C = \left|3,1,1,\frac{1}{2},\frac{1}{2}\right\rangle\,,$$

we get

$$|\varphi\rangle = \sqrt{\frac{2}{5}}\left|3,2,2,\frac{1}{2},-\frac{1}{2}\right\rangle - \sqrt{\frac{1}{10}}\left|3,2,1,\frac{1}{2},\frac{1}{2}\right\rangle + e^{i\alpha}\sqrt{\frac{1}{2}}\left|3,1,1,\frac{1}{2},\frac{1}{2}\right\rangle\,.$$

c) The spectrum of the system can be easily computed in the uncoupled basis, where the Hamiltonian is diagonal,

$$E_{n,m_s} = -\frac{E_0}{n^2} + \hbar\omega m_s\,.$$

In the state $|\varphi\rangle$, the only possible energy outcomes are

$$E_{3,-1/2} = -\frac{E_0}{9} - \frac{\hbar\omega}{2}\,, \qquad E_{3,1/2} = -\frac{E_0}{9} + \frac{\hbar\omega}{2}\,,$$

with respective probabilities

$$P_{3,-1/2} = \left|\left\langle 3,2,2,\frac{1}{2},-\frac{1}{2}\middle|\varphi\right\rangle\right|^2 = \frac{2}{5}\,, \qquad P_{3,1/2} = \left|\left\langle 3,2,1,\frac{1}{2},\frac{1}{2}\middle|\varphi\right\rangle\right|^2 + \left|\left\langle 3,1,1,\frac{1}{2},\frac{1}{2}\middle|\varphi\right\rangle\right|^2 = \frac{1}{10}+\frac{1}{2} = \frac{3}{5}\,.$$

d) In the uncoupled basis where the Hamiltonian is diagonal, the time evolution of the system's state is given by

$$|\varphi,t\rangle = e^{-\frac{i}{\hbar}E_{3,-1/2}t}\sqrt{\frac{2}{5}}\left|3,2,2,\frac{1}{2},-\frac{1}{2}\right\rangle - e^{-\frac{i}{\hbar}E_{3,1/2}t}\sqrt{\frac{1}{10}}\left|3,2,1,\frac{1}{2},\frac{1}{2}\right\rangle + e^{-\frac{i}{\hbar}E_{3,1/2}t}e^{i\alpha}\sqrt{\frac{1}{2}}\left|3,1,1,\frac{1}{2},\frac{1}{2}\right\rangle\,.$$

To determine the probability of obtaining a certain value when measuring J^2, the vectors of this linear combination must be expressed in terms of the coupled basis ones. Using once again the table of Clebsch-Gordan coefficients,

$$\left|3,2,2,\frac{1}{2},-\frac{1}{2}\right\rangle = \sqrt{\frac{1}{5}}\left|3,2,\frac{1}{2},\frac{5}{2},\frac{3}{2}\right\rangle_C + \sqrt{\frac{4}{5}}\left|3,2,\frac{1}{2},\frac{3}{2},\frac{3}{2}\right\rangle_C ,$$
$$\left|3,2,1,\frac{1}{2},\frac{1}{2}\right\rangle = \sqrt{\frac{4}{5}}\left|3,2,\frac{1}{2},\frac{5}{2},\frac{3}{2}\right\rangle_C - \sqrt{\frac{1}{5}}\left|3,2,\frac{1}{2},\frac{3}{2},\frac{3}{2}\right\rangle_C ,$$
$$\left|3,1,1,\frac{1}{2},\frac{1}{2}\right\rangle = \left|3,1,\frac{1}{2},\frac{3}{2},\frac{3}{2}\right\rangle_C ,$$

we get

$$|\varphi,t\rangle = \sqrt{\frac{2}{25}}\left(e^{-\frac{i}{\hbar}E_{3,-1/2}t} - e^{-\frac{i}{\hbar}E_{3,1/2}t}\right)\left|3,2,\frac{1}{2},\frac{5}{2},\frac{3}{2}\right\rangle + \left(\sqrt{\frac{8}{25}}e^{-\frac{i}{\hbar}E_{3-\frac{1}{2}}t} - \sqrt{\frac{1}{50}}e^{-\frac{i}{\hbar}E_{3,1/2}t}\right)\left|3,2,\frac{1}{2},\frac{3}{2},\frac{3}{2}\right\rangle_C$$
$$+ e^{-\frac{i}{\hbar}E_{3,1/2}t}e^{i\alpha}\sqrt{\frac{1}{2}}\left|3,1,\frac{1}{2},\frac{3}{2},\frac{3}{2}\right\rangle_C .$$

The probability of obtaining the value $35\hbar^2/4$ when measuring J^2 (corresponding to the quantum number $j=\frac{5}{2}$), is then

$$P_{j=\frac{5}{2}} = \left|\left\langle 3,2,\frac{1}{2},\frac{5}{2},\frac{3}{2}\middle|\varphi\right\rangle\right|^2 = \frac{2}{25}\left|e^{-\frac{i}{\hbar}E_{3,-1/2}t} - e^{-\frac{i}{\hbar}E_{3,1/2}t}\right|^2 = \frac{2}{25}\left|e^{\frac{i}{\hbar}\frac{E_0}{9}t}\right|^2\left|e^{i\omega t/2} - e^{-i\omega t/2}\right|^2 = \frac{8}{25}\sin^2\frac{\omega t}{2} .$$

Two fermions in a magnetic field

Two particles with spins $s_1 = 3/2$ and $s_2 = 1/2$ are placed in a magnetic field $\vec{B} = (0,0,B)$ along the z-axis. The Hamiltonian of the system is given by

$$H = -\left(\gamma_1\vec{S}_1 + \gamma_2\vec{S}_2\right)\cdot\vec{B} ,$$

with $\gamma_1 \neq \gamma_2$ the gyromagnetic ratios of the particles. At time $t=0$, we measure the squared magnitude of the total spin $|\vec{S}_1+\vec{S}_2|^2$ and its third component $S_z = S_{1,z}+S_{2,z}$, obtaining the values $6\hbar^2$ and $-\hbar$, respectively.

a) Determine the state of the system $|\varphi,0\rangle$ immediately after the measurement.

b) Determine the state of the system $|\varphi,t\rangle$ at time t.

c) If $\gamma_2 = 2\gamma_1$ and we measure again the squared magnitude of the total spin at a time $t_{\rm m} = \pi/(B\gamma_1)$, what values can we obtain, and with what probability?

a) The value $6\hbar^2 = \hbar^2 j(j+1)$ implies a quantum number $j=2$ for the total spin, while the value $-\hbar = m\hbar$ implies $m=-1$ for the quantum number corresponding to the third component of the total spin. Therefore, in the coupled basis $|j,m\rangle_C$, we have

$$|\varphi,0\rangle = |2,-1\rangle_C .$$

b) The Hamiltonian $H = -(\gamma_1 S_{1,z} + \gamma_2 S_{2,z})\, B$ is diagonal in the uncoupled basis $|s_1, m_{s,1}; s_2, m_{s_2}\rangle$. Therefore, in order to determine the evolution of the system, we must express the initial state in this basis. Using the table of Clebsch-Gordan coefficients, we get

$$|\varphi, 0\rangle = \sqrt{\frac{3}{4}}\,|3/2, -1/2; 1/2, -1/2\rangle + \sqrt{\frac{1}{4}}\,|3/2, -3/2; 1/2, 1/2\rangle\,.$$

Taking now into account that the energy of the state $|s_1, m_{s,1}; s_2, m_{s,2}\rangle$ is $-B\hbar\,[\gamma_1 m_{s,1} + \gamma_2 m_{s,2}]$, we get

$$|\varphi, t\rangle = \sqrt{\frac{3}{4}}e^{-iB(\gamma_1+\gamma_2)t/2}\,|3/2, -1/2; 1/2, -1/2\rangle + \sqrt{\frac{1}{4}}e^{-iB(3\gamma_1-\gamma_2)t/2}\,|3/2, -3/2; 1/2, 1/2\rangle\ .$$

c) To calculate the possible values of the total spin magnitude, we need to return to the coupled basis. Using the Clebsch-Gordan coefficients and setting $\gamma_2 = 2\gamma_1$ as indicated, we get

$$\begin{aligned}|\varphi, t\rangle &= \sqrt{\frac{3}{4}}e^{-3iB\gamma_1 t/2}\left[\sqrt{\frac{3}{4}}|2, -1\rangle_C + \sqrt{\frac{1}{4}}|1, -1\rangle_C\right] + \sqrt{\frac{1}{4}}e^{-iB\gamma_1 t/2}\left[\sqrt{\frac{1}{4}}|2, -1\rangle_C - \sqrt{\frac{3}{4}}|1, -1\rangle_C\right] \\ &= \frac{3e^{-3iB\gamma_1 t/2} + e^{-iB\gamma_1 t/2}}{4}|2, -1\rangle_C + \frac{\sqrt{3}\left(e^{-3iB\gamma_1 t/2} - e^{-iB\gamma_1 t/2}\right)}{4}|1, -1\rangle_C\,.\end{aligned}$$

The two possible values for the square of the total spin magnitude are $6\hbar^2$, corresponding to the quantum number $j = 2$, and $2\hbar^2$, corresponding to $j = 1$. The corresponding probabilities at time $t_{\rm m} = \pi/(B\gamma_1)$ are given by

$$P_{j=2}(t_{\rm m}) = \frac{1}{16}\left|3e^{-3i\pi/2} + e^{-i\pi/2}\right|^2 = \frac{1}{16}|3i - i|^2 = \frac{1}{4}\,,$$

$$P_{j=1}(t_{\rm m}) = \frac{3}{16}\left|e^{-3i\pi/2} - e^{-i\pi/2}\right|^2 = \frac{3}{16}|i + i|^2 = \frac{3}{4}.$$

A spin-orbit interaction

The Hamiltonian of a particle of spin 3/2 is given by $H = \frac{\omega}{\hbar}\vec{L}\cdot\vec{S}$, with $\vec{L}$ is the angular momentum, $\vec{S}$ the spin, and ω a constant with units of angular frequency.

a) Determine the spectrum of the Hamiltonian.

b) If the initial state of the particle is $|\varphi, 0\rangle = |l = 2, m_l = 2; s = 3/2, m_s = 1/2\rangle$ determine the state at time t.

c) If we measure the third component of the spin at time $t_{\rm m} = 2\pi/\omega$, what values can we obtain, and with what probability?

a) The set of operators $\{L^2, S^2, L_z, S_z\}$ do not form CSCO since neither L_z nor S_z commute with H. In order to diagonalize the Hamiltonian we consider the set of commuting observables $\{L^2, S^2, J^2, J_z\}$ with $\vec{J} = \vec{L} + \vec{S}$ the total angular momentum and J_z its third component. In terms of these quantities, we have

$$H = \frac{\omega}{2\hbar}\left[J^2 - L^2 - S^2\right]\,.$$

In the coupled basis common to all these operators $|j,m\rangle_C \equiv |l.s;j,m\rangle$, the Hamiltonian is diagonal and the associated energy spectrum can be trivially computed. Taking into account that the eigenvalues of J^2 are $\hbar^2 j(j+1)$ with $j = |l-\frac{3}{2}|,\ldots,l+\frac{3}{2}$, we obtain

$$E_{l,j} = \frac{\hbar\omega}{2}\left[j(j+1) - l(l+1) - \frac{15}{4}\right],$$

with $l = 0,1,\ldots$ and $j = l-\frac{3}{2}, l-\frac{1}{2}, l+\frac{1}{2}, l+\frac{3}{2}$ if $l \geq 2$. If $l = 0$, then $j = \frac{3}{2}$ and if $l = 1$, $j = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}$.

b) Using the Clebsch-Gordan coefficients, we can write the initial state in the coupled basis $|j,m\rangle_C$,

$$|\varphi,0\rangle = \sqrt{\frac{3}{7}}\left|\frac{7}{2},\frac{5}{2}\right\rangle_C + \sqrt{\frac{4}{7}}\left|\frac{5}{2},\frac{5}{2}\right\rangle_C .$$

The energy of the state $\left|\frac{7}{2},\frac{5}{2}\right\rangle_C$ is $E_{2,\frac{7}{2}} = 3\hbar\omega$ and that of the state $\left|\frac{5}{2},\frac{5}{2}\right\rangle_C$ is $E_{2,\frac{5}{2}} = -\frac{\hbar\omega}{2}$. Therefore, we have

$$|\varphi,t\rangle = \sqrt{\frac{3}{7}}e^{-3i\omega t}\left|\frac{7}{2},\frac{5}{2}\right\rangle_C + \sqrt{\frac{4}{7}}e^{i\omega t/2}\left|\frac{5}{2},\frac{5}{2}\right\rangle_C .$$

c) At $t_{\rm m} = \frac{2\pi}{\omega}$ the exponential in the previous expression equal 1 and -1, such that

$$|\varphi,t_{\rm m}\rangle = \sqrt{\frac{3}{7}}\left|\frac{7}{2},\frac{5}{2}\right\rangle_C - \sqrt{\frac{4}{7}}\left|\frac{5}{2},\frac{5}{2}\right\rangle_C .$$

In order to determine the values and probabilities of obtaining a given spin projection, we must return to the uncoupled basis $|l,m_l;s,m_s\rangle$. Using again the Clebsch-Gordan coefficients, we have

$$\left|\frac{7}{2},\frac{5}{2}\right\rangle_C = \sqrt{\frac{3}{7}}\,|2,2;3/2,1/2\rangle + \sqrt{\frac{4}{7}}\,|2,1;3/2,3/2\rangle\,,$$

$$\left|\frac{5}{2},\frac{5}{2}\right\rangle_C = \sqrt{\frac{4}{7}}\,|2,2;3/2,12\rangle - \sqrt{\frac{3}{7}}\,|2,1;3/2,32\rangle\ ,$$

and therefore

$$|\varphi,t_{\rm m}\rangle = -\frac{1}{7}\left|2,2;3/2,\frac{1}{2}\right\rangle + \frac{4\sqrt{3}}{7}\,|2,1;3/2,32\rangle\,.$$

Therefore, when measuring the third component of the spin, we can obtain $\hbar/2$ with probability $1/49$ and $3\hbar/2$ with probability $48/49$.

Three identical spin-1/2 particles in a box with spin-spin interactions

Three identical particles of mass m and spin $1/2$ are confined to a one-dimensional infinite square well of width L and subject to a spin-spin interaction

$$V = -\frac{\alpha}{\hbar^2}(\vec{S}_1\cdot\vec{S}_2 + \vec{S}_2\cdot\vec{S}_3 + \vec{S}_3\cdot\vec{S}_1)\,,$$

with $\alpha > 0$ a constant and $\vec{S}_i$ the spin operator of the i-th particle. Knowing that a measurement of the total spin along the z-axis yields the value $+3/2\hbar$, write down the possible eigenfunctions of the system and their corresponding eigenvalues. Which is the energy of the ground state in this specific spin configuration?

The presence of the spin-spin interaction precludes the use of the uncoupled basis since the different terms $S_i \cdot S_j$ do not commute with the z components of S_i. On the other hand, the set of operators $\left\{S^2, S_1^2, S_2^2, S_3^2\right\}$ with $\vec{S} = \vec{S}_1 + \vec{S}_2 + \vec{S}_3$ the total spin of the system, commute with each other, thus providing an appropriate basis for diagonalizing V. Rewriting the interaction in terms of these operators,

$$\vec{S}_1 \cdot \vec{S}_2 + \vec{S}_2 \cdot \vec{S}_3 + \vec{S}_3 \cdot \vec{S}_1 = \frac{1}{2}\left(S^2 - S_1^2 - S_2^2 - S_3^2\right) \quad \longrightarrow \quad V = -\frac{\alpha}{2\hbar^2}\left(S^2 - S_1^2 - S_2^2 - S_3^2\right) ,$$

the energy eigenfunctions become the product of the eigenfunctions related to the spatial coordinates and those of the total spin.
Since, according to the problem, the total spin along the z-axis is $s_z = +3/2\hbar$, we necessarily have $s = 3/2$, which is the maximum value s can take. This implies that the spins of the three particles have equal z-component, making the spin eigenfunction completely symmetric. Requiring the total wave functions of this fermionic system to be completely antisymmetric, we have

$$\Phi_{n_1,n_2,n_3;\frac{3}{2},+\frac{3}{2}}(x_1,x_2,x_3) = \frac{1}{\sqrt{3!}} \det \begin{vmatrix} \varphi_{n_1}(x_1) & \varphi_{n_1}(x_2) & \varphi_{n_1}(x_3) \\ \varphi_{n_2}(x_1) & \varphi_{n_2}(x_2) & \varphi_{n_2}(x_3) \\ \varphi_{n_3}(x_1) & \varphi_{n_3}(x_2) & \varphi_{n_3}(x_3) \end{vmatrix} \otimes |3/2, 3/2\rangle_C .$$

The corresponding eigenvalues are given by

$$E = \frac{\pi^2\hbar^2}{2mL^2}\left(n_1^2 + n_2^2 + n_3^2\right) - \frac{3\alpha}{4} ,$$

with $n_1, n_2, n_3 = 1, 2, \ldots$. Since these three quantum numbers must be necessarily different for Φ to be antisymmetric, the energy is minimized for $n_1 = 1, n_2 = 2, n_3 = 3$,

$$E_{\text{ground}} = \frac{7\pi^2\hbar^2}{mL^2} - \frac{3\alpha}{4} .$$

6.9 Exercises

1. **Eigenvalues of orbital angular momentum**

   Consider the operators

   $$\begin{aligned} Q_1 &= \tfrac{1}{\sqrt{2}}\left(X + \tfrac{a^2}{\hbar}P_y\right), \qquad & P_1 &= \tfrac{1}{\sqrt{2}}\left(P_x - \tfrac{\hbar}{a^2}Y\right), \\ Q_2 &= \tfrac{1}{\sqrt{2}}\left(X - \tfrac{a^2}{\hbar}P_y\right), \qquad & P_2 &= \tfrac{1}{\sqrt{2}}\left(P_x + \tfrac{\hbar}{a^2}Y\right), \end{aligned}$$

   with a a parameter with the dimension of length.

   a) Show that $[Q_1, Q_2] = 0$, $[P_1, P_2] = 0$ and $[Q_1, P_1] = [Q_2, P_2] = i\hbar$.

   b) Express L_z in terms of the above operators, showing explicitly that it can be written as the difference between two independent harmonic oscillator Hamiltonians.

   c) By examining the energy eigenvalues of the harmonic oscillator, deduce that the L_z eigenvalues are necessarily integers.

2. **Angular momentum measurements for $j = 1$**

   Consider a system of total angular momentum $j = 1$.

   a) Find the matrices representing the operators J_x, J_y, J_z, $J_\pm$ and J^2.

   b) Which are the possible values that can be obtained when measuring J_x and J_y?

   c) Find the joint eigenstates of J^2 and J_z. Verify that they form an orthogonal basis.

   d) Calculate $\langle J_z\rangle$, $\langle J_z^2\rangle$ and ΔJ_z if the system is in a state with $j_x = -\hbar$. Repeat the exercise for $\langle J_y\rangle$, $\langle J_y^2\rangle$ and ΔJ_y.

   e) If the system is initially in the state $|\varphi\rangle = \frac{1}{\sqrt{14}}(-\sqrt{3}, 2\sqrt{2}, \sqrt{3})^t$, which values could one obtain when measuring J_x? With what probabilities?

3. **A hydrogen atom wave function**

   In the case of a hydrogen atom, it is known that: a) it is in a $l = 1$ state with $n = 2$; b) the state contains eigenstates of L_z with eigenvalues $+1$ and -1; c) the expectation value of L_z is zero; d) the probability of finding the electron in the first quadrant $\left(0 < \phi < \frac{\pi}{2}\right)$ is 25 %. Write down the corresponding wave function.

4. **L_z measurements**

   An electron is described by the wave function

   $$\varphi(r, \theta, \phi) = \frac{1}{\sqrt{4\pi}}\left(e^{i\phi}\sin\theta + \cos\theta\right) f(r), \qquad \int_0^\infty |f(r)|^2 r^2\, dr = 1\,.$$

   Is φ properly normalized? What are the possible outcomes of a measurement of the z-component of the angular momentum, L_z? What are the corresponding probabilities?

5. **Angular momentum in arbitrary directions**

   Consider a system in a state with angular momentum $\ell = 1$. After measuring the component of the angular momentum along a direction $\vec{n}$ that makes an angle θ with the z-axis and obtaining the result $+\hbar$, we measure the component of the angular momentum along the z-axis. What is the probability of finding the value $+\hbar$?

6. **A spin 1 particle in a magnetic field**

Consider a particle of spin $s = 1$ subject to a constant magnetic field along the x-direction. If the particle is initially in the state $|1, 1\rangle$, which is the probability of finding the particle in the state $|1, -1\rangle$? And in $|1, 0\rangle$? Ignore all spatial degrees of freedom.

7. **A rotor with angular momentum $j = 1$**

The Hamiltonian of a system with angular momentum $j = 1$ is given by

$$H = \frac{\omega_0}{\hbar}\left(J_x^2 + J_y^2\right) ,$$

with ω_0 a constant with dimensions of frequency.

a) Determine the matrix representation of H in the basis of eigenvectors of J_z. Find the possible energy eigenvalues and the stationary states of the system. Is the Hamiltonian a CSCO?

b) Calculate the probabilities of obtaining the different possible values of energy and the x-component of the angular momentum as a result of a measurement if the system is initially prepared in a state

$$|\varphi(0)\rangle = \frac{1}{2}|1\rangle + \frac{i}{\sqrt{2}}|0\rangle + \frac{1}{2}|-1\rangle,$$

c) Determine the same probabilities as in part b) at any time $t > 0$.

8. **A spin $s = 3/2$ particle**

Consider a particle of spin 3/2.

a) Find the matrices representing the operators S_x, S_y, S_z, S_x^2, S_y^2 and S_z^2 within the basis of S^2 and S_z.

b) Find the energy levels of this particle when is Hamiltonian is given by

$$H = \frac{\epsilon_0}{\hbar^2}(S_x^2 - S_y^2) - \frac{\epsilon_0}{\hbar}S_z\,,$$

with ϵ_0 a constant with energy dimensions. Are these level degenerate?

c) If the system was initially in an eigenstate $|\varphi_0\rangle = (1, 0, 0, 0)$, determine the state of the system at time t.

9. **Rotational States of a Symmetric Triatomic System**

Three identical spin-0 particles are placed at the vertices of an equilateral triangle in the $x - y$ plane, with the z-axis passing through the center of the triangle, perpendicular to its plane. The system is free to rotate about the z-axis. Using the symmetry of the configuration and the indistinguishability of the particles, determine the allowed values of the magnetic quantum number J_z corresponding to the rotational states of the system.

10. **Clebsch-Gordan coefficients: Spins 1 and 1/2**

Consider the coupling of two angular momenta $j_1 = 1/2$ and $j_2 = 1$. Write down all elements of the uncoupled and coupled bases. Verify that both bases have the same number of elements. Work out the corresponding Clebsch-Gordan coefficients, checking explicitly the result against the table of Clebsch-Gordan coefficients.

11. **Clebsch-Gordan coefficients: Spins 1 and 3/2**

Consider the coupling of two angular momenta $j_1 = 3/2$, $j_2 = 1$. Write down all elements of the uncoupled and coupled bases. Verify that both bases have the same number of elements. Work out the corresponding Clebsch-Gordan coefficients, checking explicitly the result against the table of Clebsch-Gordan coefficients.

12. **Clebsch-Gordan coefficients: Two spins 1**

Consider the coupling of two angular momenta $j_1 = 1$, $j_2 = 1$. Write down all elements of the uncoupled and coupled bases. Verify that both bases have the same number of elements. Work out the corresponding Clebsch-Gordan coefficients, checking explicitly the result against the table of Clebsch-Gordan coefficients.

13. **Pauli's Neutrino Hypothesis**

Consider the so-called beta decay, an observed process in which a neutron is transformed into a proton by emitting an electron or beta particle, $n \to p + e^-$, having all these objects spin 1/2. Prove that this reaction is forbidden by spin conservation, arguing also that the apparent puzzle cannot be solved by adding some orbital angular momentum to the final decay products. Demonstrate, however, that the addition of a third invisible spin-1/2 particle $\bar{\nu}_e$ to the final state, $n \to p + e^- + \bar{\nu}_e$, solves the problem. This is precisely what Pauli did, postulating the existence of what we call nowadays the neutrino.. Can you postulate other properties this particle should have?

14. **Decay modes of the ρ^0-Meson**

The neutral ρ^0-meson has spin 1 and a relatively short lifetime. It decays predominantly into two spinless pions of opposite charge $\rho^0 \to \pi^+ + \pi^-$.

a) Explain why the decay $\rho^0 \to \pi^0 + \pi^0$ is forbidden, even though it appears to conserve energy and charge.

b) The photon has spin 1 and odd parity. Could the decay $\rho^0 \to \pi^0 + \pi^0 + \gamma$ be allowed?

c) Given that the intrinsic parity of the pion is negative, determine the intrinsic parity of the ρ^0-meson, assuming parity is conserved in its dominant decay.

15. **Decay modes of a particle X**

A particle X is observed to decay through the channels $X \to \rho^+ + \pi^-$ and $X \to K + K$, where all final-state particles are spin-0 mesons, except the ρ^+, which has spin 1. Assume that total angular momentum and parity are conserved in both decays.

a) What is the minimum possible spin of the particle X consistent with these decay modes?

b) Given that the intrinsic parity of both π and K is negative, determine the intrinsic parity of X.

16. **Two non-interacting electrons in a Coulomb potential**

Consider a system of two electrons in the Coulomb potential produced by a charge Ze located at the origin (neglect the Coulomb repulsion between the electrons). The state of the system is such that:

i) Its energy is well-defined and equal to $E = E_1/2$, with E_1 the energy of the ground state of a hydrogenic atom.

ii) The squares of the orbital angular momentum of each electron, $|\vec{L}_1|^2$, $|\vec{L}_2|^2$, and the total $|\vec{L}_1 + \vec{L}_2|^2$ are well-defined, and they all have values of $2\hbar^2$.

17. **Two interacting spin-1/2 particles**

Consider the following Hamiltonian for a system of two spin-1/2 particles

$$H = \frac{\omega}{\hbar}\left[\vec{S}_1 \cdot \vec{S}_2 - \hbar\left(S_{1z} + \beta S_{2z}\right)\right] .$$

a) Determine the values of ω and β for which the Hamiltonian describes a system of two identical particles.

b) Assuming the found values for β and ω, determine the eigenvalues of H, their degeneracy and the eigenstates of the system. Write the eigenstates both in terms of the uncoupled and the coupled bases.

c) If a measurement of the third spin component of each particle at time $t = 0$ provides values $-\hbar/2$ and $\hbar/2$. Which is the state of the system at time t?

d) Which values can we obtain if we measure again S_{1z} and S_{2z} at time $t_m = \pi/\omega$? Which are the associated probabilities?

18. **Two interacting particles of spin 3/2 and 1**

Repeat the previous exercise for two particles of spin 3/2 and 1. Assume that the measurement of the third spin component of each particle at time $t = 0$ provides values $-\hbar/2$ and $\hbar$.

19. **Three distinct spin-1/2 particles with spin-spin interactions**

Three distinct spin-1/2 particles are subject to an interaction Hamiltonian

$$H = \frac{\epsilon_0}{\hbar^2}(\vec{S}_1 \cdot \vec{S}_2 + \vec{S}_2 \cdot \vec{S}_3 + \vec{S}_3 \cdot \vec{S}_1) ,$$

with $\epsilon_0 > 0$ a constant a constant and $\vec{S}_i$ the spin operator of the i-th particle. Determine the eigenvalues of H, their degeneracy, and the corresponding eigenstates.

20. **Symmetric and Antisymmetric Spin States of Two Identical Particles**

Show that, for a system of two identical particles each with spin ss, the ratio of the number of spin states that are symmetric under particle exchange to those that are antisymmetric is $(s + 1)/s$.

21. **Three identical spin-1/2 particles in a box with spin-spin interactions**

Three identical particles of mass m and spin 1/2 are confined to a one-dimensional infinite square well of width L and subject to a spin-spin interaction

$$V = -\frac{\alpha}{\hbar^2}(\vec{S}_1 \cdot \vec{S}_2 + \vec{S}_2 \cdot \vec{S}_3 + \vec{S}_3 \cdot \vec{S}_1) ,$$

with $\alpha > 0$ a constant and $\vec{S}_i$ the spin operator of the i-th particle. Knowing that a measurement of the total spin along the z-axis yields the value $+3/2\hbar$, write down the possible eigenfunctions of the system and their corresponding eigenvalues. Which is the energy of the ground state in this specific spin configuration?

CHAPTER 7

APPROXIMATE METHODS FOR STATIONARY SYSTEMS

Assume the cow is a sphere...

A random theorist, upon first approaching a problem

So far, we have been dealing with quantum versions of the spherical cow, i.e. dynamical problems such as the harmonic oscillator or the Coulomb potential that can be exactly solved using their symmetries. However, the Hamiltonians associated with realistic physical problems are usually too complicated to determine their eigenvalues and eigenstates exactly, being necessary in many cases the use of approximate methods to solve them. In general, there is no miraculous recipe or universal solution for all problems of this kind, so in this chapter, we will be satisfied with building a good collection of tools:

- *Variational approach*: This non-systematical technique is useful for the study of strongly correlated systems (such as multielectron atoms, molecules, and crystalline solids) where perturbative methods do not usually apply. The success of the approach relies on an initial guess of the ground state eigenstate, which needs itself a certain understanding of the physical properties of the system. Various seminal theories, such as the Hartree-Fock and Thomas-Fermi atomic theories, the Hohenberg-Kohn density functional theory, and the Bardeen-Cooper-Schrieffer superconductivity theory, make use of the variational method.

- *Perturbation theory*: This systematical technique is useful when the precise Hamiltonian of the system can be separated into an unperturbed part with known solution and a perturbation part whose energy scale is relatively small as compared to the unperturbed piece. The exact energy eigenvalues and states can then be expressed as formal power series of the ratio of the perturbation to the exact Hamiltonian energy scales, whose coefficients can be recursively computed.

- *WKB approximation:* This technique, named after the physicists Wentzel, Kramers and Brillouin, is used to solve the one-dimensional Schrödinger equation in scenarios where the potential varies gradually in relation to the de Broglie wavelength of the particle. Most common applications of the WKB approximation are the computation of bound state energies and tunneling rates.

7.1 Variational methods

The (Ritz) variation principle provides a simple way to place an upper bound on the ground state energy of a given quantum system, being particularly useful in proving the existence of bound states. Moreover, in certain situations, it can also provide estimates for higher energy levels.

The idea behind the method is quite straightforward. Imagine for simplicity that we are confronted with an arbitrary quantum system with time-independent Hamiltonian H and discrete spectrum E_n, namely

$$H|n\rangle = E_n|n\rangle \qquad n = 0, 1, 2, \ldots \tag{7.1}$$

with $E_n \leq E_{n+1}$. Guided by the physical properties of the system, we then select a non-necessarily normalized trial state $|\varphi(\alpha)\rangle$ depending on one or multiple model parameters α_i. This defines a subset of the Hilbert space where the approximate solution is sought. Given the expansion of this state in the previous basis,

$$|\varphi\rangle = \sum_n a_n|n\rangle\,, \tag{7.2}$$

we have[1]

$$\langle\varphi(\alpha)|H|\varphi(\alpha)\rangle = \sum_{n,m=0}^{\infty} a_m^\star a_n \langle m|H|n\rangle = \sum_{n=0}^{\infty} |a_n|^2 E_n \geq E_0 \sum_{n=0}^{\infty} |a_n|^2 = E_0\langle\varphi(\alpha)|\varphi(\alpha)\rangle\,. \tag{7.3}$$

From this, we conclude that the so-called Rayleigh-Ritz quotient,

$$E(\alpha) = \frac{\langle\varphi(\alpha)|H|\varphi(\alpha)\rangle}{\langle\varphi(\alpha)|\varphi(\alpha)\rangle}\,, \tag{7.4}$$

determining the expectation value of the Hamiltonian for the general family of states $|\varphi(\alpha)\rangle$ must necessarily exceed the ground state energy E_0 for all α,

$$E(\alpha) = \frac{\langle\varphi(\alpha)|H|\varphi(\alpha)\rangle}{\langle\varphi(\alpha)|\varphi(\alpha)\rangle} \geq E_0\,. \tag{7.5}$$

The most stringent upper bound for a given set of states $|\varphi(\alpha)\rangle$ is obtained by minimizing the above expression with respect to the model parameters α_i, namely

$$\left.\frac{\partial E}{\partial \alpha}\right|_{\alpha=\alpha_\star} = 0\,, \qquad \longrightarrow \qquad E_0 \leq E\left(\alpha_\star\right) \qquad \text{(Upper bound)}\,. \tag{7.6}$$

Note that, although the variational method provides a bound on E_0, it does not tell us how much $E\left(\alpha_\star\right)$ and E_0 actually differ. Indeed, unless someone tells us the true answer, we have no way of telling how good our approximation actually is. Everything will depend on our ability to choose a physically relevant ansatz. In doing so, it is important to take into account several considerations, such as:

[1] For a non-degenerate ground state, the equality in this expression is achieved only for $a_0 = 1$, thereby implying that $a_n = 0$ for all $n \neq 0$.

- Symmetries: The trial function should respect the symmetries of the problem, such as parity, rotational, or translational invariance. This symmetry matching can simplify calculations dramatically, leading also often to more accurate approximations.
- Number of nodes: The trial function should have a number of nodes (points where the function passes through zero) consistent with the expected behavior of the state being approximated.
- Smoothness and continuity: The trial function should be smooth and continuous, especially in regions where abrupt changes would be unphysical.
- Asymptotic behavior: The trial function should approach zero or a sensible limit as it extends to infinity or separates from the region of interest, reflecting characteristic boundary conditions or decay properties.

Example 41: Ground state of the Helium atom

The neutral Helium atom is the simplest multielectron atom: a three-body system of two electrons and a nucleus containing two protons and one or two neutrons. After factoring out the center of mass motion and neglecting the spin-orbit interaction, the associated quantum Hamiltonian takes the form ($e^2 = q_e^2/(4\pi\epsilon_0)$)

$$H = \frac{\vec{P}_1^2}{2m} - \frac{e^2}{r_1} + \frac{\vec{P}_2^2}{2m} - \frac{e^2}{r_2} + \frac{e^2}{|\vec{r}_1 - \vec{r}_2|}\,, \tag{7.7}$$

with $Z = 2$ the nuclear charge number, m the reduced mass of the two electrons and $\vec{r}_1$ and $\vec{r}_2$ their coordinates under the provisional assumption that these two spin 1/2 particles can be actually distinguished. This premise is of course false, but, since electrons have only two possible substates, the fiction of two (and only two!) indistinguishable electrons may be maintained while solving (7.7), provided of course that we think of one electron as having spin up and the other spin down.
Due to the last term in the Hamiltonian (7.7), the associated Schrödinger equation has no exact solution. The variational method can be used, however, to obtain an approximate expression for the ground state energy of the system and the associated eigenfunction. To do this, let us notice that in the absence of the repulsive electrostatic between electrons, the Hamiltonian (7.7) becomes just the sum of two independent copies of a hydrogenic atom (i.e. that of a nucleus with nuclear charge Z and a single electron). In this limit, the Schrödinger equation becomes separable, admitting thus a solution

$$E_{n_1 n_2} = E_{n_1} + E_{n_2} = -Z^2 \left(\frac{1}{n_1^2} + \frac{1}{n_2^2}\right) Ry\,, \tag{7.8}$$

$$\varphi(\vec{r}_1, \vec{r}_2) = \phi_{n_1 l_1 m_1}(\vec{r}_1)\phi_{n_2 l_2 m_2}(\vec{r}_2)\,, \tag{7.9}$$

with

$$Ry = \frac{e^2}{2a_0} = \frac{me^4}{2\hbar^2} \approx 13.6\,\text{eV} \tag{7.10}$$

the *Rydberg* constant, $a_0 = \hbar^2/(me^2) \approx 0.5\mathring{A}$ the Bohr radius, n, l and m the usual quantum numbers of the hydrogenic atom and $\phi_{nlm}(r)$ the associated energy eigenstates. [a] In the ground state of the system, the two electrons occupy the lowest state

$n = 1$, $l = 0$, $m = 0$, i.e., $n_1 = n_2 = 1$. The associated spatial wave function is then given by

$$\varphi_{11}(\vec{r}_1, \vec{r}_2, \vec{S}_2) = \phi_{100}(\vec{r}_1)\,\phi_{100}(\vec{r}_2)\ , \tag{7.11}$$

with

$$\phi_{100}(\vec{r}) = R_{10}(r)Y_{00}(\Omega) = \frac{1}{\sqrt{\pi}}\left(\frac{Z}{a_0}\right)^{3/2} e^{-\frac{Zr}{a_0}}\ . \tag{7.12}$$

Equations (7.11) and (7.12) will be the starting point for our variational method. In particular, let us consider a trial wave function coinciding in form with (7.11), but with the atomic number Z replaced by a general free parameter α aiming to accommodate the physical idea that, due to the screening of the other electron, each electron sees in practice a lower effective charge, namely

$$\varphi(\vec{r}_1, \vec{r}_2) = \phi(\vec{r}_1)\,\phi(\vec{r}_2)\ , \qquad \phi(\vec{r};\alpha) = \sqrt{\frac{\alpha^3}{\pi a_0^3}}e^{-\alpha r/a_0}\ . \tag{7.13}$$

For this trial wave function, the energy expectation value takes the form

$$E(\alpha) = \int d^3\vec{r}_1 d^3\vec{r}_2\ \phi^\star(\vec{r}_1)\,\phi^\star(\vec{r}_2)\,H\phi(\vec{r}_1)\,\phi(\vec{r}_2)\ . \tag{7.14}$$

To evaluate this expression, let us rewrite Eq. (7.7) as

$$H = H_\alpha\left(\vec{P}_1, \vec{r}_1\right) + H_\alpha\left(\vec{P}_2, \vec{r}_2\right) + e^2\left[(\alpha - Z)\left(\frac{1}{r_1} + \frac{1}{r_2}\right) + \frac{1}{|\vec{r}_1 - \vec{r}_2|}\right]\ ,$$

with

$$H_\alpha = \frac{\vec{P}^2}{2m} - \frac{\alpha\, e^2}{r}\ . \tag{7.15}$$

Taking into account that $\phi(\vec{r}, \alpha)$ is by construction the ground state of the Hamiltonian H_α, the expectation value (7.14) becomes

$$E(\alpha) = -2\alpha^2 Ry + e^2\left[2(\alpha - Z)\underbrace{\int d^3\vec{r}\frac{|\phi(\vec{r})|^2}{r}}_{I_1} + \underbrace{\int d^3\vec{r}_1 d^3\vec{r}_2 \frac{|\phi(\vec{r}_1)|^2\,|\phi(\vec{r}_2)|^2}{|\vec{r}_1 - \vec{r}_2|}}_{I_2}\right]\ . \tag{7.16}$$

The computation of the first integral in this expression is straightforward,

$$I_1 = 4\pi\frac{\alpha^3}{\pi a_0^3}\int dr\, r\, e^{-2\alpha r/a_0} = \frac{\alpha}{a_0}\ . \tag{7.17}$$

For computing the second integral we choose a reference system in which the z axis for the second particle aligns with the direction $\vec{r}_1$ of the first particle, such that the angle θ between the two particles coincides with the polar angle θ_2 of the second particle and

$$|\vec{r}_1 - \vec{r}_2| = \sqrt{\left(\vec{r}_1 - \vec{r}_2\right)^2} = \sqrt{r_1^2 + r_2^2 - 2r_1 r_2\cos\theta_2}\ . \tag{7.18}$$

Performing the integration over the angular variables for the first particle, we then have

$$\begin{aligned} I_2 &= 8\pi^2 \left(\frac{\alpha^3}{\pi a_0^3}\right)^2 \int dr_1 r_1^2 e^{-2\alpha r_1/a_0} \int dr_2 r_2^2 e^{-2\alpha r_2/a_0} \int_{-1}^{+1} \frac{d(\cos\theta_2)}{\sqrt{r_1^2+r_2^2-2r_1 r_2\cos\theta_2}} \\ &= 8\pi^2 \left(\frac{\alpha^3}{\pi a_0^3}\right)^2 \int dr_1 r_1^2 e^{-2\alpha r_1/a_0} \int dr_2 r_2^2 e^{-2\alpha r_2/a_0} \frac{|r_1+r_2|-|r_1-r_2|}{r_1 r_2} \\ &= 2\times 8\pi^2 \left(\frac{\alpha^3}{\pi a_0^3}\right)^2 \int_{r_2}^{\infty} dr_1 \left(\frac{2}{r_1}\right) r_1^2 e^{-2\alpha r_1/a_0} \int_0^{\infty} dr_2\, r_2^2 e^{-2\alpha r_2/a_0} \\ &= 32\pi^2 \left(\frac{\alpha^3}{\pi a_0^3}\right)^2 \int_0^{\infty} dr_2\, r_2^2 \left(\frac{a_0 r_2}{2\alpha}+\frac{a_0^2}{4\alpha^2}\right) e^{-4\alpha r_2/a_0} \\ &= \frac{5\alpha}{8a_0}\,, \end{aligned} \tag{7.19}$$

where in the third step (in blue) we have taken into account that the integral is symmetric in r_1 and r_2 to focus on the regime $r_1 > r_2$ where

$$\frac{|r_1+r_2|-|r_1-r_2|}{r_1 r_2} = \frac{2r_2}{r_1 r_2} = \frac{2}{r_1}\,, \tag{7.20}$$

while doubling our result. Collecting all terms, we get

$$E(\alpha) = \left[-2\alpha^2 + 4(\alpha - Z)\alpha + \frac{5}{4}\alpha\right] Ry = \left[-2\alpha^2 + 4\left(\alpha - Z + \frac{5}{16}\right)\alpha\right] Ry\,, \tag{7.21}$$

which becomes minimal for $\alpha_* = Z - 5/16$. This leads to a ground energy estimate formally equal to twice the energy of an electron orbiting a nucleus of charge $(Z - 5/16)e$,

$$E(\alpha_\star) = -2\left(Z - \frac{5}{16}\right)^2 Ry \simeq -77.45\,\text{eV}\,. \tag{7.22}$$

Note that this estimate is much closer to the experimental value $E_0^{\text{exp}} = -79.003$ eV than the one obtained when neglecting the Coulomb repulsion of electrons, $E_0^{(0)} = -108.8$ eV. This conclusion can be easily extended to other atoms containing two electrons (cf. Table 7.1). The agreement with experiments can be further improved by considering a better trail function accounting, for instance, for the dependence of the wave function on the distance between electrons or the motion of the nucleus.

[a] Note, once again, that the same spatial wave function can be shared by the two electrons only if one of them has spin up and the other has spin down.

7.1.1 Bound states

In spite of their limitations, the variational method is able to provide a conclusive answer to a specific type of question: the existence of bound states. In particular, if we are able to

Atom	Z	$E_0^{\text{non-int}}$ (eV)	E_0^{var} (eV)	E_0^{exp} (eV)
He	2	− 108.8	−77.45	−79.00
Li (1)	3	− 244.8	−196.46	−198.08
Be (2)	4	− 435.2	−369.86	−371.57
B (3)	5	− 680.0	−597.66	−599.50
C (4)	6	− 979.2	−879.86	−881.88
N (5)	7	− 1332.8	−1216.46	−1218.74
O (6)	8	− 1740.8	−1607.46	−1610.07

Table 7.1: Comparison between the ground state energy values $E_0^{\text{non-int}}$ for non-interacting electrons, the variational result E_0^{var} and the experimental value E_0^{exp} for several atoms with two electrons. The number of removed electrons is indicated in parenthesis in the first column.

find a trial wave function leading to an energy eigenvalue smaller than zero, the relation (7.6) guarantees that the actual ground state of the system has also an energy value below zero! Consider for instance a particle moving in a one-dimensional attractive potential $V(x)$, with $V(x) \leq 0$ for all x and $V(x) \to 0$ as $|x| \to \infty$. Given the corresponding Hamiltonian

$$H = -\frac{\hbar^2}{2m}\partial_x^2 + V(x)\,,$$

and choosing a normalized Gaussian trial function $\varphi = (2\alpha/\pi)^{1/4} e^{-\alpha x^2}$, we have

$$E(\alpha) = \langle\varphi|H|\varphi\rangle = \frac{\hbar^2}{2m}\alpha + \underbrace{\sqrt{\frac{2\alpha}{\pi}}\int_{-\infty}^{\infty} dx\, V(x) e^{-2\alpha x^2}}_{I} = \frac{\hbar^2}{2m}\alpha + I\,.$$

Minimizing this result with respect to α, $\partial E/\partial\alpha|_{\alpha=\alpha_\star} = 0$, we obtain an implicit equation for $\alpha_\star$,

$$0 = \frac{\hbar^2}{2m} + \frac{I}{2\alpha_\star} + \sqrt{\frac{2\alpha_\star}{\pi}}\int_{-\infty}^{\infty} dx\, V(x)\left(-2x^2\right) e^{-2\alpha_\star x^2}\,,$$

where the second term arises from differentiating the normalization prefactor in I with respect to α, and the third term from differentiating the integrand. Solving for I,

$$I = -2\frac{\hbar^2}{2m}\alpha_\star + 2\alpha_\star\sqrt{\frac{2\alpha_\star}{\pi}}\int_{-\infty}^{\infty} V(x)\left(+2x^2\right) e^{-2\alpha_\star x^2}\,,$$

and substituting the result into the equation for the expectation value above, we obtain

$$E(\alpha_*) = \frac{\hbar^2}{2m}\alpha_\star + I = -\frac{\hbar^2}{2m}\alpha_\star + 2\alpha_\star\sqrt{\frac{2\alpha_\star}{\pi}}\int dx V(x)\left(2x^2\right) e^{-2\alpha_\star x^2} \leq 0$$

for $V(x) \leq 0$. Since this is an upper bound on the ground state energy, this must be also negative, proving that exist, at least, one bound state.

7.1.2 Excited states

The variational method can also be applied to estimate the energies of the first few excited states. To this end, one should consider trial functions that are orthogonal to the previously computed ones. For instance, having determined the ground state $|\varphi_0(\alpha_*)\rangle$, an energy

estimate for the first excited state $|\varphi_1\rangle$ follows from minimizing the quotient

$$\frac{\partial}{\partial\alpha}\frac{\langle\varphi_1(\alpha)|H|\varphi_1(\alpha)\rangle}{\langle\varphi_1(\alpha)|\varphi_1(\alpha)\rangle}=0\,, \tag{7.23}$$

with

$$\langle\varphi_1(\alpha)|\varphi_0(\alpha_*)\rangle=0\,. \tag{7.24}$$

While this method could be a priori extended to any excited state, the process becomes increasingly complex as the number of states increases, being usually necessary to incorporate the different orthogonality constraints as Lagrange multipliers.

Example 42: The first excited state of the harmonic oscillator

Imagine that you know the ground state of the harmonic oscillator or have estimated it previously by the variational method,

$$\varphi_0\left(x\right)=\left(\frac{m\omega}{\pi\hbar}\right)^{1/4}e^{-m\omega x^2/2\hbar}\,. \tag{7.25}$$

To estimate the energy E_1 of the first excited state, we can select a normalized odd trial function

$$\varphi_1(x,\alpha)=Bxe^{-\alpha x^2}\,,\qquad B=\left(32\alpha^3/\pi\right)^{1/4}\,, \tag{7.26}$$

vanishing as $x\to\pm\infty$, containing exactly one node and ensuring automatically the orthogonality condition

$$\langle\varphi_0|\varphi_1\rangle=B\left(\frac{m\omega}{\pi\hbar}\right)^{1/4}\int_{-\infty}^{+\infty}xe^{-\alpha x^2}e^{-m\omega x^2/2\hbar}\,dx=0\,. \tag{7.27}$$

Applying the Rayleigh-Ritz quotient (7.23) to this trial wave function,

$$E_1(\alpha)=B^2\int_{-\infty}^{+\infty}xe^{-\alpha x^2}\left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}+\frac{1}{2}m\omega^2x^2\right]xe^{-\alpha x^2}\,dx=\frac{3\hbar^2\alpha}{2m}+\frac{3m\omega^2}{8\alpha}\,, \tag{7.28}$$

and minimizing the result with respect to α, we get $\alpha_*=m\omega/2\hbar$ and therefore

$$E_1\left(\alpha_*\right)=\frac{3\hbar\omega}{2}\,,\qquad \varphi_1\left(x,\alpha_*\right)=\left(\frac{4m^3\omega^3}{\pi\hbar^3}\right)^{1/4}xe^{-m\omega x^2/2\hbar}\,. \tag{7.29}$$

These results are indeed in complete agreement with the exact solution, since the chose family of trails functions (7.26) includes the correct wave function.

7.2 Perturbation theory for non-degenerate states

Let us consider a Hamiltonian H separable into a unperturbed part H_0 with a known independent-of-time solution, and a static perturbation term λH_1,

$$H(\lambda)=H_0+\lambda H_1\,, \tag{7.30}$$

with λ a small parameter characterizing the relative magnitude of the perturbation against H_0 and associated, for instance, with a coupling constant in the theory. Let's assume for the time being that the states of the so-called *unperturbed Hamiltonian* H_0 are non-degenerate and discrete, meaning that for each normalized state $|n^{(0)}\rangle$, there exists a unique eigenvalue $E_n^{(0)}$,

$$H_0|n^{(0)}\rangle = E_n^{(0)}|n^{(0)}\rangle\,, \tag{7.31}$$

as it happens, for instance, in the one-dimensional harmonic oscillator. Our objective is to determine the eigenvalues $E_n(\lambda)$ and eigenvectors $|n(\lambda)\rangle$ of $H(\lambda)$, specifically

$$H(\lambda)|n(\lambda)\rangle = E_n(\lambda)|n(\lambda)\rangle\,, \tag{7.32}$$

with the *adiabatic continuity* conditions

$$\lim_{\lambda\to 0} E_n(\lambda) = E_n^{(0)}\,, \qquad \lim_{\lambda\to 0}|n(\lambda)\rangle = |n^{(0)}\rangle\,. \tag{7.33}$$

To do this, let us assume that $E_n(\lambda)$ and $|n(\lambda)\rangle$ admit a perturbative series expansion in λ around the known quantities $E_n^{(0)}$ and $|n^{(0)}\rangle$,

$$E_n(\lambda) = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \ldots + E_n^{(L)}\,, \tag{7.34}$$

$$|n(\lambda)\rangle = |n^{(0)}\rangle + \lambda|n^{(1)}\rangle + \lambda^2|n^{(2)}\rangle + \ldots + \lambda^L|n^{(L)}\rangle\,, \tag{7.35}$$

with the superscripts $(1), (2), \ldots (L)$ indicating the order of the approximation in powers of λ. Of course, this type of approximation is only valid if the states $|n\rangle$ and $|n^{(0)}\rangle$ do not differ much qualitatively, since we are treating λ as a variable parameter that in the limit $\lambda \to 0$ returns us to the original system. Additionally, it is noteworthy that this procedure fails miserably if the unperturbed and perturbed energy levels cross or if we deal with bound states, which cannot be perturbatively obtained from unbound states like those of the free Schrödinger equation.

Although the perturbation series (7.34) and (7.35) are not necessarily convergent in the mathematical sense, their first terms can provide a good approximation to the actual result. Substituting the expressions (7.34) and (7.35) into the equation (7.32),

$$\begin{aligned}&(H_0 + \lambda H_1)\left(|n^{(0)}\rangle + \lambda|n^{(1)}\rangle + \lambda^2|n^{(2)}\rangle + \ldots + \lambda^L|n^{(L)}\rangle\right) \\ &= (E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \ldots + \lambda^N E_n^{(L)})(|n^{(0)}\rangle + \lambda|n^{(1)}\rangle + \lambda^2|n^{(2)}\rangle + \ldots + \lambda^{(L)}|n^{(L)}\rangle)\,,\end{aligned} \tag{7.36}$$

and identifying terms order by order in λ, we obtain

$$(H_0 - E_n^{(0)})|n^{(0)}\rangle = 0\,, \tag{7.37}$$

$$(H_0 - E_n^{(0)})|n^{(1)}\rangle + H_1|n^{(0)}\rangle - E_n^{(1)}|n^{(0)}\rangle = 0\,, \tag{7.38}$$

$$(H_0 - E_n^{(0)})|n^{(2)}\rangle + H_1|n^{(1)}\rangle - E_n^{(1)}|n^{(1)}\rangle - E_n^{(2)}|n^{(0)}\rangle = 0\,. \tag{7.39}$$

$$(H_0 - E_n^{(0)})|n^{(3)}\rangle + H_1|n^{(2)}\rangle - E_n^{(1)}|n^{(2)}\rangle - E_n^{(2)}|n^{(1)}\rangle - E_n^{(3)}|n^{(0)}\rangle = 0\,, \tag{7.40}$$

$$\vdots \qquad\qquad \vdots$$

$$(H_0 - E_n^{(0)})|n^{(L)}\rangle + H_1|n^{(L-1)}\rangle - \sum_{K=1}^{L} E_n^{(K)}|n^{(L-K)}\rangle = 0\,. \tag{7.41}$$

Let us now see how we can calculate the quantities $E_n^{(L)}$ and $|n^{(L-1)}\rangle$ in these expressions. Since the equation (7.32) is homogeneous in $|n(\lambda)\rangle$, determining these solutions requires fixing the normalization of the states in some way. For reasons that will soon become clear, it is convenient to adopt a normalization

$$\langle n^{(0)}|n(\lambda)\rangle = 1\,, \tag{7.42}$$

instead of the standard normalization $\langle n(\lambda)|n(\lambda)\rangle = 1$, being of course possible to recover the standard normalization afterwards. In fact, since the Schrödinger equation is homogeneous, given a solution $|n(\lambda)\rangle$, it is always possible to construct another solution $c(\lambda)|n(\lambda)\rangle$, so that the normalized state $1/a(\lambda)|n(\lambda)\rangle$ with $a(\lambda) = \langle n^{(0)}|n(\lambda)\rangle$ satisfies (7.32).

7.2.1 First order solution

Multiplying Eq. (7.38) by $\langle n^{(0)}|$ and taking into account that H_0 is a self-adjoint operator,

$$\langle n^{(0)}|\left(H_0 - E_n^{(0)}\right) = 0\,, \tag{7.43}$$

we obtain

$$\underbrace{\langle n^{(0)}|\left(H_0 - E_n^{(0)}\right)|n^{(1)}\rangle}_{=0} + \langle n^{(0)}|\left(H_1 - E_n^{(1)}\right)|n^{(0)}\rangle = 0\,, \tag{7.44}$$

or equivalently

$$\boxed{E_n^{(1)} = \langle n^{(0)}|H_1|n^{(0)}\rangle\,,} \tag{7.45}$$

where we have taken into account the normalization condition $\langle n^{(0)}|n^{(0)}\rangle = 1$. This is the simplest and most important result in perturbation theory: to calculate the energy shift of a state, we only need to calculate the expectation value of the change in the Hamiltonian $\Delta H \equiv \lambda H_1 = H - H_0$ in the unperturbed state. To calculate the associated eigenvector $|n^{(1)}\rangle$, we expand it in the orthonormal basis $\{|m^{(0)}\rangle\}$ of the Hilbert space of H_0,

$$H_0|m^{(0)}\rangle = E_m^{(0)}|m^{(0)}\rangle\,. \tag{7.46}$$

Note that, due to the normalization (7.42),

$$\lambda\langle n^{(0)}|n^{(1)}\rangle + \lambda^2\langle n^{(0)}|n^{(2)}\rangle + \ldots = 0\,, \quad \longrightarrow \quad \langle n^{(0)}|n^{(1)}\rangle = \langle n^{(0)}|n^{(2)}\rangle = \ldots = 0\,. \tag{7.47}$$

this expansion does not include the term with $m = n$,

$$|n^{(1)}\rangle = \sum_{m\neq n}|m^{(0)}\rangle\underbrace{\langle m^{(0)}|n^{(1)}\rangle}_{\equiv c_m} = \sum_{m\neq n} c_m|m^{(0)}\rangle\,. \tag{7.48}$$

Multiplying now Eq. (7.38) by $\langle m^{(0)}|$, we obtain $(m \neq n)$

$$E_m^{(0)}\underbrace{\langle m^{(0)}|n^{(1)}\rangle}_{=c_m} + \langle m^{(0)}|H_1|n^{(0)}\rangle = E_n^{(0)}\underbrace{\langle m^{(0)}|n^{(1)}\rangle}_{=c_m} + E_n^{(1)}\underbrace{\langle m^{(0)}|n^{(0)}\rangle}_{=0}\,, \tag{7.49}$$

from which we get

$$c_m = \frac{\langle m^{(0)}|H_1|n^{(0)}\rangle}{E_n^{(0)} - E_m^{(0)}}, \qquad \text{and} \qquad |n^{(1)}\rangle = \sum_{m\neq n}\frac{\langle m^{(0)}|H_1|n^{(0)}\rangle}{E_n^{(0)} - E_m^{(0)}}|m^{(0)}\rangle\,, \tag{7.50}$$

with $E_n^{(0)} \neq E_m^{(0)}$ for $n \neq m$.

7.2.2 Iterative solution

The general solution of the system of equations (7.37)-(7.41) can be obtained iteratively. In particular, given the solution at order $L-1$, we can multiply equation (7.41) by $\langle n^{(0)}|$ to eliminate the first term in that expression using (7.43) and then use again the normalization conditions (7.47). This procedure gives us the following correction at order λ^L,

$$E_n^{(L)} = \langle n^{(0)}|H_1|n^{(L-1)}\rangle\,. \tag{7.51}$$

Multiplying (7.41) by $\langle m^{(0)}|$ with $m \neq n$ and using $\langle m^{(0)}|n^{(0)}\rangle = 0$, we can also obtain the approximate solution for the eigenvector $|n^{(L)}\rangle$ in terms of eigenvectors of the unperturbed Hamiltonian,

$$\langle m^{(0)}|n^{(L)}\rangle = \frac{\langle m^{(0)}|H_1|n^{(L-1)}\rangle}{E_n^{(0)} - E_m^{(0)}} - \frac{\sum_{K=1}^{L-1} E_n^{(K)}\langle m^{(0)}|n^{(L-K)}\rangle}{E_n^{(0)} - E_m^{(0)}}\,. \tag{7.52}$$

In this course, we will mainly be interested in calculating first-order corrections in the eigenstates and second-order corrections in the energies. In this case, the combination of the expressions (7.34), (7.45), (7.50), and (7.51) gives us

$$|n(\lambda)\rangle = |n^{(0)}\rangle + \sum_{m\neq n} \frac{\langle m^{(0)}|\Delta H|n^{(0)}\rangle}{E_n^{(0)} - E_m^{(0)}}|m^{(0)}\rangle + \dots\,. \tag{7.53}$$

$$E_n^{(2)} = E_n^{(0)} + \langle n^{(0)}|\Delta H|n^{(0)}\rangle + \sum_{m\neq n} \frac{|\langle m^{(0)}|\Delta H|n^{(0)}\rangle|^2}{E_n^{(0)} - E_m^{(0)}} + \dots \tag{7.54}$$

where we have absorbed λ into the definition of the perturbation Hamiltonian, $\Delta H \equiv \lambda H_1$. Note that:

- At order λ, the state $|n(\lambda)\rangle$ is automatically normalized, as follows from (7.47),

$$\langle n(\lambda)|n(\lambda)\rangle = (\langle n^{(0)}| + \lambda\langle n^{(1)}|)(|n^{(0)}\rangle + \lambda|n^{(1)}\rangle) + \mathcal{O}(\lambda^2) = 1 + \mathcal{O}(\lambda^2)\,. \tag{7.55}$$

- The energy of the (non-degenerate) ground state at first order in perturbation theory exceeds generically the true ground energy. This can be easily seen by expressing $E_0^{(0)} + \lambda E_0^{(1)}$ in terms of expectation values over the unperturbed ground state $|0^{(0)}\rangle$, namely

$$E_0^{(0)} + \lambda E_0^{(1)} = \langle 0^{(0)}|H^{(0)} + \lambda H_1|0^{(0)}\rangle = \langle 0^{(0)}|H(\lambda)|0^{(0)}\rangle\,. \tag{7.56}$$

  Since, according to the variational principle, the expectation value of the Hamiltonian on *any* normalized state exceeds the ground state energy $E_0(\lambda)$, this implies

$$E_0^{(0)} + \lambda E_0^{(1)} = \langle 0^{(0)}|H(\lambda)|0^{(0)}\rangle \geq E_0(\lambda) \tag{7.57}$$

- Since the unperturbed excited states exceed the unperturbed ground state energy, $E_m^{(0)} > E_0^{(0)}$, the second-order correction to the ground state is always negative. This is especially relevant in situations where the first-order correction vanishes, since in this case the perturbation inevitably decreases the energy of the ground state.

- If the matrix elements of ΔH are of similar magnitude, neighboring energy levels contribute more significantly than distant ones.
- The consistency of the perturbative series in expressions (7.53) and (7.54) requires the first energy correction to be small as compared to the ground state for $m = n$,

$$\langle n^{(0)}|\Delta H|n^{(0)}\rangle \ll E_n^{(0)}\,, \tag{7.58}$$

and that the matrix elements $|\langle m^{(0)}|H_1|n^{(0)}\rangle|$ with $m \neq n$ are smaller than the separation between the unperturbed levels $E_n^{(0)} - E_m^{(0)}$,

$$\left|\frac{\langle m^{(0)}|\Delta H|n^{(0)}\rangle}{E_n^{(0)} - E_m^{(0)}}\right| \ll 1\,. \tag{7.59}$$

- If H_0 would contain also a continuous spectrum, the formulas (7.53) and (7.54) would be still valid, at least formally. However, the summatories would contain an integral part.

7.2.3 Matricial form

It is common to refer to the process in the previous section as the *diagonalization of the Hamiltonian*. To understand this name, let's revisit what we just did, rewriting everything in matrix notation. In the basis of the Hilbert space generated by the eigenstates $\{|n^{(0)}\rangle\}$, the unperturbed Hamiltonian is diagonal,

$$[H_0]_{ij} = \langle n_i^{(0)}|H_0|n_j^{(0)}\rangle = \delta_{ij}E_j^{(0)}\,, \quad \rightarrow \quad H_0 = \begin{pmatrix} E_1^{(0)} & 0 & 0 & \dots \\ 0 & E_2^{(0)} & 0 & \dots \\ 0 & 0 & E_3^{(0)} & \dots \\ \cdot & \cdot & \cdot & \dots \\ \cdot & \cdot & \cdot & \dots \end{pmatrix}, \tag{7.60}$$

and the eigenvalue problem $H_0|n^{(0)}\rangle = E_n^{(0)}|n^{(0)}\rangle$ is trivial. In particular, the eigenvectors are simply

$$|n_1^{(0)}\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \cdot \\ \cdot \end{pmatrix} \quad , \quad |n_2^{(0)}\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \cdot \\ \cdot \end{pmatrix} \quad , \quad |n_3^{(0)}\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ \cdot \\ \cdot \end{pmatrix} \quad , \quad \dots \tag{7.61}$$

and the eigenvalues are the diagonal elements of the matrix (7.60). When we include the perturbation ΔH, the total Hamiltonian becomes non-diagonal,

$$H = H_0 + \Delta H = \begin{pmatrix} E_1^{(0)} + \Delta H_{11} & \Delta H_{12} & \Delta H_{13} & \cdots \\ \Delta H_{21} & E_2^{(0)} + \Delta H_{22} & \Delta H_{23} & \cdots \\ \Delta H_{31} & \Delta H_{32} & E_3^{(0)} + \Delta H_{33} & \cdots \\ \cdot & \cdot & \cdot & \cdots \\ \cdot & \cdot & \cdot & \dots \end{pmatrix}, \tag{7.62}$$

requiring an additional diagonalization to determine the eigenvalues and eigenvectors of the new perturbed system. At first order in λ for the eigenstates and second order in the energies, we obtain

$$|n\rangle = |n^{(0)}\rangle + \sum_{m\neq n} \frac{\Delta H_{mn}}{E_n^{(0)} - E_m^{(0)}} |m^{(0)}\rangle + \dots \,. \tag{7.63}$$

$$E_n^{(2)} = E_n^{(0)} + \Delta H_{nn} + \sum_{m\neq n} \frac{|\Delta H_{mn}|^2}{E_n^{(0)} - E_m^{(0)}} + \dots \tag{7.64}$$

with

$$H_{mn} = \langle m^{(0)}|H_0 + \Delta H|n^{(0)}\rangle = E_n \delta_{mn} + \Delta H_{mn}\,, \tag{7.65}$$

This is simply the result (7.54) in matrix notation. [2] In the new basis (after normalizing the states (7.63)), we have

$$H = \begin{pmatrix} E_1 & 0 & 0 & \dots \\ 0 & E_2 & 0 & \dots \\ 0 & 0 & E_3 & \dots \\ . & . & . & \cdots \\ . & . & . & \cdots \end{pmatrix}, \quad \text{or} \quad H_{mn} = \langle m|H|n\rangle = E_n \delta_{mn}\,. \tag{7.66}$$

Example 43: Charged harmonic oscillator in an electric field

To illustrate the advantages of perturbation theory and verify the perturbative results in our first encounter with it, we will apply it to a problem that can be solved exactly: the Hamiltonian of a particle of mass m and charge q subjected to an elastic force and a constant electric field, namely

$$H = \frac{P^2}{2m} + \frac{1}{2} m\omega^2 X^2 - q\mathcal{E}X\,. \tag{7.67}$$

The exact solution involves something as trivial as completing the square to obtain the Hamiltonian of a harmonic oscillator with equilibrium position shifted to $X \to X - q\mathcal{E}/(m\omega^2)$ and energies modified by a constant downward shift,

$$H = \frac{P^2}{2m} + \frac{1}{2} m\omega^2 X^2 - q\mathcal{E}X = \frac{P^2}{2m} + \frac{1}{2} m\omega^2 \left(X - \frac{q\mathcal{E}}{m\omega^2}\right)^2 - \frac{1}{2}\frac{q^2\mathcal{E}^2}{m\omega^2}\,. \tag{7.68}$$

Since a translation does not modify the commutation relations on which the spectrum depends,

$$\left[X - \frac{q\mathcal{E}}{m\omega^2}, P\right] = [X, P] = i\hbar \tag{7.69}$$

[2] In fact, the formulas we have just found are useful for diagonalizing any matrix of the form $M = M_0 + Q$ with $M_{ij} = m_i\delta_{ij}$, $|Q_{ii}| \ll |m_i - m_j|$ and $Q_{ii} \ll |m_i|$!

the energies of this system can be obtained by a simple rescaling of the energies of the harmonic oscillator,

$$E_n = E_n^{(0)} - \frac{q^2\mathcal{E}^2}{2m\omega^2}\,. \tag{7.70}$$

Let's see if perturbation theory is able to partially or fully capture this exact result. To do so, we will assume that the last term in (7.67) is energetically subdominant, identifying

$$H_0 = \frac{P^2}{2m} + \frac{1}{2}m\omega^2 X^2\,, \qquad \Delta H = -q\mathcal{E}X\,. \tag{7.71}$$

At first order in the expansion parameter $\lambda = -q\mathcal{E}$, we have

$$E_n^{(1)} = E_n^{(0)} + \langle n^{(0)}|\Delta H|n^{(0)}\rangle = \left(n+\frac{1}{2}\right)\hbar\omega - q\mathcal{E}\langle n^{(0)}|X|n^{(0)}\rangle\,, \tag{7.72}$$

$$|n\rangle = |n^{(0)}\rangle - q\mathcal{E}\sum_{m\neq n}\frac{\langle m^{(0)}|X|n^{(0)}\rangle}{E_n^{(0)} - E_m^{(0)}}|m^{(0)}\rangle\,. \tag{7.73}$$

The perturbative problem is then reduced to the calculation of the matrix element

$$\langle m^{(0)}|X|n^{(0)}\rangle. \tag{7.74}$$

for the distinct cases $m = n$ and $m \neq n$. To do so, we express the position operator in terms of creation and annihilation operators.

$$X = \sqrt{\frac{\hbar}{2m\omega}}(a + a^\dagger)\,, \tag{7.75}$$

with

$$a^\dagger|n^{(0)}\rangle = \sqrt{n+1}|(n+1)^{(0)}\rangle\,, \qquad a|n^{(0)}\rangle = \sqrt{n}|(n-1)^{(0)}\rangle\,, \tag{7.76}$$

to obtain

$$\langle m^{(0)}|X|n^{(0)}\rangle = \sqrt{\frac{\hbar}{2m\omega}}\langle m^{(0)}|a + a^\dagger|n^{(0)}\rangle = \sqrt{\frac{\hbar}{2m\omega}}(\sqrt{n}\delta_{n,m+1} + \sqrt{n+1}\delta_{n,m-1})\,. \tag{7.77}$$

Substituting this result in the expressions (7.72) and (7.73), we obtain

$$E_n^{(1)} = 0\,, \qquad \forall n \in \mathbb{N}\,, \tag{7.78}$$

and

$$\begin{aligned}|n\rangle &= |n^{(0)}\rangle - q\mathcal{E}\sqrt{\frac{\hbar}{2m\omega}}\left[\frac{\sqrt{n}|(n-1)^{(0)}\rangle}{\hbar\omega(n-n+1)} + \frac{\sqrt{n+1}|(n+1)^{(0)}\rangle}{\hbar\omega(n-n-1)}\right] \\ &= |n^{(0)}\rangle + q\mathcal{E}\sqrt{\frac{1}{2m\hbar\omega^3}}\left[\sqrt{n+1}|(n+1)^{(0)}\rangle - \sqrt{n}|(n-1)^{(0)}\rangle\right]\,.\end{aligned}$$

The effect of the first-order perturbation is to mix each stationary state with the adjacent states.

Since the corrections to the energy levels are zero at first order, it is appropriate to calculate the second-order corrections.

$$E_n^{(2)} = \sum_{m\neq n} \frac{\left|\langle m^{(0)}|\Delta H|\, n^{(0)}\rangle\right|^2}{E_n^{(0)} - E_m^{(0)}} = q^2\mathcal{E}^2 \frac{\hbar}{2m\omega}\left[\frac{n+1}{-\hbar\omega} + \frac{n}{\hbar\omega}\right] = -\frac{q^2\mathcal{E}^2}{2m\omega^2}\,. \tag{7.79}$$

Surprisingly, this second-order perturbative result coincides with the exact solution (7.70). *Not so bad!*

7.3 Perturbation theory for degenerate states

The previous results assume that the states of the unperturbed Hamiltonian are non-degenerate, a condition that may not occur in highly symmetric scenarios, such as the hydrogen atom. This limitation is apparent, for example, in the expression (7.63), which diverges if there exists an energy $E_m^{(0)}$ such that $E_n^{(0)} = E_m^{(0)}$, which clearly cannot be classified as a *small* perturbation. What do we do in this typical case in atomic physics? To answer this question, let us consider a Hamiltonian

$$H = H_0 + \lambda H_1 \tag{7.80}$$

with a basis of known eigenstates that are partially or fully degenerate. Expanding eigenvalues and eigenstates in a perturbation series in λ around $E_n^{(0)}$ and $|n^{(0)}\rangle$,

$$E_{n,i} = E_n^{(0)} + \lambda E_{n,i}^{(1)} + \lambda^2 E_{n,i}^{(2)} + \dots\,, \qquad |n_i\rangle = |n_i^{(0)}\rangle + \lambda|n_i^{(1)}\rangle + \lambda^2|n_i^{(2)}\rangle + \dots\,, \tag{7.81}$$

and substituting into the Schrödinger equation, we obtain a result similar to the non-degenerate case,

$$(H_0 - E_n^{(0)})|n_i^{(0)}\rangle = 0\,, \tag{7.82}$$

$$(H_0 - E_n^{(0)})|n_i^{(1)}\rangle + (H_1 - E_{n,i}^{(1)})|n_i^{(0)}\rangle = 0\,, \tag{7.83}$$

$$(H_0 - E_n^{(0)})|n_i^{(2)}\rangle + (H_1 - E_{n,i}^{(1)})|n_i^{(1)}\rangle - E_{n,i}^{(2)}|n^{(0)}\rangle = 0\,, \tag{7.84}$$

where, for simplicity, we have stopped this time at order λ^2. Multiplying Eq. (7.83) by $\langle n_i^{(0)}|$ and taking into account that H_0 is a self-adjoint operator, we find the energy values at first order,

$$\underbrace{\langle n_i^{(0)}|(H_0 - E_n^{(0)})|n_i^{(1)}\rangle}_{\langle n_i^{(0)}|(H_0-E_n^{(0)})=0} + \langle n_i^{(0)}|(H_1 - E_{n,i}^{(1)})|n_i^{(0)}\rangle = 0\,, \longrightarrow\ E_{n,i}^{(1)} = \langle n_i^{(0)}|H_1|n_i^{(0)}\rangle\,, \tag{7.85}$$

so that the matrix H_1 is diagonal in the $|n_i^{(0)}\rangle$ basis, with eigenvalues $E_{n,i}^{(1)}$. Note, however, that our assumption for $|n_i^{(1)}\rangle$ is now

$$|n_i^{(1)}\rangle = \sum_{m\neq n}\sum_k c_{mk}|m_k^{(0)}\rangle + \sum_{j\neq i} c_{nj}^{(i)}|n_j^{(0)}\rangle\,, \tag{7.86}$$

with the first term summing over the elements of the basis corresponding to non-degenerate energy eigenvalues $E_m^{(0)} \neq E_n^{(0)}$ and the second term summing over the elements of the basis

orthogonal to $|n_i^{(0)}\rangle$, corresponding to the same unperturbed energy eigenvalue $E_n^{(0)}$. The coefficients c_{mk} can be found, as before, by multiplying (7.83) by $\langle m_k^{(0)}|$. We obtain

$$c_{mk} = \frac{\langle m_k^{(0)}|H_1|n_i^{(0)}\rangle}{E_n^{(0)} - E_m^{(0)}}\,. \tag{7.87}$$

If instead we apply $\langle n_j^{(0)}|$ to that equation,

$$E_n^{(0)}\langle n_j^{(0)}|n_i^{(1)}\rangle - E_n^{(0)}\langle n_j^{(0)}|n_i^{(1)}\rangle = -\langle n_j^{(0)}|H_1|n_i^{(0)}\rangle \;\; (i \neq j)\,, \tag{7.88}$$

we obtain the following consistency relation

$$\langle n_i^{(0)}|H_1|n_j^{(0)}\rangle = 0 \qquad \text{for} \qquad i \neq j\,. \tag{7.89}$$

For an arbitrarily chosen base $|n_j^{(0)}\rangle$, the fulfillment of (7.89) is not guaranteed a priori. If this occurs, a new basis must be found that satisfies this condition. Note that, although in non-trivial problems H_0 and H_1 do not necessarily commute and therefore cannot be simultaneously diagonalized, it is always possible to diagonalize (7.89) in the subspace of degenerate states G, where every state has the same energy eigenvalue and H_0 is simply the identity matrix,

$$H_0|_G = E_n^{(0)}\,\mathbb{I}|_G\,. \tag{7.90}$$

More specifically, if we denote the matrix elements in this subspace of dimension g as

$$[\bar{H}_1]_{ij} = \langle n_i^{(0)}|H_1|n_j^{(0)}\rangle\,, \quad i,j \leq g\,, \tag{7.91}$$

the eigenvalue problem

$$\bar{H}_1|\varphi_i^{(0)}\rangle = \bar{E}_i|\varphi_i^{(0)}\rangle\,, \qquad |\varphi_i^{(0)}\rangle = \sum_{j=0}^{g} a_{ij}|n_j^{(0)}\rangle\,, \tag{7.92}$$

provides us with a set of states that give rise to the same g-dimensional Hilbert space as $\{|n_i^{(0)}\rangle\,, i \leq q\}$ but with

$$\langle\varphi_i|\bar{H}_1|\varphi_j^{(0)}\rangle = 0 \quad \text{for} \quad i \neq j\,, \tag{7.93}$$

as long as all the eigenvalues $\bar{E}_i$ are different, which we assume to avoid complicating things further. Combining (7.85) with the result of this process, we obtain the first-order correction to the energy in perturbation theory, namely

$$E_{n,i}^{(1)} = \begin{cases} \bar{E}_i \quad \text{para} \quad i \leq g\,, \\ \langle n_i^{(0)}|H_1|n_i^{(0)}\rangle \quad \text{para} \quad i > g\,. \end{cases} \tag{7.94}$$

Having done that, we still need to determine the coefficients $c_{nj}^{(i)}$. For their derivation, one can consult chapter 5.1 of Sakurai's book.

7.4 Application to atoms in electromagnetic fields

7.4.1 The Stark effect

The so-called Stark effect is a spectral line splitting in atomic systems that appears in the presence of an external electric field. It was discovered in 1913 by Johannes Stark, who got the Nobel Prize for it in 1922. To analize this effect using the techniques of time-independent perturbation theory we just learned, let us consider the Hamiltonian of the hydrogen atom mildly perturbed by the presence of a constant electric field $\mathcal{E}$ in the z-direction,[3]

$$H = -\frac{\hbar^2}{2m}\nabla^2 - \frac{e^2}{r} + e\mathcal{E}Z\,, \tag{7.95}$$

where we can easily identify the small coupling constant $\lambda = -e\mathcal{E}$ and the perturbation Hamiltonian $H_1 = Z$. This new electromagnetic interaction breaks explicitly the hidden $SO(4)$ symmetry responsible for the degeneracy in the angular momentum quantum number l and the rotational symmetry responsible for the degeneracy in m. Therefore, we should expect these degeneracies to disappear, although not necessarily at first order in perturbation theory.

Since the energy levels of the hydrogen atom at $\mathcal{E} = 0$ have a degeneracy $g = n^2$ (ignoring spin), we should work a priori with degenerate perturbation theory in order to calculate the relevant matrix elements

$$\langle n, l', m'|H_1|n, l, m\rangle = \langle n, l', m'|Z|n, l, m\rangle\,. \tag{7.96}$$

Fortunately for us, the vast majority of the states $|n, l, m\rangle$ do not receive corrections at first order in perturbation theory, as can be easily deduced from simple symmetry arguments.[4] First, using the fact that the perturbation operator Z commutes with the L_z component of the angular momentum,

$$\begin{aligned} m\hbar\langle n, l', m'|Z|n, l, m\rangle &= \langle n, l', m'|ZL_z|n, l, m\rangle \\ &= \langle n, l', m'|L_zZ|n, l, m\rangle = m'\hbar\langle n, l', m'|Z|n, l, m\rangle\,, \end{aligned} \tag{7.97}$$

we conclude that the matrix elements with $m \neq m'$ are identically zero. On top of that, considering the transformation of the hydrogen atom states under the parity operator Π,

$$\Pi|n, l, m\rangle = (-1)^l|n, l, m\rangle\,, \tag{7.98}$$

we can write

$$\langle n, l', m'|Z|n, l, m\rangle = (-1)^{l+l'}\langle n, l', m'|\Pi Z\Pi^\dagger|n, l, m\rangle = (-1)^{l+l'+1}\langle n, l', m'|Z|n, l, m\rangle\,, \tag{7.99}$$

from where we also conclude that the considered matrix elements are zero for $l + l'$ equal to zero or even. This is sufficient to determine the corrections to the four degenerate states with $n = 2$,

$$|2, 0, 0\rangle\,, \quad |2, 1, 0\rangle\,, \quad |2, 1, 1\rangle\,, \quad |2, 1, -1\rangle\,, \tag{7.100}$$

[3] We neglect spins and the potential energy of the nucleus in the electric field.

[4] However, all states receive corrections at second order.

which span a 4×4 subspace in the full Hilbert space of the system. In particular, while the states $|2,1,\pm1\rangle$ remain unchanged, the state $|2,0,0\rangle$ mixes with the state $|2,1,0\rangle$. To calculate the associated matrix element $\langle 2,0,0|Z|2,1,0\rangle$ and its Hermitian conjugate, we make use of the the associated wave functions

$$\varphi_{200} = \frac{1}{\sqrt{8\pi}} e^{-r/(2a_0)} \left(1 - \frac{r}{2a_0}\right) , \qquad \varphi_{210} = \frac{re^{-r/(2a_0)}}{2\sqrt{6}a_0} \sqrt{\frac{3}{4\pi}} \cos z = \frac{ze^{-r/(2a_0)}}{2\sqrt{8\pi}} , \quad (7.101)$$

to get

$$\langle 2,0,0|\Delta H|2,1,0\rangle = -3e\mathcal{E}a_0 .$$

The first energy corrections are therefore given by the eigenvalues of the 4×4 matrix

$$[\bar{H}_1]_{ij} = 3a_0 \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} , \quad (7.102)$$

namely

$$E^{(1)}_{n=2} = 0, \, +3a_0 \, , -3a_0 \, . \quad (7.103)$$

The $n = 2$ level is thus split into three symmetrically-distributed energy levels (not four!) with spacing $3e\mathcal{E}a_0$,

$$E^{(0)}_{n=2} \longrightarrow \begin{cases} E^{(0)}_{n=2} + 3e\mathcal{E}a_0 \, , \\ E^{(0)}_{n=2} \, , \\ E^{(0)}_{n=2} - 3e\mathcal{E}a_0 \, , \end{cases} \quad (7.104)$$

not completely eliminating therefore the degeneracy of that level. Taking into account the Bohr radius $a_0 \sim 0.5 \times 10^{-8}$ cm, the consistency of perturbation theory requires

$$3e\mathcal{E}a_0 \simeq 1.5 \times 10^{-8} \mathcal{E} \, \text{eV} , \qquad \mathcal{E} \ll 10^9 \, \text{V/cm} . \quad (7.105)$$

In order to calculate the eigenvectors associated with the eigenvalues (7.103), we express them in terms of the subspace G of degenerate states at zeroth order,

$$|\varphi\rangle = a\,|\phi_{200}\rangle + b\,|\phi_{210}\rangle + c\,|\phi_{211}\rangle + d\,|\phi_{20-1}\rangle \, , \qquad |\varphi\rangle = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} , \quad (7.106)$$

and solve the corresponding eigenvalue problem $\bar{H}_1|\varphi^{(0)}\rangle = \bar{E}|\varphi^{(0)}\rangle$,

$$\begin{pmatrix} 0 & 3a_0 & 0 & 0 \\ 3a_0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \bar{E} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \quad \longrightarrow \quad 3a_0 \begin{pmatrix} b \\ a \\ 0 \\ 0 \end{pmatrix} = \bar{E} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \quad (7.107)$$

for the three roots in (7.103),

$$\bar{E} = \pm 3a_0 \quad \longrightarrow \quad a = b = 0 \quad \longrightarrow \quad |2,\pm\rangle = \frac{1}{\sqrt{2}}(|2,0,0\rangle \pm |2,1,0\rangle) \, , \quad (7.108)$$

$$\bar{E} = 0 \quad \longrightarrow \quad a = b = 0 \quad \longrightarrow \quad |2,1\rangle = c\,|2,1,1\rangle + d\,|2,1,-1\rangle \, , \quad (7.109)$$

with the numerical prefactors in these expressions related to the normalization conditions $\langle 2,\pm|2,\pm\rangle = 2a^2 = 1$ and $c^2 + d^2 = 1$. Note that, as a consequence of the parity violation of the interaction Hamiltonian H_1, the states $|2,\pm\rangle$ have non-zero dipole moment

$$\langle 2,\pm|\, z\, |2,\pm\rangle \simeq \pm\frac{1}{2}\left(\langle 2,0,0|Z|2,1,0\rangle + \langle 2,1,0|Z|2,0,0\rangle\right) = \pm 3a_0\,. \tag{7.110}$$

A short digression on parity

Classically, a parity transformation Π in three spatial dimensions can be understood as a kind of mirror symmetry $\Pi : \vec{x} \mapsto -\vec{x}$. [a] At this level, we also have $\Pi : \vec{P} \mapsto -\vec{P}$, since $\vec{P} = m\dot{\vec{x}}$.
In quantum mechanics, the parity operator Π is naturally defined by its action on the position basis,

$$\Pi|\vec{x}\rangle = |-\vec{x}\rangle\,,$$

or equivalently by its action on the wave function in that representation,

$$\Pi : \varphi(\vec{x}) \mapsto \varphi(-\vec{x})\,.$$

As in this basis the momentum operator is given by $\vec{P} = -i\hbar\partial/\partial\vec{X}$, we have also

$$\Pi|\vec{P}\rangle = |-\vec{P}\rangle\,.$$

Parity is therefore a discrete transformation, not involving a parameter to be varied at will, as happens for the continuous symmetries we have encountered till now. The parity operator is indeed not only unitary but also Hermitian, as follows immediately from

$$\Pi^\dagger\Pi = 1 \quad \text{and} \quad \Pi^2 = 1 \quad \Rightarrow \quad \Pi = \Pi^\dagger = \Pi^{-1}\,.$$

Action of parity on operators

Taking into account the action of parity on the position and momentum operators

$$\Pi\,\vec{X}\,\Pi^\dagger = -\vec{X}\,, \qquad \Pi\,\vec{P}\,\Pi^\dagger = -\vec{P}\,,$$

we can deduce its action on the angular momentum operator $\vec{L}$, namely

$$\Pi\,\vec{L}\,\Pi^\dagger = +\vec{L}\,.$$

Similarly, the spin and total angular momentum operators $\vec{S}$ and $\vec{J} = \vec{L} + \vec{S}$ are taken to transform as

$$\Pi\,\vec{S}\,\Pi^\dagger = +\vec{S}\,, \qquad \Pi\,\vec{J}\,\Pi^\dagger = +\vec{J}\,.$$

In general, an object $\vec{V}$ that transforms under both rotations and parity in the same way as the position operator $\vec{X}$, i.e. $\Pi\,\vec{V}\,\Pi^\dagger = -\vec{V}$, is called a vector. In contrast, an object like the angular momentum operator which rotates like $\vec{X}$ but transforms under parity as $\Pi\,\vec{V}\,\Pi = +\vec{V}$ is called a pseudo-vector. Similarly, an object K that is invariant under both rotations and parity, i.e. $\Pi\,K\,\Pi^\dagger = K$, is called a scalar. However, if it is invariant under rotations but odd under parity transformations, i.e.

$\Pi K \Pi^\dagger = -K$, it is called a pseudo-scalar. An example of a pseudo-scalar in quantum mechanics is $\vec{P} \cdot \vec{S}$.

Parity as a quantum number

Since the parity operator is Hermitian, it is technically an observable. The corresponding eigenstates and eigenvalues are determined by the characteristic equation

$$\Pi|\varphi\rangle = \eta_\varphi|\varphi\rangle\,,$$

with η_φ the so-called *parity of the state* $|\varphi\rangle$, which can take only two values,

$$\Pi^2|\varphi\rangle = \eta_\varphi^2|\varphi\rangle = |\varphi\rangle \quad \Rightarrow \quad \eta_\varphi = \pm 1\,.$$

States with $\eta_\varphi = +1$ are called parity even. States with $\eta_\varphi = -1$ are called parity odd. Note that if the parity operator Π commutes with the Hamiltonian, the energy eigenstates can be also assigned definite parity. This follows immediately when the energy level is non-degenerate. When the energy level is degenerate, it is always possible to pick up a basis within the eigenspace which has definite parity.

Action on spherical harmonics

To understand how parity acts on the spherical harmonics $Y_{l,m}(\theta,\phi)$, let us start by considering the case $l = m$ where these are particularly simple,

$$Y_{l,l}(\theta,\phi) \sim e^{il\phi}\sin^l\theta\,.$$

Taking then into account the action of parity on spherical coordinates,

$$\Pi : (r,\theta,\phi) \mapsto (r,\pi-\theta,\phi+\pi)\,,$$

we get

$$\Pi : Y_{l,l}(\theta,\phi) \mapsto Y_{l,l}(\pi-\theta,\phi+\pi) = e^{il\phi}e^{il\pi}\sin^l(\pi-\theta) = (-1)^l Y_{l,l}(\theta,\phi)\,.$$

To obtain the states with lower values of m, we can use the lowering operator $L_- = L_x - iL_y$. In particular, taking into account that $[\Pi, \vec{L}] = 0$ and therefore $[\Pi, L_-] = 0$, we get

$$\Pi|n,l,l-1\rangle = \Pi L_-|n,l,l\rangle = L_-\Pi|n,l,l\rangle = (-1)^l L_-|n,l,l\rangle = (-1)^l|n,l,l-1\rangle\,.$$

Repeating this argument, we can easily conclude that

$$\Pi|n,l,m\rangle = (-1)^l|n,l,m\rangle\,,$$

for all m values.

[a] This definition is indeed valid for odd dimensions. For even ones, it coincides with a rotation.

7.4.2 The Paschen-Back and Zeeman effects

The so-called Paschen-Back and Zeeman effects are spectral line splittings in atomic systems appearing in the presence of external magnetic fields. Pieter Zeeman received a Nobel prize for their discovery in 1896. As shown in Chapter 3, the interaction of a magnetic field $\vec{B}$ with a hydrogen atom is effectively described by the Pauli Hamiltonian

$$H = \frac{1}{2m}(\vec{P} + e\vec{A})^2 + \frac{ge}{2m}\vec{B}\cdot\vec{S} - \frac{e^2}{r}\,, \tag{7.111}$$

with $\vec{A}$ the vector potential, $\vec{B} = \nabla\times\vec{A}$ and $g = 2$. Expanding the first term in this expression and reordering the different pieces, we get

$$H = \frac{\vec{P}^2}{2m} - \frac{e^2}{r} + \frac{ge}{2m}\vec{B}\cdot\vec{S} + \frac{e}{2m}(\vec{P}\cdot\vec{A} + \vec{A}\cdot\vec{P}) + \frac{q^2}{2m}\vec{A}^2\,. \tag{7.112}$$

While the commutator of $\vec{P}$ and $\vec{A}$ is generically different from zero,

$$\vec{P}\cdot\vec{A} - \vec{A}\cdot\vec{P} = -i\hbar\vec{\nabla}\cdot\vec{A} + i\hbar\vec{A}\cdot\vec{\nabla} = -i\hbar(\vec{\nabla}\cdot\vec{A}) - i\hbar\vec{A}\cdot\vec{\nabla} + i\hbar\vec{A}\cdot\vec{\nabla} = -i\hbar(\vec{\nabla}\cdot\vec{A})\,, \tag{7.113}$$

it vanishes in the symmetric or Landau gauge $\vec{A} = \frac{1}{2}\vec{B}\times\vec{X}$, where

$$\vec{\nabla}\cdot\vec{A} = \frac{1}{2}\vec{\nabla}\cdot(\vec{B}\times\vec{X}) = \frac{1}{2}\epsilon_{ijk}\underbrace{\partial_i(B_jX_k)}_{\partial_iB_j=0} = \frac{1}{2}\epsilon_{ijk}B_j\delta_{ik} = 0\,. \tag{7.114}$$

For the specific case of a magnetic field along the z-direction, we have

$$\vec{A} = \frac{B}{2}(-Y, X, 0)\,, \qquad \vec{P}\cdot\vec{A} = \vec{A}\cdot\vec{P} = \frac{B}{2}\left(XP_y - Yp_x\right) = \frac{1}{2}BL_z\,, \tag{7.115}$$

and therefore

$$H = \frac{\vec{P}^2}{2m} - \frac{e^2}{r} + \frac{e}{2m}\vec{B}\cdot(\vec{L} + 2\vec{S}) + \frac{e^2}{2m}B^2\left(x^2 + y^2\right)\,. \tag{7.116}$$

The terms linear in $\vec{B}$ in this expression define the Zeeman Hamiltonian

$$H_Z = \frac{e}{2m}\vec{B}\cdot(\vec{L} + 2\vec{S})\,. \tag{7.117}$$

The quadratic term in $\vec{B}$ is the so-called diamagnetic term and can be ignored for sufficiently small magnetic fields. In terms of dimensionless quantities, this requires that $B \ll \hbar/(ea_0^2) \sim \mathcal{O}(10)\,T$, with a_0 the Bohr radius. When dealing with the remaining Hamiltonian

$$H = H_0 + H_Z = \frac{\vec{P}^2}{2m} - \frac{e^2}{r} + \frac{e}{2m}\vec{B}\cdot(\vec{L} + 2\vec{S})\,, \tag{7.118}$$

we can distinguish two situations:

- **The Paschen-Back effect**

  For magnetic fields in the range $5\,T \lesssim B \lesssim 10\,T$, the relativistic fine-structure corrections to the hydrogen atom (cf. Chapter 9), and in particular those associated to

the spin-orbit interaction already discussed in Chapter 6, are energetically subdominant as compared to the Zeeman contribution and can be safely ignored. In this situation, and in spite of the many degeneracies involved, the usual eigenstates of H_0, $|n,l,m_l,m_s\rangle$ are also eigenstates of the Zeeman Hamiltonian H_Z, making the matrix representation of (7.118) diagonal in that basis. The associated spectrum of the system at first order in perturbation theory is then given by $E = E_0 + \Delta E_Z^{(1)}$, with

$$\Delta E_Z^{(1)} = \langle n,l,m_l,m_s\,|H_Z|\,n,l,m_l,m_s\rangle = \frac{e\hbar}{2m}\,(m_l+2m_s)\,B\,. \tag{7.119}$$

Note that the two original states with $n=0$, $l=0$, $s=\pm 1/2$ are no longer degenerate. The same applies to other excited states such as $n=2$, $l=1$, $s=\pm 1/2$, which split now into five separated states.

- **The Zeeman effect**

For magnetic fields of strength $B \lesssim 5\,T$, the fine-structure corrections to the hydrogen atom, and in particular those associated to the spin-orbit interaction $H_{\rm SO} \propto \vec{L}\cdot\vec{S}$, become larger than the ones induced by the Zeeman Hamiltonian. In this case, the spin-orbit coupling $\vec{L}\cdot\vec{S}$ locks down effectively the orbital angular momentum and spin, in such a way that only the total angular momentum $\vec{J} = \vec{L} + \vec{S}$ interacts with the magnetic field. Mathematically, this requires us to consider an initial Hamiltonian $\tilde{H}_0 = H_0 + H_{\rm SO}$ to be perturbed by the Zeeman Hamiltonian H_Z. Since the spin-orbit coupling precludes the use of the uncoupled basis, the eigenstates of the system are no longer labeled by $|n,l,m_l,m_s\rangle$, but rather by $|n,j,m_j;l\rangle_C$ with $j = |l\pm 1/2|$. In this coupled basis, the corrections associated to the Zeeman Hamiltonian take the form

$$\Delta E_Z^{(1)} = \frac{eB}{2m}{}_C\langle n,j,m_j;l\,|L_z+2S_z|\,n,j,m_j;l\rangle_C\,. \tag{7.120}$$

Note again that, in spite of the many degeneracies involved, we do not need to make use of degenerate perturbation theory, since both J_z and L^2 commute with H_Z,

$$[J_z,H_Z] = 0\,, \qquad [L^2,H_Z] = 0\,, \tag{7.121}$$

making the more general matrix elements ${}_C\langle n,j,m_j';l'|L_z+2S_z|n,j,m_j;l\rangle_C$ vanish identically unless $l=l'$ and $m_j = m_j'$.

To compute (7.120), let us consider the identity

$$i\hbar\vec{S}\times\vec{L} = (\vec{L}\cdot\vec{S})\vec{S} - \vec{S}(\vec{L}\cdot\vec{S})\,, \tag{7.122}$$

following from the commutation relations $[S_i,S_j] = i\hbar\epsilon_{ijk}S_k$ and $[L_i,S_j]=0$. Taking the cross product of this expression with $\vec{J}$ (remember that $[\vec{L}\cdot\vec{S},\vec{J}]=0$),

$$i\hbar(\vec{S}\times\vec{L})\times\vec{J} = (\vec{L}\cdot\vec{S})\vec{S}\times\vec{J} - \vec{S}\times\vec{J}(\vec{L}\cdot\vec{S})\,, \tag{7.123}$$

and using standard vector identities,

$$(\vec{S}\times\vec{L})\times\vec{J} = \vec{L}(\vec{S}\cdot\vec{J}) - \vec{S}(\vec{L}\cdot\vec{J}) = (\vec{J}-\vec{S})(\vec{S}\cdot\vec{J}) - \vec{S}((\vec{J}-\vec{S})\cdot\vec{J}) = \vec{J}(\vec{S}\cdot\vec{J}) - \vec{S}J^2\,, \tag{7.124}$$

we get

$$(\vec{L}\cdot\vec{S})\vec{S}\times\vec{J} - \vec{S}\times\vec{J}(\vec{L}\cdot\vec{S}) = i\hbar(\vec{J}(\vec{S}\cdot\vec{J}) - \vec{S}J^2)\,. \tag{7.125}$$

Since the spin-orbit interaction $\vec{L}\cdot\vec{S} = 1/2\,(J^2 - L^2 - \vec{S}^2)$ is diagonal in the coupled basis $|n, j, m_j; l\rangle_C$, the expectation value of the left-hand side of this expression vanishes identically for the considered states and so must do the one of the right-hand side. This gives us a relation

$$ {}_C\langle n, j, m_j; l|\vec{S}J^2|n, j, m_j; l\rangle_C = {}_C\langle n, j, m_j; l|\vec{J}(\vec{S}\cdot\vec{J})|n, j, m_j; l\rangle_C\,, \tag{7.126}$$

which, together with $\vec{S}\cdot\vec{J} = 1/2(J^2 + \vec{S}^2 - L^2)$, allows us to relate the expectation value of the spin operator $\vec{S}$ to that of $\vec{J}$,

$$ {}_C\langle n, j, m_j; l|\vec{S}|n, j, m_j; l\rangle = \frac{j(j+1) + s(s+1) - l(l+1)}{2j(j+1)}\,{}_C\langle n, j, m_j; l|\vec{J}|n, j, m_j; l\rangle_C\,. \tag{7.127}$$

Combining this result with $\vec{L} = \vec{J} - \vec{S}$, the first order correction to the energy levels in (7.120) becomes

$$ \Delta E_Z^{(1)} = \frac{eB}{2m}\left(m_j\hbar + \langle n, j, m_j; l\,|S_z|\,n, j, m_j; l\rangle\right) = \frac{eg_J}{2m}\,\hbar m_j B\,, \tag{7.128}$$

with

$$ g_J = 1 + \frac{j(j+1) + s(s+1) - l(l+1)}{2j(j+1)}\,. \tag{7.129}$$

In the presence of the spin-orbit interaction, the magnetic field couples effectively to a magnetic dipole moment $\vec{\mu}_J = eg_J/(2m)\vec{J}$!.

Seeing magnetic fields from the distance

The Zeeman splitting is one of the most direct and reliable methods for measuring astrophysical magnetic fields, particularly in environments where the fields are either intrinsically strong or the sources are sufficiently nearby to provide detectable signals. It has been most effectively applied to the Sun, where high-resolution observations reveal intricate magnetic structures from diffuse quiet regions to intense fields in sunspots. In nearby stars, the Zeeman splitting allows measurements of magnetic activity across various stellar types, offering insights into stellar dynamos and their evolution. In molecular clouds, the technique is used to probe the role of magnetic fields in star formation, revealing how they influence the dynamics of collapsing gas and the formation of protostars.

7.5 The Wentzel-Kramers-Brillouin (WKB) approximation

Consider the time-independent Schrödinger's equation for a particle moving in one dimension under the influence of a potential $V(x)$,

$$ -\frac{\hbar^2}{2m}\frac{d^2\varphi}{dx^2} + V(x)\varphi = E\varphi\,, \quad \longrightarrow \quad \frac{d^2\varphi}{dx^2} + \frac{2m}{\hbar^2}(E - V(x))\varphi = 0\,. \tag{7.130}$$

In regions where this potential varies slowly as compared to the de Broglie's wavelength $\lambda = 2\pi\hbar/p$,[5] we can approximate the wave function by that of a free particle with varying phase,

$$\varphi(x) = e^{iW(x)/\hbar}\,. \tag{7.131}$$

To determine $W(x)$, we substitute this ansatz into the Schrödinger equation and identify the local "classical" momentum

$$p(x) = \sqrt{2m(E - V(x))}\,, \tag{7.132}$$

obtaining with it a differential equation

$$i\hbar\frac{d^2W}{dx^2} - \left(\frac{dW}{dx}\right)^2 + p(x)^2 = 0\,. \tag{7.133}$$

The nonlinearity of this expression is a natural consequence of the exponential mapping (7.131). Note also that, due to the presence of the imaginary unit i, the phase W cannot be real. Among the solutions of (7.133), we will be mainly interested in those involving a negligible second derivative of $W(x)$, namely

$$\hbar\left|\frac{d^2W}{dx^2}\right| \ll \left|\frac{dW}{dx}\right|^2\,. \tag{7.134}$$

This *adiabaticity requirement* is technically equivalent to consider the formal limit $\hbar \to 0$, suggesting itself the use of $\hbar$ as a perturbative expansion parameter. Expanding $W(x)$ in powers of $\hbar$,

$$W(x) = W_0(x) + \hbar W_1(x) + \hbar^2 W_2(x) + \dots\,, \tag{7.135}$$

and substituting the result into Eq. (7.133), we get an expression

$$\left[-W_0'(x)^2 + p(x)^2\right] + \hbar\left[iW_0''(x) - 2W_0'(x)W_1'(x)\right] + \mathcal{O}\left(\hbar^2\right) = 0 \tag{7.136}$$

that can be solved order by order in $\hbar$. At the leading orders $\mathcal{O}(\hbar^0)$ and $\mathcal{O}(\hbar^1)$, we have, respectively,

$$W_0'(x) = \pm p(x) \qquad \longrightarrow \qquad W_0(x) = \pm\int^x dx' p(x')\,, \tag{7.137}$$

$$W_1'(x) = \frac{i}{2}\frac{W_0''(x)}{W_0'(x)} = \frac{i}{2}\frac{p'(x)}{p(x)} \qquad \longrightarrow \qquad W_1(x) = \frac{i}{2}\log p(x) + c\,, \tag{7.138}$$

with c an integration constant to be adjusted. Plugging these expressions into Eq. (7.131), we obtain finally the first-order WKB wave function[6]

$$\varphi(x) \approx \frac{A}{\sqrt{p(x)}}\exp\left(\pm\frac{i}{\hbar}\int^x dx' p(x')\right)\,, \tag{7.139}$$

[5] Note that this condition is always satisfied for classical particles.

[6] Since the region of validity of this expression is still unclear, we will not attempt to normalize it.

with the $\pm$ signs corresponding to waves moving to the right and to the left, respectively. According to this formula, the probability density $|\varphi(x)|^2$ of finding a particle at the position x is proportional to $1/p(x)$ or, equivalently, to the time the particle spends in the vicinity of x. Note also that the momentum $p(x)$ becomes purely imaginary in regions where $E < V(x)$. When that happens, the WKB wave function (7.139) displays either an exponential decay or an exponential growth,

$$\varphi(x) \approx \frac{A}{\sqrt{|p(x)|}} \exp\left(\pm \frac{1}{\hbar} \int^x dx' |p(x')|\right) , \tag{7.140}$$

with the $\pm$ signs set typically by normalisability requirements.

7.5.1 Validity of the WKB approximation

To understand the validity of our approximations, let us we reconsider the adiabaticity condition (7.134). At leading order in the WKB approximation, this can be written as

$$\hbar \left|\frac{dp(x)}{dx}\right| \ll |p(x)|^2 \quad \to \quad \lambda \left|\frac{dp}{dx}\right| \ll |p| \,, \tag{7.141}$$

making explicit that the changes in the local momentum over a distance equal to the de Broglie wavelength $\lambda(x) = 2\pi\hbar/p(x)$ must be small as compared to the local classical momentum. This expression can be alternatively written as

$$\left|\lambda \frac{d\lambda}{dx}\right| \ll \lambda \,, \tag{7.142}$$

showing clearly that the variation of the de Broglie wavelength λ over a distance λ must be much smaller than λ. In terms of the potential, this requires the change over a distance equal to the de Broglie wavelength to be much smaller than the kinetic energy, namely

$$\lambda(x) \left|\frac{dV}{dx}\right| \ll \frac{p^2}{2m} \,. \tag{7.143}$$

7.5.2 Matching conditions

The WKB approximation becomes invalid in the vicinity of classical turning points x_0 where $E = V(x_0)$, $p(x_0) = 0$ and the wave functions (7.139) and (7.140) diverge. To deal with this situation, let us expand the potential in the vicinity of x_0

$$V(x) \approx E + V'(x_0)\,(x - x_0) + \dots \tag{7.144}$$

This leads to a Schrödinger equation

$$-\frac{\hbar^2}{2m}\frac{d^2\varphi}{dx^2} + V'(x_0)\,(x - x_0)\,\varphi = 0 \,, \tag{7.145}$$

which can be trivially rewritten as an Airy equation

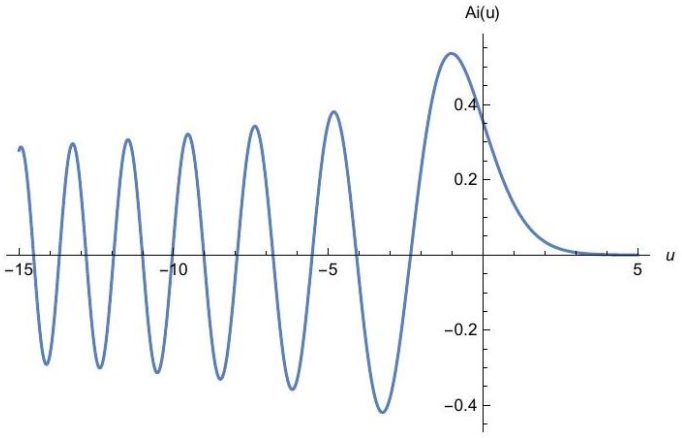
$$\frac{d^2\varphi}{du^2} - u\varphi = 0 \,, \tag{7.146}$$

for the rescaled variable

$$u = \left(\frac{2mV'(x_0)}{\hbar^2}\right)^{1/3}(x - x_0)\,. \tag{7.147}$$

The solution of this seminal differential equation is given by the linear combination of two independent Airy functions $\mathrm{Ai}(u)$ and $\mathrm{Bi}(u)$, being the latter divergent as $u \to \infty$ or equivalently as $x \to \infty$. Discarding $\mathrm{Bi}(u)$ on physical grounds, we are left then with a solution

$$\varphi(u) = \mathrm{Ai}(u) = \frac{1}{\pi}\int_0^\infty dt \cos\left(\frac{t^3}{3} + ut\right)\,. \tag{7.148}$$



The asymptotic behavior of this expression at positive and negative u values is given by

$$\mathrm{Ai}(u) \sim \frac{1}{2}\left(\frac{1}{\pi\sqrt{u}}\right)^{1/2}\exp\left(-\frac{2}{3}u^{3/2}\right)\,, \qquad u \gg 0\,, \tag{7.149}$$

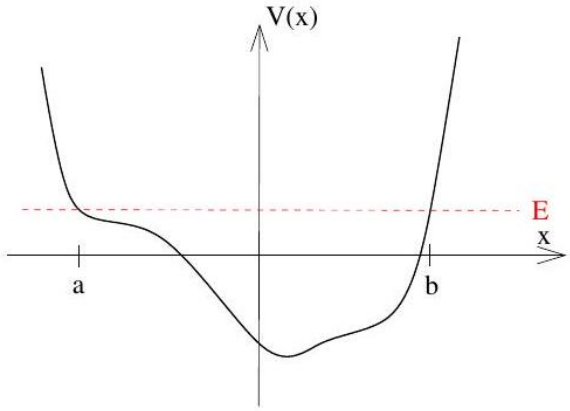
$$\mathrm{Ai}(u) \sim \left(\frac{1}{\pi\sqrt{-u}}\right)^{1/2}\cos\left(\frac{2}{3}u\sqrt{-u} + \frac{\pi}{4}\right)\,, \qquad u \ll 0\,, \tag{7.150}$$

As physically expected, $\mathrm{Ai}(u)$ oscillates for $u < 0$, but decays exponentially for $u > 0$. In terms of the original variables, we have

$$\mathrm{Ai}(x) \sim \begin{cases} \left(\frac{(2mV'(x_0)\hbar)^{1/3}}{\pi\sqrt{p(x)}}\right)^{1/2}\cos\left(\frac{1}{\hbar}\int_{x_0}^x dx'\sqrt{p(x)} + \frac{\pi}{4}\right)\,, & E \gg V(x)\,, \\ \frac{1}{2}\left(\frac{(2mV'(x_0)\hbar)^{1/3}}{\pi\sqrt{|p(x)|}}\right)^{1/2}\exp\left(-\frac{1}{\hbar}\int_{x_0}^x dx'|p(x')|\right)\,, & E \ll V(x)\,. \end{cases} \tag{7.151}$$

7.5.3 Spectrum of bound states

To illustrate the process of matching, let us assume $V(x)$ to be a monotonically growing potential for positive and negative x values, like the one in the figure.



Given the turning points $x_{\min} = a$ and $x_{\max} = b$, we can distinguish three regions where the WKB approximation can be trusted,

$$\text{Region 1: } x \ll a\,, \qquad \text{Region 2: } a \ll x \ll b\,, \qquad \text{Region 3: } x \gg b\,.$$

Deep into Region 1, the solution dies exponentially as

$$\varphi_1(x) \approx \frac{A}{\sqrt{|p(x)|}} \exp\left(-\int_x^a dx' \sqrt{|p(x')|}\right) . \tag{7.152}$$

However, as we approach the turning point at $x = a$, the potential becomes approximately linear, forcing $\varphi_1(x)$ to approach the exponential limit of the Airy function $\mathrm{Ai}(-u)$. Continuing this solution to Region 2, we get

$$\varphi_2(x) \approx \frac{2A}{\sqrt{|p(x)|}} \cos\left(\frac{1}{\hbar}\int_a^x dx'\, p(x') - \frac{\pi}{4}\right) , \tag{7.153}$$

with the minus sign in the phase shift $-\pi/4$ due to the fact that we are working with $\mathrm{Ai}(-u)$ instead of $\mathrm{Ai}(u)$.

The same procedure can be repeated for Regions 2 and 3. As before, deep into Region 3, the solution dies exponentially,

$$\varphi_3(x) \approx \frac{A'}{\sqrt{|p(x)|}} \exp\left(-\int_b^x dx' \sqrt{|p(x')|}\right) . \tag{7.154}$$

Matching this expression to the Airy function across the turning point $x = b$, we obtain

$$\varphi_2(x) \approx \frac{2A'}{|p(x)|} \cos\left(\frac{1}{\hbar}\int_b^x dx' \sqrt{p(x')} + \frac{\pi}{4}\right) . \tag{7.155}$$

We are left therefore with two expressions (7.153) and (7.155), which, being both valid in Region 2, must necessarily agree. Requiring $|A| = |A'|$ and forcing the two cos functions to agree, up to sign, leads to the condition

$$\frac{1}{\hbar}\int_a^x dx' \sqrt{p(x)} - \frac{\pi}{4} = \frac{1}{\hbar}\int_b^x dx' \sqrt{p(x)} + \frac{\pi}{4} + n\pi\,, \tag{7.156}$$

or more compactly

$$\int_{x_{\min}}^{x_{\max}} dx' p\left(x'\right) = \left(n + \frac{1}{2}\right) \hbar\pi \,, \tag{7.157}$$

with n an integer. This condition quantizes the energy levels.

Example 44: Harmonic oscillator spectrum

For the harmonic oscillator $V(x) = m^2\omega^2 x^2$, the quantization condition (7.157) provides precisely the exact spectrum,

$$\int_{x_{\min}}^{x^{\max}} dx \sqrt{2m\left(E - m^2\omega^2 x^2\right)} = \frac{2mE}{m\omega}\frac{\pi}{2} = \left(n + \frac{1}{2}\right)\hbar\pi \quad \rightarrow \quad E = \left(n + \frac{1}{2}\right)\hbar\omega \,. \tag{7.158}$$

Note, however, that this is just a coincidence. In general, we will not get the exact result.

7.5.4 Tunneling probabilities- Gammow's model for α decay

The WKB approximation allows also to estimate tunneling probabilities. Consider for instance the toy version of the seminal Gammow's model for α decay,

$$V(x) = \begin{cases} -V_0 & x < R \,, \\ +\frac{\alpha}{x} & x > R \,, \end{cases} \tag{7.159}$$

with R parametrizing the size of the nucleus, V_0 encoding the nuclear binding energy and $\alpha = Zqe$, with $q = 2e$ the charge of the alpha particle and Ze the charge of the nucleus that remains after the decay. States of negative energy in this potential are bound forever. On the other hand, states with $0 < E < \alpha/R$ can tunnel through the barrier, escaping to infinity. Matching the solution before the tunneling ($x < R$),

$$\varphi_{\text{inside}}(x) = Ae^{ikx} \,, \qquad E = \frac{\hbar^2 k^2}{2m} \,, \tag{7.160}$$

to the exponentially decaying WKB solution far away from the turning point $x = x_\star = \alpha/E$ at which the particle emerges, we obtain

$$\varphi(x_\star) = \varphi(R) e^{-S/\hbar} \,, \tag{7.161}$$

with

$$S = \int_R^{x_\star} dx' \sqrt{2m\left(\frac{\alpha}{x'} - E\right)} \tag{7.162}$$

extending in between the two turning points and capturing information about the height and the width of the barrier. Evaluating this integral in the long-distance, $x_* \gg R$, or narrow-well approximation, $R \to 0$, we get

$$S = \sqrt{\frac{2m}{E}}\pi\alpha \,. \tag{7.163}$$

The associated transmission probability is then given by

$$T = \frac{|\varphi(x_\star)|^2}{|\varphi(R)|^2} = e^{-2S/\hbar}\,. \tag{7.164}$$

Since $W \propto \alpha \propto Zq$, the larger the charge of the nucleus, the less likely the decay.

7.6 Some worked-out exercises

A forced harmonic oscillator

Consider a particle of mass m subjected to a one-dimensional potential

$$V(x) = \frac{1}{2}m\omega^2 x^2 + bx\,,$$

with b a positive constant.

a) Calculate an upper bound for the ground state energy using the trial function family $\varphi(x;\alpha) = Ae^{-\alpha x^2}$, with parameter $\alpha > 0$.

b) Compare the result with the exact solution. Argue how the trial function family should be modified for the variational method to yield the exact ground state energy.

a) We start by determining the normalization constant A,

$$\int_{-\infty}^{\infty} |\varphi(x;\alpha)|^2\,dx = A^2\int_{-\infty}^{\infty} e^{-2\alpha x^2}\,dx = A^2\sqrt{\frac{\pi}{2\alpha}} \quad\longrightarrow\quad A = \left(\frac{2\alpha}{\pi}\right)^{1/4}.$$

The expectation value of the Hamiltonian for the given ansatz contains a kinetic contribution and a potential contribution, namely

$$E_{\rm kin}(\alpha) = -\frac{\hbar^2}{2m}\int_{-\infty}^{\infty}\varphi(x;\alpha)\frac{d^2}{dx^2}\varphi(x;\alpha)\,dx = -\frac{A^2\hbar^2}{2m}\int_{-\infty}^{\infty}\left[(2\alpha x)^2 - 2\alpha\right]e^{-2\alpha x^2}\,dx = \frac{\hbar^2\alpha}{2m}\,.$$

$$E_{\rm pot}(\alpha) = \int_{-\infty}^{\infty}|\varphi(x;\alpha)|^2\left(\frac{m\omega^2}{2}x^2 + bx\right)dx = A^2\int_{-\infty}^{\infty}\left(\frac{m\omega^2}{2}x^2 + bx\right)e^{-2\alpha x^2}\,dx = \frac{m\omega^2}{8\alpha}\,,$$

where in the last step we have taken into account that the integral of the odd function $x\,e^{-2\alpha x^2}$ is zero by symmetry considerations. Minimizing the resulting expression for the total energy

$$E(\alpha) = \frac{\hbar^2\alpha}{2m} + \frac{m\omega^2}{8\alpha}$$

with respect to α,

$$\left.\frac{dE(\alpha)}{d\alpha}\right|_{\alpha=\alpha_*} = \frac{\hbar^2}{2m} - \frac{m\omega^2}{8\alpha_*^2} = 0\,, \quad\longrightarrow\quad \alpha_* = \frac{m\omega}{2\hbar}\,,$$

and plugging the optimized value α_* into the original expression, we get

$$E_{\rm min} = E(\alpha_*) = \frac{\hbar\omega}{4} + \frac{\hbar\omega}{4} = \frac{\hbar\omega}{2} \geq E_0\,,$$

with E_0 the true ground state energy.

b) Completing the square in the potential, we can rewrite it as

$$V = \frac{m\omega^2}{2}\left(x + \frac{b}{m\omega^2}\right)^2 - \frac{b^2}{2m\omega^2}\,.$$

Since a shift in the position operator x does not modify the commutation relations for x and p, the energy of the ground state in this quadratic potential will be that of an harmonic oscillator of frequency ω shifted down by an amount $\frac{b^2}{2m\omega^2}$, i.e.

$$E_0 = \frac{\hbar\omega}{2} - \frac{b^2}{2m\omega^2}\,.$$

Note that for $b \neq 0$, this is lower than the value $E_{\min}$ found in the previous section. The reason is that the potential is centered at $-\frac{b}{m\omega^2}$ and, therefore, the exact ground state is that of the harmonic oscillator displaced in x by an amount $\frac{b}{m\omega^2}$, namely

$$\varphi_0(x) = Ae^{-\alpha\left(x+\frac{b}{m\omega^2}\right)^2}\,,$$

which does not belong to the chosen trial function family. To obtain the exact result, the family should contain this function, which occurs if, for example, we take the family of functions depending on two parameters α and x_0,

$$\varphi(x;\alpha,x_0) = Ae^{-\alpha(x-x_0)^2}$$

Another possibility is to take a family of Gaussian functions with the same symmetry as the potential V, that is, centered at $-\frac{b}{m\omega^2}$,

$$\varphi(x;\alpha) = Ae^{-\alpha\left(x+\frac{b}{m\omega^2}\right)^2}\,.$$

In both cases, the variational method would yield the exact result.

The one-sided harmonic oscillator

Consider a particle of mass m in a one-dimensional half-harmonic oscillator potential

$$V(x) = \begin{cases} \frac{1}{2}m\omega^2x^2 & \text{for} \quad x > 0\,, \\ \infty & \text{for} \quad x \leq 0\,. \end{cases}$$

To estimate the ground state energy of the system using the variational method you are given four different families of trial wave functions at $x \geq 0$,

$$\varphi_1(x,\alpha) = e^{-\alpha x^2}\,, \qquad \varphi_2(x,\beta) = x\,e^{-\beta x^2}\,,$$
$$\varphi_3(x,\gamma,\delta) = \sin(\gamma x)\,e^{\delta x^2}\,, \qquad \varphi_4(x,\zeta,\eta) = \cos(\zeta x)\,e^{-\eta x^2}\,,$$

with α, β, γ, δ, ζ and η *real and positive* variational parameters. All these functions are taken to vanish identically for $x < 0$.

a) Which of the given options are not suitable for the requested task? Justify the reasoning.
b) Estimate the ground-state energy of the system using the remaining viable trial function(s).

Hint: All needed integrals can be trivially computed by deriving $\int_{-\infty}^{\infty} dx\, e^{-ax^2} = \sqrt{\frac{\pi}{a}}$ *with respect to* a.

a)
- $\varphi_1(x,\alpha) = e^{-\alpha x^2}$: This function does not satisfy the boundary condition at $x = 0$, as in $\varphi_1(0,\alpha) = 1$. Thus, this test function is not suitable for the task.

- $\varphi_2(x,\beta) = x\,e^{-\beta x^2}$: This function vanishes at $x = 0$ and behaves well as $x \to \infty$. It also has no additional nodes, making it a viable candidate for the ground state.
- $\varphi_3(x,\gamma,\delta) = \sin(\gamma x)\,e^{\delta x^2}$: Although this function vanishes at $x = 0$, the sine function introduces additional nodes depending on the value of γ, which is not characteristic of the ground state (node theorem). On top of that, the presence of $e^{+\delta x^2}$ leads to a non-physical, unbounded behavior [a physical wave function must be normalizable]. Therefore, this trial function is not suitable for estimating the ground-state energy.
- $\varphi_4(x,\zeta,\eta) = \cos(\zeta x)\,e^{-\eta x^2}$: This function does not vanish at $x = 0$ unless ζ is chosen in such a way that $\cos(\zeta \cdot 0) = 0$, which would not generally be the case for a ground-state approximation. Thus, this test function is not suitable for the task.

b) The remaining trial function $\varphi_2(x,\beta) = xe^{-\beta x^2}$ vanishes at $x = 0$, while remaining finite as $x \longrightarrow +\infty$. By trivially computing the needed integrals by parametric differentiation of the given Gaussian integral,

$$\int_{-\infty}^{\infty} dx\, x^2 e^{-2\beta x^2} = -\frac{1}{2}\frac{\partial}{\partial\beta}\int_{-\infty}^{\infty} dx\, e^{-2\beta x^2} = \frac{1}{4}\sqrt{\frac{\pi}{2\beta^3}} = \frac{1}{4\beta}\sqrt{\frac{\pi}{2\beta}}\,,$$

$$\int_{-\infty}^{\infty} x^4 e^{-2\beta x^2} dx = \frac{1}{4}\frac{\partial^2}{\partial\beta^2}\int_{-\infty}^{\infty} dx\, e^{-2\beta x^2} = \frac{3}{4}\sqrt{\frac{\pi}{(2\beta)^5}} = \frac{3}{16\beta^2}\sqrt{\frac{\pi}{2\beta}}\,,$$

we have

$$\langle\varphi_2|\varphi_2\rangle = \int_0^{+\infty} x^2 e^{-2\beta x^2} dx = \frac{1}{2}\int_{-\infty}^{\infty} dx\, x^2 e^{-2\beta x^2} = \frac{1}{8\beta}\sqrt{\frac{\pi}{2\beta}}\,,$$

$$\left\langle\varphi_2\left|\frac{1}{2}m\omega^2 x^2\right|\varphi_2\right\rangle = \frac{1}{2}m\omega^2\int_0^{+\infty} x^4 e^{-2\beta x^2} dx = \frac{1}{4}m\omega^2\int_{-\infty}^{+\infty} x^4 e^{-2\beta x^2} dx = \frac{3m\omega^2}{64\beta^2}\sqrt{\frac{\pi}{2\beta}}\,,$$

$$\left\langle\varphi_2\left|-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\right|\varphi_2\right\rangle = \frac{\hbar^2}{2m}\int_0^{+\infty}\left(3\beta x^2 - 2\beta^2 x^4\right) e^{-2\beta x^2} dx = \frac{3\hbar^2}{16m}\sqrt{\frac{\pi}{2\beta}}\,,$$

leading to the ground state energies

$$E_0(\beta) = \frac{\langle\varphi_2(\beta)|H|\varphi_2(\beta)\rangle}{\langle\varphi_2|\varphi_2\rangle} = \frac{3\hbar^2}{2m}\beta + \frac{3m\omega^2}{8\beta}\,.$$

The minimization of $E_0(\beta)$ with respect to β yields $\beta_0 = m\omega/(2\hbar)$ and hence

$$E_0\left(\beta_0\right) = \frac{3}{2}\hbar\omega\,.$$

As compared to the exact ground state energy of a full harmonic oscillator, the presence of the infinite barrier at $x = 0$ increases the estimated ground state energy by an additional $\hbar\omega$, which reflects the fact that the particle is now confined to half the space, effectively increasing the confinement and thus the energy.

Variational method versus perturbative theory

A given quantum system is described by a Hamiltonian $H = H^{(0)} + \mu H^{(1)}$ with $0 \le \mu \ll 1$. In order to find its ground state energy, two methods are employed:

1. A first-order perturbative calculation yields the result $E_{\text{pert}} = (94 - 5\mu)$ eV.
2. A variational method applied to a family of states depending on a parameter $c \in \mathbb{R}$ yields the result $E_{\text{var}}(c) = \left(c^2 - 2c\mu + 100\right)$ eV.

a) Discuss whether the exact ground state energy of the unperturbed Hamiltonian $H^{(0)}$ is greater or smaller than 90 eV.
b) Determine which of the two methods provides a closer approximation to the true ground state energy of the Hamiltonian H *in their common range of validity.*
c) If one were to perform a second order perturbative computation, would the associated correction raise or lower the one obtained at first order in perturbation theory?

a) From the perturbative result, when $\mu = 0$, the ground state energy is $E_0 = 94$ eV. This is the energy of the unperturbed Hamiltonian $H^{(0)}$. Therefore, the exact ground state energy of the unperturbed Hamiltonian $H^{(0)}$ is greater than 90 eV.

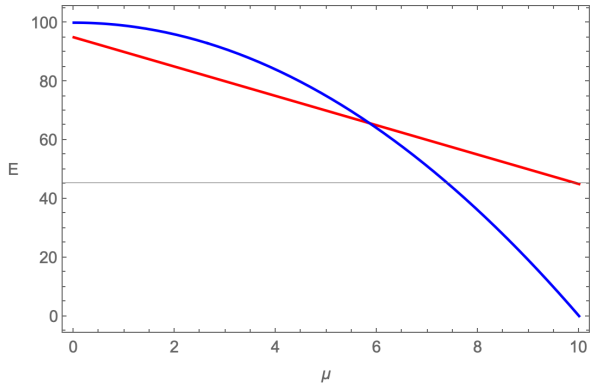
b) To find the best variational approximation, we minimize $E_{\rm var}(c)$ with respect to c,

$$\frac{dE_{\rm var}(c)}{dc} = 2c - 2\mu = 0 \qquad \longrightarrow \qquad c = \mu\,.$$

Substituting this result into the variational expression gives

$$E_{\rm var} = (100 - \mu^2)\ \text{eV}\,.$$

To determine which of the two methods provides the closer approximation to the true ground state energy *in their common range of validity*, we compare the results $E_{\rm pert} = (94 - 5\mu)$ eV (in red) and $E_{\rm var} = (100 - \mu^2)$ eV (in blue) as a function of μ.



We observe that for $0 \leq \mu < 6$, and, in particular, for the *common* range of validity of the two methods, i.e. $0 \leq \mu \ll 1$, the perturbative result provides a smaller energy than the variational one, $E_{\rm pert} < E_{\rm var}$. Combining this result with the fact that the energy of the (non-degenerate) ground state at first order in perturbation theory exceeds generically the true ground state energy, we conclude that the perturbative method provides a better approximation *in the common range of validity.*

c) The energies at second order in perturbation theory are given by

$$E_n^{(2)} = E_{\rm pert} + \sum_{m\neq n} \frac{|\langle m^{(0)}|\Delta H|n^{(0)}\rangle|^2}{E_n^{(0)} - E_m^{(0)}}\,.$$

Since the unperturbed excited states exceed by definition the unperturbed ground state energy, $E_m^{(0)} > E_0^{(0)}$, the second-order correction to the ground state is always negative. Therefore, the resulting second-order energy estimate is generically smaller than the one at first order.

A particle in the earth gravitational field

Consider a quantum particle of mass m in a potential

$$V(z) = \begin{cases} mgz & \text{if } z > 0\,, \\ \infty & \text{if } z \leq 0\,, \end{cases}$$

with g the gravitational constant. Estimate the ground state energy of the system using the variational method with a trial function properly matching the boundary conditions of the problem. Compare the results with the exact solution of the problem.

The ground state wave function of this particle has no nodes and must vanish at $z = 0$ and be finite as $z \to +\infty$. A trial function satisfying these conditions is for instance

$$\varphi_0(z, \alpha) = Aze^{-\alpha z}\,,$$

with α a variational parameter and A a normalization constant. A trivial computation fixes the latest quantity to a value $A = 2\alpha^{3/2}$ and hence

$$\varphi_0(z, \alpha) = 2\sqrt{\alpha^3}ze^{-\alpha z}\,.$$

For this antsatz, we have

$$\begin{aligned} E_0(\alpha) &= 4\alpha^3 \int_0^{+\infty} ze^{-\alpha z}\left[-\frac{\hbar^2}{2m}\frac{d^2}{dz^2} + mgz\right] ze^{-\alpha z}dz \\ &= 4\alpha^3\frac{\hbar^2}{2m}\int_0^{+\infty}\left(2\alpha z - \alpha^2 z^2\right)e^{-2\alpha z}dz + 4\alpha^3 mg\int_0^{+\infty} z^3 e^{-2\alpha z}dz \\ &= 2\alpha^3\frac{\hbar^2}{m}\left(\frac{1}{2\alpha} - \frac{1}{4\alpha}\right) + 4mg\alpha^3\frac{3}{8\alpha^4}\,. \end{aligned}$$

or

$$E_0(\alpha) = \frac{\hbar^2}{2m}\alpha^2 + \frac{3}{2\alpha}mg\,,$$

Minimization wrt to α yields

$$\alpha_* = \left(3m^2 g/2\hbar^2\right)^{1/3}$$

and hence

$$E_0\left(\alpha_*\right) = \frac{3}{2}\left(\frac{9}{2}\right)^{1/3}\left(\frac{1}{2}mg^2\hbar^2\right)^{1/3}\,.$$

To compare this with the exact result,we need now to solve the Schrödinger equation with the boundary conditions $\varphi(0) = 0$ and $\varphi(+\infty) = 0$, namely

$$-\frac{\hbar^2}{2m}\frac{d^2\varphi(z)}{dz^2} + mgz\varphi(z) = E\varphi(z) \quad \longrightarrow \quad \frac{d^2\varphi(z)}{dz^2} - \frac{2m}{\hbar^2}(mgz - E)\varphi(z) = 0\,.$$

With the change of variable $x = \left(\hbar^2/\left(2m^2 g\right)\right)^{2/3}\left(2m/\hbar^2\right)(mgz - E)$, we can reduce this equation to

$$\frac{d^2\phi(x)}{dx^2} - x\phi(x) = 0\,,$$

a standard differential equation with well-known solution vanishing at $x \to = +\infty$, i.e., $\phi(+\infty) = 0$,

$$\phi(x) = B\,\mathrm{Ai}(x) \quad \text{where} \quad \mathrm{Ai}(x) = \frac{1}{\pi}\int_0^{\infty}\cos\left(\frac{1}{3}t^3 + xt\right)dt\,,$$

with $\mathrm{Ai}(x)$ the Airy function. When $z = 0$ we have $x = -\left(2/\left(mg^2\hbar^2\right)\right)^{1/3} E$. The boundary condition $\varphi(0) = 0$ yields $\phi\left[-\left(2/\left(mg^2\hbar^2\right)\right)^{1/3} E\right] = 0$ or $\mathrm{Ai}\left[-\left(2/\left(mg^2\hbar^2\right)\right)^{1/3} E\right] = 0$. The Airy function has zeros only at certain values of R_n : $\mathrm{Ai}\left(R_n\right) = 0$ with $n = 0, 1, 2, 3, \ldots$ The roots R_n of the Airy function can be found in standard tables. For instance, the first few roots are $R_0 = -2.338$, $R_1 = -4.088, R_2 = -5.521, R_3 = -6.787$

The boundary condition $\varphi(0) = 0$ therefore gives a discrete set of energy levels which can be expressed in terms of the roots of the Airy function,

$$\mathrm{Ai}\left[-\left(\frac{2}{mg^2\hbar^2}\right)^{1/3} E\right] = 0 \quad \longrightarrow \quad -\left(\frac{2}{mg^2\hbar^2}\right)^{1/3} E_n = R_n \,.$$

hence

$$E_n = -\left(\frac{1}{2}mg^2\hbar^2\right)^{1/3} R_n, \qquad \varphi_n(z) = B_n \mathrm{Ai}\left[-\left(\frac{2m^2g^2}{\hbar^2}\right)^{1/3} z - R_n\right] .$$

The ground state energy is therefore

$$E_0 = 2.338\left(\frac{1}{2}mg^2\hbar^2\right)^{1/3} .$$

For our trial function, the variational method overestimates the energy by a 5.9% only

$$E_0 = \frac{3}{2}\left(\frac{9}{2}\right)^{1/3} \frac{E_0^{\text{exact}}}{2.338} \simeq 1.059 E_0^{\text{exact}} \ .$$

A spin-3/2 particle

The Hamiltonian of a single spin-3/2 particle is given by

$$H = -E_0\left(\frac{1}{\hbar}S_z + \frac{1}{\hbar^2}S_x^2\right) ,$$

with $E_0 > 0$. Estimate the ground state energy of the system using the variational method with a trial family of states

$$|\varphi(\theta)\rangle = \cos\theta\,|3/2, 3/2\rangle + \sin\theta\,|3/2, -1/2\rangle \,,$$

with θ a variational parameter and $|s, m_s\rangle = |3/2, 3/2\rangle, |3/2, -1/2\rangle$ eigentates of S_z.

The proposed state is already normalized. To calculate the required matrix elements, we express S_x

as $S_x = \frac{1}{2}(S_+ + S_-)$, so

$$\begin{aligned}
S_x \left|\frac{3}{2}\frac{1}{2}\right\rangle &= \frac{1}{2}S_- \left|\frac{3}{2}\frac{1}{2}\right\rangle = \frac{\sqrt{3}}{2}\hbar \left|\frac{3}{2}\frac{1}{2}\right\rangle , \\
S_x \left|\frac{3}{2}\frac{1}{2}\right\rangle &= \frac{1}{2}S_+ \left|\frac{3}{2}\frac{1}{2}\right\rangle + \frac{1}{2}S_- \left|\frac{3}{2}\frac{1}{2}\right\rangle \\
&= \frac{\sqrt{3}}{2}\hbar \left|\frac{3}{2}\frac{3}{2}\right\rangle + \frac{\hbar}{2}\sqrt{\frac{3}{2}\cdot\frac{5}{2} - \frac{1}{2}\left(-\frac{1}{2}\right)} \left|\frac{3}{2}-\frac{1}{2}\right\rangle = \frac{\sqrt{3}}{2}\left|\frac{3}{2}\frac{3}{2}\right\rangle + \hbar\left|\frac{3}{2}-\frac{1}{2}\right\rangle , \\
S_x \left|\frac{3}{2}-\frac{1}{2}\right\rangle &= \frac{1}{2}S_+ \left|\frac{3}{2}-\frac{1}{2}\right\rangle + \frac{1}{2}S_- \left|\frac{3}{2}-\frac{1}{2}\right\rangle \\
&= \hbar\left|\frac{3}{2}\frac{1}{2}\right\rangle + \frac{\hbar}{2}\sqrt{\frac{3}{2}\cdot\frac{5}{2} - \left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)} = \hbar\left|\frac{3}{2}\frac{1}{2}\right\rangle + \frac{\sqrt{3}}{2}\hbar\left|\frac{3}{2}-\frac{3}{2}\right\rangle ,
\end{aligned}$$

and

$$\begin{aligned}
&\left\langle\frac{3}{2}\frac{3}{2}\middle|H\middle|\frac{3}{2}\frac{3}{2}\right\rangle = -E_0\left(\frac{3}{2} + \frac{1}{\hbar}\left|S_x\left|\frac{3}{2}\frac{3}{2}\right\rangle\right|^2\right) = -E_0\left(\frac{3}{2}+\frac{3}{4}\right) = -\frac{9}{4}E_0 , \\
&\left\langle\frac{3}{2}-\frac{1}{2}\middle|H\middle|\frac{3}{2}-\frac{1}{2}\right\rangle = -E_0\left(-\frac{1}{2} + \frac{1}{\hbar^2}\left|S_x\left|\frac{3}{2}-\frac{1}{2}\right\rangle\right|^2\right) = -E_0\left(-\frac{1}{2}+1+\frac{3}{4}\right) = -\frac{5}{4}E_0 , \\
&\left\langle\frac{3}{2}-\frac{1}{2}\middle|H\middle|\frac{3}{2}\frac{3}{2}\right\rangle = -E_0\cdot\frac{1}{\hbar^2}\left\langle\frac{3}{2}-\frac{1}{2}\middle|S_x\cdot S_x\middle|m=\frac{3}{2}\right\rangle = -E_a\frac{\sqrt{3}}{2} , \\
&\left\langle\frac{3}{2}\frac{3}{2}\middle|H\middle|\frac{3}{2}-\frac{1}{2}\right\rangle = \left\langle\frac{3}{2}-\frac{1}{2}\middle|H\middle|\frac{3}{2}\frac{3}{2}\right\rangle^\dagger = -E_0\frac{\sqrt{3}}{2} .
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
E_{\text{ground}} &\le \cos^2\theta\left\langle\frac{3}{2}\frac{3}{2}\middle|H\middle|\frac{3}{2}\frac{3}{2}\right\rangle + \sin^2\theta\left\langle\frac{3}{2}-\frac{1}{2}\middle|H\middle|\frac{3}{2}-\frac{1}{2}\right\rangle + 2\sin\theta\cos\theta\left\langle\frac{3}{2}-\frac{1}{2}\middle|H\middle|\frac{3}{2}\frac{3}{2}\right\rangle \\
&= -E_0\left(\frac{9}{4}\cos^2\theta + \frac{5}{4}\sin^2\theta + \sqrt{3}\sin\theta\cos\theta\right) .
\end{aligned}$$

To find the best lower bound, we maximize the expression in brackets, obtaining

$$-\frac{9}{2}\cos\theta\sin\theta + \frac{5}{2}\sin\theta\cos\theta + \sqrt{3}\left(\cos^2\theta - \sin^2\theta\right) = 0 \quad\longrightarrow\quad \tan(2\theta) = \sqrt{3} \quad\longrightarrow\quad \theta = \frac{\pi}{6} \text{ or } \frac{2\pi}{3} .$$

For $\theta = \frac{\pi}{6}$, we have

$$\frac{9}{4}\cos^2\theta + \frac{5}{4}\sin^2\theta + \sqrt{3}\sin\theta\cos\theta = \frac{11}{4} .$$

For $\theta = \frac{2\pi}{3}$, we have

$$\frac{9}{4}\cos^2\theta + \frac{5}{4}\sin^2\theta + \sqrt{3}\sin\theta\cos\theta = \frac{3}{4} ,$$

so the best bound on the ground state energy is

$$E_{\text{ground}} \leqslant -\frac{11}{4}E_0 .$$

A perturbed anharmonic oscillator: perturbation vs. variational methods

The Hamiltonian of a one-dimensional quantum system is given by

$$H = H_0 + \Delta H\,, \qquad H_0 = E_0\, a^\dagger a^\dagger a a\,, \qquad \Delta H = \lambda E_0 \left(a + a^\dagger\right)\,,$$

with a and $a^\dagger$ the creation and annihilation operators of a harmonic oscillator of mass m and frequency ω, E_0 a constant and $\lambda \ll 1$.

a) Show that the usual eigenstates $|n\rangle$ of the harmonic oscillator with $n = 0, 1, 2, \ldots$ are also eigenstates of H_0, determining the corresponding eigenvalues and degeneracies.

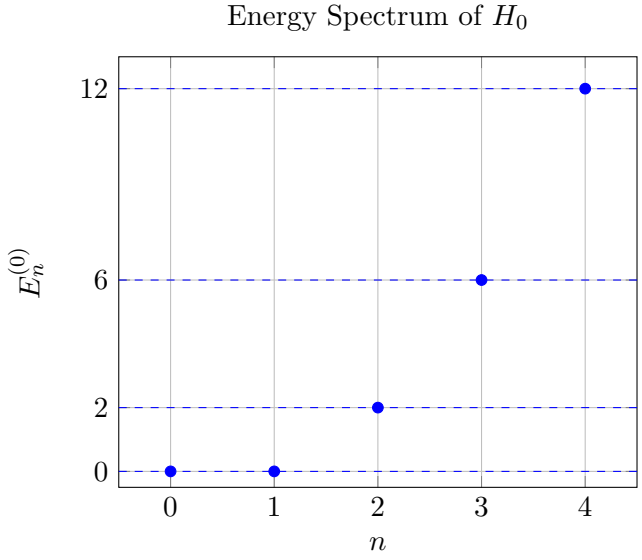
b) Calculate the energy of the lowest energy state of H at first order in perturbation theory.

c) Compare the previous result with that resulting from the use of the variational method with a trial function $|\varphi\rangle = \cos\theta|0\rangle + \sin\theta|1\rangle$, with θ a variational parameter.

a)

$$\begin{aligned} H_0|n\rangle = E_0 a^\dagger a^\dagger a a|n\rangle &= E_0 a^\dagger a^\dagger a\sqrt{n}|n-1\rangle = E_0 a^\dagger a^\dagger \sqrt{n}\cdot\sqrt{n-1}|n-2\rangle \\ &= E_0 a^\dagger \sqrt{n}\cdot\sqrt{n-1}\cdot\sqrt{n-1}|n-1\rangle = E_0\sqrt{n}\cdot\sqrt{n-1}\cdot\sqrt{n-1}\cdot\sqrt{n}|n\rangle \\ &= E_0\cdot n\cdot(n-1)|n\rangle\,. \end{aligned}$$

Therefore, $|n\rangle$ is an energy eigenstate of H_0 with eigenvalue $E_0 \cdot n(n-1)$.



b) ΔH has odd parity and can therefore mix states of different parity. Since the two lowest energy states of the unperturbed Hamiltonian H_0 have both zero energy ($n = 0$ and $n = 1$), we make use of degenerate perturbation theory

$$\Delta H|0\rangle = \lambda E_0 \left(a + a^\dagger\right)|0\rangle = \lambda E_0|1\rangle\,, \qquad \Delta H|1\rangle = \lambda E_0 \left(a + a^\dagger\right)|1\rangle = \lambda E_0(|0\rangle + \sqrt{2}|2\rangle)\,,$$

$$\begin{pmatrix} \langle 0\,|\Delta H|\,0\rangle & \langle 0\,|\Delta H|\,1\rangle \\ \langle 1\,|\Delta H|\,0\rangle & \langle 1\,|\Delta H|\,1\rangle \end{pmatrix} = \lambda E_0 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} .$$

The corresponding eigvenvalues and eigenstates are $\pm E_0$ and

$$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}, \qquad \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}.$$

The lowest energy at first order in perturbation theory is therefore $-\lambda E_0$.

c) The energy expectation value for the trial wave function is $E(\theta) = \langle\varphi|H|\varphi\rangle = \langle\varphi|H_0 + \Delta H|\varphi\rangle$. However, since H_0 gives zero energy for both $|0\rangle$ and $|1\rangle$, $\langle\varphi|H_0|\varphi\rangle = 0$ and the expectation value reduces to

$$E(\theta) = \langle\varphi|\Delta H|\varphi\rangle = \cos\theta\sin\theta\,(\langle 0|\Delta H|1\rangle + \langle 1|\Delta H|0\rangle) = 2\lambda E_0\cos\theta\sin\theta = \lambda E_0\sin(2\theta)\,,$$

where we have taken into account the previously computed matrix element $\langle 1|\Delta H|0\rangle = \lambda E_0$. The minimum value of $\sin(2\theta)$ is -1, which occurs at $2\theta = \frac{3\pi}{2}$, or $\theta = \frac{3\pi}{4}$. Thus, the minimum energy estimate obtained through this method is also $-\lambda E_0$.

A perturbed 2D harmonic oscillator

The Hamiltonian of a particle of mass m oscillating harmonically in two dimensions is given by

$$H_0 = \frac{1}{2m}\left(p_x^2 + p_y^2\right) + \frac{1}{2}m\omega^2\left(x^2 + y^2\right).$$

Which is the energy and degeneracy of the first excited level? Using first-order perturbation theory, calculate the energy of this state after introducing a perturbation $\Delta H = \alpha\, xy$, with α a constant.

The eigenvalues and eigenstates of H_0 are given by

$$E_n^{(0)} = \hbar\omega(n_x + n_y + 1)\,, \qquad |n^{(0)}\rangle = |n_x\rangle \otimes |n_y\rangle = |n_x n_y\rangle\,,$$

with $n_{x,y} = 0, 1, 2, \ldots$ and $|n_x\rangle$ and $|n_y\rangle$ the eigenstates of the one-dimensional harmonic oscillator for each of the two directions. The first excited level corresponds to $n_x + n_y = 1$, which can be obtained in two possible ways, $n_x = 1, n_y = 0$ or $n_x = 0, n_y = 1$. The first excited level is therefore 2-fold degenerate. In an abbreviated notation, we can label the corresponding states as $|10\rangle$ and $|01\rangle$. To calculate the splitting induced by the perturbation, we evaluate the matrix elements of ΔH between the above degenerate states, namely

$$\begin{pmatrix} \langle 10|\Delta H|10\rangle & \langle 10|\Delta H|01\rangle \\ \langle 01|\Delta H|10\rangle & \langle 01|\Delta H|01\rangle \end{pmatrix} = \alpha \begin{pmatrix} \langle 10|xy|10\rangle & \langle 10|xy|01\rangle \\ \langle 01|xy|10\rangle & \langle 01|xy|01\rangle \end{pmatrix}.$$

Expressing x and y in terms of raising and lowering operators

$$x = \sqrt{\frac{\hbar}{2m\omega}}\left(a_x + a_x^\dagger\right), \qquad y = \sqrt{\frac{\hbar}{2m\omega}}\left(a_y + a_y^\dagger\right),$$

we get

$$xy = \frac{\hbar}{2m\omega}\left(a_x a_y + a_x a_y^\dagger + a_x^\dagger a_y + a_x^\dagger a_y^\dagger\right),$$

and

$$\langle 10|xy|01\rangle = \frac{\hbar}{2m\omega}\langle 10|\left(a_x a_y + a_x a_y^\dagger + a_x^\dagger a_y + a_x^\dagger a_y^\dagger\right)|01\rangle = \langle 10|a_x^\dagger a_y|01\rangle = \frac{\hbar}{2m\omega}\,,$$

where in the last step we have used the fact that $a_x|01\rangle = 0$ and invoked the orthogonality properties of the eigenstates. By similar arguments, it is easy to show that the two diagonal terms in the matrix are zero, while $\langle 01|xy|10\rangle = \langle 10|xy|01\rangle$. The 2×2 submatrix of ΔH is then given by

$$\begin{pmatrix} \langle 10|\Delta H|10\rangle & \langle 10|\Delta H|01\rangle \\ \langle 01|\Delta H|10\rangle & \langle 01|\Delta H|01\rangle \end{pmatrix} = \frac{\alpha\hbar}{2m\omega} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

which has eigenvalues

$$\Delta E = \pm \frac{\alpha\hbar}{2m\omega}.$$

A peturbed system of two identical particles of mass

A system of two identical particles of mass m and spin 1 is described, in the center of mass frame, by the Hamiltonian

$$H = \frac{p^2}{2\mu} + \frac{1}{2}\mu\omega^2 r^2 + \frac{\omega}{\hbar}J^2,$$

with $\mu = m/2$ the reduced mass of the system, $\vec{r} = \vec{r}_1 - \vec{r}_2$ the relative coordinate between the particles, $\vec{p}$ the corresponding conjugate momentum and $\vec{J} = \vec{L} + \vec{S}$, where $\vec{L}$ is the total angular momentum, and $\vec{S} = \vec{S}_1 + \vec{S}_2$ is the total spin. Calculate, at first order in perturbation theory, the energy shift of the ground state due to the perturbation

$$\Delta H = \lambda \left[\left(\vec{S}_1 - \vec{S}_2 \right) \cdot \vec{r} \right]^2,$$

with λ a small constant.

The spatial part in H is the Hamiltonian of a 3-dimensional harmonic oscillator. Therefore, one has

$$E = \hbar\omega \left[n_x + n_y + n_z + \frac{3}{2} + j(j+1) \right].$$

Since the two particles have spin 1, one can have $s = 0, 1, 2$. Therefore, the ground state of the system is ($n_x = n_y = n_z = l = 0$, $|n_x\, n_y\, n_z\rangle|s\, s_z\rangle$)

$$E = \frac{3}{2}\hbar\omega, \qquad |\varphi\rangle = |n_x\, n_y\, n_z\rangle|s\, s_z\rangle = |\,0\,0\,0\,\rangle|\,0\,0\,\rangle.$$

Defining for compactness $\vec{S}_{12} = \vec{S}_1 - \vec{S}_2$, we have

$$\begin{aligned} &\langle\varphi|\lambda \left[\left(\vec{S}_1 - \vec{S}_2 \right) \cdot \tilde{r} \right]^2 |\varphi\rangle = \lambda \left\langle \varphi \left| (S_{12x}x + S_{12y}y + S_{12z}z)^2 \right| \varphi \right\rangle \\ &= \lambda \langle\varphi| \left(S_{12x}^2 x^2 + S_{12y}^2 y^2 + S_{12z}^2 z^2 + 2S_{12x}S_{12y}xy + 2S_{12x}S_{12z}xz + 2S_{12y}S_{12z}yz \right) |\varphi\rangle . \end{aligned}$$

Now, since

$$\langle 000|x^2|000\rangle = \langle 0|x^2|0\rangle = \frac{\hbar}{2m\omega}, \quad \langle 000|y^2|000\rangle = \langle 0|y^2|0\rangle = \frac{\hbar}{2m\omega}, \quad \langle 000|z^2|000\rangle = \langle 0|z^2|0\rangle = \frac{\hbar}{2m\omega},$$
$$\langle 000|xy|000\rangle = \langle 0|x|0\rangle\langle 0|y|0\rangle = 0, \quad \langle 000|xz|000\rangle = \langle 0|x|0\rangle\langle 0|z|0\rangle = 0, \quad \langle 000|yz|000\rangle = \langle 0|y|0\rangle\langle 0|z|0\rangle = 0,$$

the above expectation value becomes

$$\langle\varphi|\lambda\left[\left(\vec{S}_1-\vec{S}_2\right)\cdot\vec{r}\right]^2|\varphi\rangle=\lambda\frac{\hbar}{2m\omega}\langle\varphi|\,S_{12x}^2+S_{12y}^2+S_{12z}^2\,|\varphi\rangle=\frac{\lambda\hbar}{2m\omega}\langle\varphi|\,S_{12}^2\,|\varphi\rangle\ .$$

Expanding this expression,

$$S_{12}^2=(S_1-S_2)^2=S_1^2+S_2^2-2\vec{S}_1\cdot\vec{S}_2=S_1^2+S_2^2-\left[S^2-S_1^2-S_2^2\right]=2S_1^2+2S_2^2-S^2\,,$$

and acting on the spin part of $|\varphi\rangle$

$$\begin{aligned}(2S_1^2+2S_2^2-S^2)|s\,s_z\rangle&=\hbar^2\left[2s_1(s_1+1)+2s_2(s_2+1)-s(s+1)\right]|s\,s_z\rangle\,,\\(2S_1^2+2S_2^2-S^2)|0\,0\rangle&=\hbar^2\left(4+4-0\right)|0\,0\rangle=8\hbar^2|0\,0\rangle\,,\end{aligned}$$

we get

$$\langle\varphi|\Delta H|\varphi\rangle=\frac{\lambda\hbar}{2m\omega}\cdot 8\hbar^2=\frac{4\lambda\hbar^3}{m\omega}\,.$$

A perturbed spin-1 system

A spin-1 system is governed by an effective Hamiltonian

$$H=H_0+\Delta H\,,\qquad H_0=\frac{\omega}{\hbar}\left(S^2-S_y^2-\hbar S_y\right)\,,\qquad \Delta H=-\lambda\omega S_z\,,$$

with ω a positive frequency, $0<\lambda\ll 1$ a dimensionless parameter and $\vec{S}=(S_x,S_y,S_z)$, with S_i $(i=x,y,z)$ the spin operators along the three Cartesian axes.

a) Find a suitable set of commuting observables leading to the diagonalization of H_0. Determine the eigenvalues and eigenstates of H_0 in the corresponding basis.

b) Compute the first-order corrections to the ground state in terms of λ. Interpret the result.

a) Since the unperturbed Hamiltonian H_0 depends only on S^2 and S_y and $[S^2,S_y]=0$, we have $[H_0,S^2]=[H_0,S_y]=0$, forming these three operators a CSCO. In the associated basis $|s\,s_y\rangle_y$, the Hamiltonian is diagonal

$$H_0|s\,s_y\rangle_y=\hbar\omega\left(2-s_y^2-s_y\right)|s\,s_y\rangle_y\,.$$

b) To determine the correction to the ground state,

$$\Delta E={}_y\langle 1,1|\Delta H|1,1\rangle_y=-\lambda\omega_y\langle 1,1|S_z|1,1\rangle_y\,,$$

we write down the eigenvalue equation $S_y|1,1\rangle_y=\hbar|1,1\rangle_y$ in terms of the eigenvectors of S_z,

$$|1,1\rangle_z=\begin{pmatrix}1\\0\\0\end{pmatrix}\,,\qquad |1,0\rangle_z=\begin{pmatrix}0\\1\\0\end{pmatrix}\,,\qquad |1,-1\rangle_z=\begin{pmatrix}0\\0\\1\end{pmatrix}\,.$$

namely

$$\frac{1}{\sqrt{2}}\begin{pmatrix}0&-i&0\\i&0&-i\\0&i&0\end{pmatrix}\begin{pmatrix}a\\b\\c\end{pmatrix}=\begin{pmatrix}a\\b\\c\end{pmatrix}\quad\longrightarrow\quad\begin{aligned}-\frac{i}{\sqrt{2}}b&=a\\-\frac{i}{\sqrt{2}}(-a+c)&=b\\+\frac{i}{\sqrt{2}}b&=c\end{aligned}\quad\longrightarrow\quad|1,1\rangle_y=b\left(-\frac{i}{\sqrt{2}},1,\frac{i}{\sqrt{2}}\right)\,.$$

Normalizing the result,

$$ {}_y\langle 1,1|1,1\rangle_y = |b|^2(1/2+1+1/2) = 2|b|^2 = 1 \quad \longrightarrow \quad |b|^2 = 1\,, $$

and choosing $b = i/\sqrt{2}$ to minimize the number of imaginary units, we have $|1,1\rangle_y = \left(\frac{1}{2}, -\frac{i}{\sqrt{2}}, -\frac{1}{2}\right)$, and therefore

$$ \Delta E = -\lambda\omega\, {}_y\langle 1,1|S_z|1,1\rangle_y = -\lambda\hbar\omega \left(\frac{1}{2}, \frac{i}{\sqrt{2}}, -\frac{1}{2}\right) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1/2 \\ -i/\sqrt{2} \\ -1/2 \end{pmatrix} = -\lambda\omega\hbar \left(\frac{1}{4} - \frac{1}{4}\right) = 0\,. $$

The result should not be surprising and indeed can be obtained without performing a single computation. The correction to the ground state energy is zero because the state $|1,1\rangle_y$ has no net projection along the z-direction, as evidenced by the vanishing expectation value of S_z. This reflects the fact that the S_y-eigenstates are orthogonal to the S_z-measurements.

7.7 Exercises

1. **Feynman-Hellmann theorem**

   Consider the eigenvalue equation $H(\lambda)|E(\lambda)\rangle = E(\lambda)|E(\lambda)\rangle$ for an Hermitian operator H depending continuously on a parameter λ, with $E(\lambda)$ and $|E(\lambda)\rangle$ the associated eigenvalues and normalized eigenvectors. Show that

   $$\left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle = \frac{\partial E(\lambda)}{\partial \lambda} \,.$$

2. **Attractive potentials**

   Consider two identical particles moving in two attractive potentials $V_1(r)$ and $V_2(r)$, with $V_1(r) \leq V_2(r)$ for all r. Use the variational method to show that the corresponding ground state energies E_1 and E_2 satisfy $E_1 \leq E_2$. *Hint:* Consider the wave function of the particle moving in the potential $V_2(r)$ as trial wave function for that in the potential $V_1(r)$.

3. **Stuck on a desert island**

   Imagine you're stuck on a desert island with just a set of integral tables and no quantum mechanics books. For some reason you don't understand getting off the island depends on figuring out the ground state energies of a) a particle moving in a one-dimensional infinite potential well with walls at $x = \pm L$, b) the one-dimensional harmonic oscillator and c) the hydrogen atom. Unfortunately, you don't remember how to solve the Schrödinger equation for none of these cases (this is probably true, right?). Use the variational method with trial functions on your own devise to estimate these ground state energies of these systems. Justify the choice of these functions. If you manage to get off the island, check a book and compare you estimates with the exact results.

4. **A linear potential**

   Estimate the ground state energy of a particle in a one-dimensional linear potential

   $$V(x) = \alpha|x|$$

   using a Gaussian trial wave function family $\varphi(x;b) = A\exp\left(-bx^2\right)$ with free b.

5. **A central potential**

   A particle of mass m is subjected to a central potential

   $$V(r) = -Ar^{-3/2} \,,$$

   with A a positive constant.

   a) Calculate an upper bound for the ground state energy using the normalized trial function family $\varphi(\vec{r};\alpha) = \sqrt{\frac{\alpha^3}{\pi}}e^{-\alpha r}$, with $r = |\vec{r}|$ and α a positive parameter.

   b) Write the wave function for the state that minimizes the energy. Interpret the result.

   Hints: $\int_0^\infty x^n e^{-ax}dx = \frac{\Gamma(n+1)}{a^{n+1}}$, *where* $\Gamma(z+1) = z\Gamma(z)$ *and* $\Gamma(1/2) = \sqrt{\pi}$, $\Gamma(1) = 1$.

6. **The Yukawa potential** ✐

   If the photon were to have a nonzero mass, the Coulomb potential would be replaced by a Yukawa potential of the form

   $$V(r) = -V_0 \frac{e^{-\mu r}}{r},$$

   with $V_0 > 0$ and $\mu > 0$ constants. Assuming $\mu a \ll 1$ and using a trial wave function of your devise, estimate the binding energy of a "hydrogen" atom under this potential.

7. **A step perturbation in the infinite potential well** ✐

   A particle of mass m is restricted to move in a one-dimensional infinite well of width L in the presence of a small perturbation,

   $$\Delta H = \begin{cases} \infty, & \text{if} \quad x < 0 \text{ and } x > L, \\ -V_0, & \text{if} \quad 0 < x < \frac{L}{2}, \\ 0 & \text{if} \quad \frac{L}{2} < x < L, \end{cases}$$

   with V_0 a constant. Calculate the energy eigenvalues at first order in perturbation theory.

8. **Two identical fermions in a perturbed infinite potential well** ✐

   Two identical spin-$\frac{1}{2}$ fermions are confined in a one-dimensional infinite potential well of width L, with

   $$V(x) = \begin{cases} \infty, & x < 0 \text{ or } x > L, \\ 0, & 0 \le x \le L. \end{cases}$$

   a) Determine the spatial wavefunction and ground-state energy of the system when the two fermions are in a *symmetric spin configuration.*

   b) Repeat for the case where the two fermions are in an *antisymmetric spin configuration.*

   c) Suppose the particles interact via a short-range interaction

      $$\Delta H(x_1, x_2) = -\lambda \delta(x_1 - x_2),$$

      with $\lambda > 0$. Assuming first-order perturbation theory is applicable, estimate qualitatively how this interaction modifies the energy levels obtained in a) and b). Which configuration is more affected, and why?

9. **An anharmonic static perturbation in the harmonic oscillator** ✐

   Consider a perturbation $\Delta H = \lambda x^4$ on the Hamiltonian of the harmonic oscillator. Determine the energies of the system at second-order in perturbation theory and the perturbed wave function at first order for $\lambda > 0$ and $\lambda < 0$. Argue that first-order perturbation theory breaks down for large enough values of n, independently of the smallness of λ. Explain this physically.

10. **A periodic static perturbation to a particle on a circle** ✐

    A particle of mass m is restricted to move on a circle under the influence of a periodic perturbation

    $$\Delta H(\theta) = V_0 \sin\theta \cos\theta,$$

    with θ the angle measured from a fixed reference point on the circle. Find the wave functions of the unperturbed system that diagonalize $V(\theta)$. Calculate the energy levels at second order in perturbation theory.

11. **A perturbed rotor**

The Hamiltonian of a certain type of rotor is given by $H = H_0 + \Delta H$, with

$$H_0 = \frac{a}{\hbar^2}|\vec{L}|^2 + \frac{b}{\hbar^2}L_z^2\,, \qquad \Delta H = \frac{c}{\hbar^2}L_x^2\,,$$

L the orbital angular momentum operator, and a, b, and c constants with energy units.

a) Calculate the eigenvalues of H_0.

b) Assuming $a > b \gg c$, calculate the eigenvalues of H at first order in perturbation theory for states with $l = 1$. Draw an energy level diagram for the $l = 1$ case as a function of the perturbation parameter c.

c) For an arbitrary value of l, calculate the exact eigenvalues of H when $c = b$. Discuss the validity of the perturbative calculation in part b).

12. **A quadratic perturbation in a two-dimensional square box**

A particle of mass m moving in two dimensions is confined to a square box of side L centered at the origin. The associated potential energy is given by

$$V(x,y) = \begin{cases} 0 & \text{if } -\frac{L}{2} \le x, y \le \frac{L}{2}\,, \\ \infty & \text{otherwise.} \end{cases}$$

If the corresponding Hamiltonian is modified by adding a perturbation

$$\Delta H = V_0\left[1 - \frac{(x+y)^2}{L^2}\right], \quad -\frac{L}{2} \le x, y \le \frac{L}{2}\,,$$

calculate the energy of the ground state and the first excited state at first-order perturbation theory.

Hints:

$$\int_{-L/2}^{L/2} dx\, x^2\, \phi_n^2(x) = \frac{n^2\pi^2 - 6}{12n^2\pi^2}L^2, \quad n = 1, 2, 3, \ldots$$

$$\int_{-L/2}^{L/2} dx\, x\, \phi_n(x)\phi_k(x) = \frac{8nk(-1)^{(n+k-1)/2}}{\left(n^2 - k^2\right)^2\pi^2}L, \quad n \text{ odd and } k \text{ even.}$$

with

$$\phi_n(x) = \sqrt{\frac{2}{L}}\cos\left(\frac{n\pi x}{L}\right), \quad n = 1, 3, 5, \ldots\,, \qquad \phi_n(x) = \sqrt{\frac{2}{L}}\sin\left(\frac{n\pi x}{L}\right), \quad n = 2, 4, 6, \ldots\,.$$

13. **A screened Coulomb potential**

Consider a particle of mass μ and charge e in a central potential differing from the Coulomb potential only in a region $r > R$ where the Coulomb force is screened, namely

$$V(r) = \begin{cases} -\frac{e^2}{r}\,, & 0 < r < R\,, \\ -\frac{e^2}{r}e^{-\lambda(r-R)}\,, & R < r < \infty\,. \end{cases}$$

Considering this difference as a small perturbation, calculate the first-order correction to the ground state energy. What happens in the $\lambda \to 0$ and $R \to \infty$ limits?

14. **A harmonic perturbation to the hydrogen atom**

Working at first order in perturbation theory, determine the effect of a harmonic perturbation

$$\Delta H = \lambda(x^2 + y^2 + z^2)$$

on the energy levels $n = 1, 2$ of the hydrogen atom.

15. **Beta decay and energy level shifts in hydrogen-like atoms**

When a nucleus undergoes β decay, a neutron is transformed into a proton, increasing the nuclear charge by one unit. Using first-order time-independent perturbation theory, estimate the resulting energy shift for the atomic levels. You may find it useful to apply the virial theorem to simplify the calculation. Compare your result with the exact energy levels for a hydrogen-like atom with nuclear charge $Z+1$, and discuss under what conditions the perturbative approach provides a good approximation.

16. **A perturbed spin system**

Given a Hamiltonian $H_0 = \epsilon\,\sigma_z$ with $\epsilon > 0$, calculate the changes in energy induced by a term $\Delta H = \lambda(\sigma_z \cos\theta + \sigma_x \sin\theta)$ at first order in perturbation theory.

17. **A spin $1/2$ particle in an infinite potential well and a weak magnetic field**

Use perturbation theory to estimate the energy of the n-th state of a spin $1/2$ particle of mass m moving in a one-dimensional infinite potential well with walls at $x = \pm L$ under the influence of a weak magnetic field

$$\vec{B} = \begin{cases} -B_0\, z & \text{if} \quad -L \le x \le 0\,, \\ -B_0\, x & \text{if} \quad 0 \le x \le L\,. \end{cases}$$

18. **Two bosons in a one-dimensional infinite well with a contact interaction**

Two particles of mass m and spin 0 are confined to a segment of length L.

a) What are the energies, wave functions, and degeneracies of the three lowest energy states?

b) Show that the associated Hamiltonian commutes with the exchange operator Π, defined (up to a phase) by the properties

$$\Pi\, x_1\, \Pi^{-1} = x_2\,, \quad \Pi\, x_2\, \Pi^{-1} = x_1\,, \quad \Pi\, p_1\, \Pi^{-1} = p_2\,, \quad \Pi\, p_2\, \Pi^{-1} = p_1\,. \tag{7.165}$$

What are the eigenvalues of Π?

c) Identify the observables that commute with H_0 but not with Π. Use this to explain the degeneracies determined in part a).

d) Will the degeneracies be maintained in the presence of an interaction potential $\Delta H(x_1, x_2) = \lambda\delta(x_1 - x_2)$ with $\lambda > 0$ a constant with dimensions of energy? Calculate the energy levels, wave functions, and degeneracies found in part a) to first order in perturbation theory.

19. **A perturbed three-level system**

A system with three unperturbed states can be represented by the perturbed Hamiltonian matrix

$$H = \begin{pmatrix} E_1 & 0 & a \\ 0 & E_1 & b \\ a^* & b^* & E_2 \end{pmatrix} ,$$

with $E_2 < E_1$ and $|a|, |b| \ll E_2 - E_1$. Use perturbation theory to calculate the second-order corrections to E_1 and E_2. Compare your findings with the exact result expanded at second order in λ.

20. **A perturbed 5 dimensional Hilbert space**

The Hilbert space describing a quantum system is $\mathcal{H} = C^5$ with $\{|e_i\rangle\}_{i=1}^5$ an orthonormal basis of $\mathcal{H}$. The system's Hamiltonian is $H = H_0 + \Delta H$, with

$$\begin{aligned} H_0 &= -\epsilon \left[6\,|e_1\rangle\langle e_1| + 5|e_2\rangle\langle e_2| + 2|e_3\rangle\langle e_3|\right] , \\ \Delta H &= \alpha \left[|e_1\rangle\langle e_2| + |e_2\rangle\langle e_1|\right] , \end{aligned}$$

and ϵ and α real, positive constants satisfying $\alpha \ll \epsilon$.

a) Calculate the spectrum (possible energies and degeneracy of each) and the eigenstates of H_0.

b) Show that the first-order perturbative correction to the eigenenergies of H_0 due to W is zero for all levels of H_0.

c) Calculate the first non-zero correction to the energy and eigenstate of the ground level of H_0.

d) Show that the eigenstate of H calculated in the previous part does not coincide with the exact solution.

21. **An interacting two-particle system**

A particle of mass m in a one-dimensional harmonic oscillator of frequancy ω interacts with another nearby spin-1/2 particle whose location is fixed. The system is described by a Hamiltonian

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 - \frac{\omega}{2}S_z + \frac{\epsilon}{2}\left(a^\dagger S_- + a S_+\right) ,$$

with $0 < \epsilon \ll \omega$ a small parameter and $S_\pm = S_x \pm i S_y$, with S_x and S_y the spin operators.

a) Determine the energy spectrum for $\epsilon = 0$. Identify the second excited state.

b) Working at first order in perturbation theory, find the shift in the obtained state induced by considering $\epsilon \neq 0$. What is the energy of this state?

c) Assuming the system to be in the shifted second-excited state, which is the probability of measuring $S_z = \hbar/2$?

22. **Scattering in an inverse parabolic potential**

Consider the scattering of a particle of mass m in a potential

$$U(x) = \begin{cases} 0, & \text{for} \quad x < -a \\ U_0 - \alpha x^2, & \text{for} \quad -a \le x \le a \\ 0, & \text{for} \quad a < x \end{cases}$$

where U_0 a constant, $\alpha = 2mU_0^2/\hbar^2$ and $a = \sqrt{U_0/\alpha}$. Calculate the transmission coefficient $T = T(E)$ in the WKB approximation as a function of the energy E of the incoming particle.

23. **Energy levels of half-monomial potentials**

Use the WKB approximation to determine the energy levels of a particle in the one dimensional potential

$$V(x) = \begin{cases} \infty, & x < 0 \\ sign(\alpha) \cdot x^\alpha, & x \ge 0. \end{cases}$$

For what values of α is there a normalisable ground state? Discuss in detail the cases with $\alpha = 2$ and -1.

Hint: Define the integral $C(\alpha) \equiv \int_0^1 \sqrt{S(1-t^\alpha)}dt$ *with* $S \equiv sign(\alpha)$ *and study its convergence.*

24. **A particle in a one-dimensional crystal**

A particle of mass m in a one-dimensional crystal is free to move in the region $-a < x < a$ but it is subject in harmonic forces beyond this range

$$V(x) = \begin{cases} \frac{1}{2}m\omega^2(x-a)^2, & x > a, \\ 0, & -a < x < a, \\ \frac{1}{2}m\omega^2(x+a)^2, & x < -a. \end{cases}$$

Find the energy spectrum in the WKB approximation. How does the result behave in the limits $a \gg L$ and $a \ll L$, with $L \equiv \sqrt{\hbar/(m\omega)}$ the characteristic oscillator length?

25. **An electron in an electric field**

An electron moves under the influence of an electric field

$$\mathcal{E}(x) = \begin{cases} \mathcal{E}\,\vec{x} & x > a, \ \ \forall y, z \\ 0 & -a < x < a, \ \forall y, z \\ -\mathcal{E}\,\vec{x} & x < -a, \ \forall y, z \end{cases}$$

In the WKB approximation the x-dependent part of its energy eigenfunctions in the $x > 0$ region is of the form

$$\psi(x) = \begin{cases} C\,\kappa^{-1/2} \exp\left[1/\hbar \int_x^b dx\,\kappa(x)\right] & x \gg b, \\ C'k^{-1/2} \cos\left[1/\hbar \int_x^b dx\,k(x) - \pi/4\right] & x \ll b, \end{cases}$$

where $k^2 = 2m[E - V(x)]$, $\kappa^2(x) = 2m[V(x) - E]$ and b in the turning point defined by $E = V(b)$. Find the energy eigenvalues in this approximation.

CHAPTER 8

APPROXIMATE METHODS FOR TIME-DEPENDENT SYSTEMS

Time, he said, is what keeps everything from happening at once.

RAY CUMMINGS, THE GIRL IN THE GOLDEN ATOM, 1922

Up to know, we have restricted ourselves to time-independent Hamiltonians. But nothing in this world is completely time independent! In this chapter we will extend our general considerations on perturbation theory to the case of a time-dependent perturbation over a time-independent Hamiltonian. These techniques are especially useful in scattering theory, problems involving the emission and absorption of radiation and in field theoretical problems of various kinds. The standard application is to calculate the transition probabilities between given energy levels per unit interval of time.

8.1 Time-dependent Hamiltonians

Consider a Hamiltonian

$$H = H_0 + V(t) \tag{8.1}$$

containing a time-independent part H_0 whose solution is assumed to be known,

$$H_0|m\rangle = E_m|m\rangle, \tag{8.2}$$

and a time-dependent piece $V(t)$ effectively turned on at a time $t = t_0$,

$$V(t) = \begin{cases} 0 & \text{for} \quad t \leq t_0\,, \\ V(t) & \text{for} \quad t > t_0\,, \end{cases} \tag{8.3}$$

and whose analytical solution is typically unavailable. Given a state $|\varphi(t_0)\rangle = \sum_m c_m(t_0)|m\rangle$ at time $t = t_0$, the dynamics of the system at $t > t_0$ is specified by a time-dependent wave function $|\varphi(t)\rangle_S$ satisfying

$$i\hbar\frac{\partial}{\partial t}|\varphi(t)\rangle_S = (H_0 + V(t))|\varphi(t)\rangle_S\,. \tag{8.4}$$

Since the complete set of unperturbed eigenvectors $|n\rangle$ span the Hilbert space at any time t, we can always expand this wave function in terms of them, namely

$$|\varphi(t)\rangle_S = \sum_n c_n(t)\exp\left(\frac{-iE_n(t-t_0)}{\hbar}\right)|n\rangle\,. \tag{8.5}$$

Note that in writing this expression, we have explicitly separated the usual exponential factor appearing when $V = 0$, relegating therefore any perturbation effect to the time-dependent coefficients $c_n(t)$.

8.2 The interaction picture

In order to obtain the coefficients $c_n(t)$, it is convenient to introduce the so-called *Dirac or interaction picture*, which is an intermediate representation between the Schrödinger and Heisenberg ones that we introduced in Chapter 2.

A given state vector in the interaction picture is defined by a unitary transformation stripping off only the evolution due to H_0 and bringing the Schrödinger ket $|\varphi(t)\rangle_S$ back to the initial time $t = t_0$, namely

$$|\varphi(t)\rangle_{\rm I} = \exp\left(\frac{iH_0(t-t_0)}{\hbar}\right)|\varphi(t)\rangle_S\,, \tag{8.6}$$

which, taking into account Eq. (8.5), can be alternatively written as

$$|\varphi(t)\rangle_{\rm I} = \sum_n c_n(t)|n\rangle\,. \tag{8.7}$$

The corresponding unitary transformation relating a given operator in the two representations is given by

$$O_{\rm I} = \exp\left(\frac{iH_0(t-t0)}{\hbar}\right)O_S\exp\left(-\frac{iH_0(t-t0)}{\hbar}\right)\,, \tag{8.8}$$

which applies, in particular, to the time-dependent perturbation V,

$$V_{\rm I} = \exp\left(\frac{iH_0(t-t_0)}{\hbar}\right)V\exp\left(-\frac{iH_0(t-t_0)}{\hbar}\right)\,. \tag{8.9}$$

The equations of motion in the interaction picture are obtained by multiplying Eq. (8.6) by $i\hbar$ and differentiating the result with respect to time, namely

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}|\varphi(t)\rangle_{\rm I} &= -H_0\exp\left(\frac{iH_0(t-t_0)}{\hbar}\right)|\varphi(t)\rangle_S + \exp\left(\frac{iH_0(t-t_0)}{\hbar}\right)i\hbar\frac{\partial}{\partial t}|\varphi(t)\rangle_S \\ &= -H_0\exp\left(\frac{iH_0(t-t_0)}{\hbar}\right)|\varphi(t)\rangle_S + \exp\left(\frac{iH_0(t-t_0)}{\hbar}\right)(H_0+V)|\varphi(t)\rangle_S \\ &= \exp\left(\frac{iH_0(t-t_0)}{\hbar}\right)V\,|\varphi(t)\rangle_S \\ &= \exp\left(\frac{iH_0(t-t_0)}{\hbar}\right)V\,\exp\left(-\frac{iH_0(t-t_0)}{\hbar}\right)\exp\left(\frac{iH_0(t-t_0)}{\hbar}\right)|\varphi(t)\rangle_S\,, \end{aligned} \tag{8.10}$$

where we have used the Shcrödinger equation (8.4) and inserted the identity in the last step. Taking into account the definitions (8.6) and (8.9), we obtain an expression closely resembling the Schrödinger equation (8.4) but with the total Hamiltonian H replaced by $V_{\rm I}$,

$$i\hbar\frac{\partial}{\partial t}|\varphi(t)\rangle_{\rm I} = V_{\rm I}|\varphi(t)\rangle_{\rm I}\,, \tag{8.11}$$

Defining a time-evolution operator $U_{\rm I}(t,t_0)$ relating the state $|\varphi(t)\rangle_{\rm I}$ to the inital state $|\varphi(0)\rangle_{\rm I}$ in this picture,

$$|\varphi(t)\rangle_{\rm I} = U_{\rm I}(t,t_0)|\varphi(t_0)\rangle_{\rm I}\,, \tag{8.12}$$

we can rewrite Eq. (8.11) as

$$i\hbar\frac{dU_{\rm I}(t)}{dt} = V_{\rm I}(t)U_{\rm I}(t,t_0)\,, \tag{8.13}$$

with the boundary condition $U_{\rm I}(t_0,t_0)=\mathbb{I}$. The solution of this equation is given by that of Eq. (3.194), with the formal replacement $H(t)\to V_{\rm I}(t)$, namely

$$\begin{aligned} U_{\rm I}(t,t_0) &= \mathbb{I} - \frac{i}{\hbar}\int_{t_0}^{t} V_{\rm I}(t')\left[1-\frac{i}{\hbar}\int_{t_0}^{t'} V_{\rm I}(t'')U_{\rm I}(t'',t_0)dt''\right]dt' \\ &= \mathbb{I} - \frac{i}{\hbar}\int_{t_0}^{t} V_{\rm I}(t')dt' + \left(-\frac{i}{\hbar}\right)^2\int_{t_0}^{t}dt'\int_{t_0}^{t'}dt''\,V_{\rm I}(t')V_{\rm I}(t'') \\ &+ \dots \\ &+ \left(-\frac{i}{\hbar}\right)^n\int_{t_0}^{t}dt'\int_{t_0}^{t'}dt''\dots\int_{t_0}^{t^{(n-1)}}dt^{(n)}V_{\rm I}(t')V_{\rm I}(t'')\dots V_{\rm I}(t^{(n)})\,. \end{aligned} \tag{8.14}$$

Given a sufficiently small perturbation V, this expression allows to calculate the evolution operator $U_{\rm I}(t,t_0)$ to arbitrary order in perturbation theory.

8.3 Time-dependent perturbation theory

Let us assume for simplicity that the system is initially in an eigenstate of the time-independent Hamiltonian H_0, say $|\varphi(t_0)\rangle = |m\rangle$. Then, according to the definition (8.12), we have

$$|\varphi(t)\rangle_{\rm I} = U_{\rm I}(t,t_0)|m\rangle = \sum_n |n\rangle\langle n|U_{\rm I}(t,t_0)|m\rangle\,, \quad \longrightarrow \quad c_n(t) = \langle n|U_{\rm I}(t,t_0)|m\rangle\,, \tag{8.15}$$

with the complete set of H_0 eigenstates $\{|n\rangle\}$ spanning the Hilbert space of the system. This matrix element is called the *transition amplitude* between the states m and n and coincides, up to a phase, with that computed using the evolution operator $U(t,t_0)$ in the Schödinger picture. Thus, the probability for a transition from $|m\rangle$ to $|n\rangle$ can be computed by using either $U_{\rm I}$ or U, i.e.

$$P_{m\to n} = |c_n(t)|^2 = |\langle n|U_{\rm I}(t,t_0)|m\rangle|^2 = |\langle n|U(t,t_0)|m\rangle|^2\,. \tag{8.16}$$

Expanding

$$c_n(t) = c_n^{(0)}(t) + c_n^{(1)}(t) + c_n^{(2)}(t) + \dots \tag{8.17}$$

with $c_n^{(i)}(t)$ of i-th order in $V_{\rm I}$ and using the series (8.14) together with the relation (8.9), we obtain

$$c_n^{(0)}(t) = \delta_{nm}\,, \tag{8.18}$$

$$c_n^{(1)}(t) = -\frac{i}{\hbar}\int_{t_0}^{t} dt' e^{i\,\omega_{nm}(t'-t_0)}\langle n|V(t')|m\rangle\,, \tag{8.19}$$

$$c_n^{(2)}(t) = \left(-\frac{i}{\hbar}\right)^2\sum_{\rm i}\int_{t_0}^{t}dt'\int_{t_0}^{t'}dt''e^{i\,\omega_{ni}(t'-t_0)}\langle n|V(t')|i\rangle e^{i\,\omega_{im}(t''-t_0)}\langle i|V(t'')|m\rangle\,, \tag{8.20}$$

$$\vdots$$

with

$$\omega_{nm} = \frac{E_n - E_m}{\hbar} \tag{8.21}$$

frequencies of potential quantum transitions between the m-th and n-th energy levels of the unperturbed system. Note that in Eq. (8.20) we have inserted a resolution of the identity $\sum_i |i\rangle\langle i| = \mathbb{I}$ between the two V factors. The associated states $|i\rangle$ are known as intermediate states and encode the existence of a temporal sequence as we move from the right to the left in the product of matrix elements, starting with the initial state $|m\rangle$, moving through the intermediate state $|i\rangle$ and ending with the final state $|n\rangle$.

For the purposes of this course, we will restrict ourselves to applications in which the series can be truncated at first order. Assuming that the final state we are interested in is not the initial state, the associated transition probability takes the form

$$P_{m\to n} \simeq |c_n^{(1)}(t)|^2 = \left| -\frac{i}{\hbar} \int_{t_0}^{t} dt' e^{i\omega_{nm}(t'-t_0)} \langle n|V(t')|m\rangle \right|^2 , \qquad (n \neq m) \tag{8.22}$$

Note that the the *detailed balance* property $P_{m\to n} = P_{n\to m}$ follows from switching n with m and complexly conjugating this expression.

Example 45: A step-function perturbation

As an application of the previous result, let us calculate the transition probability among states for the simplest choice of $V(t)$, namely a perturbation that is switched on at a time $t_0 = 0$ and remains unchanged afterwards,

$$V(t) = \begin{cases} 0 & \text{for } t \leq 0\,, \\ V & \text{constant for } t > 0\,. \end{cases} \tag{8.23}$$

In this case, the matrix elements in (8.22) are independent of time and can be taken out of the integral,

$$P_{m\to n} = \left| -\frac{i}{\hbar} \langle n|V|m\rangle \int_{t_0}^{t} dt' e^{i\omega_{nm}t'} \right|^2 . \tag{8.24}$$

Taking into account that

$$\int_{t_0}^{t} dt' e^{i\omega_{nm}t'} = \left[\frac{e^{i\omega_{nm}t'}}{i\omega_{nm}} \right]_0^t = \frac{1}{i} \frac{e^{i\omega_{nm}t} - 1}{\omega_{nm}} , \tag{8.25}$$

$$\begin{aligned} \left|1 - e^{i\omega_{nm}t}\right|^2 &= \left(1 - e^{i\omega_{nm}t}\right)\left(1 - e^{-i\omega_{nm}t}\right) \\ &= \left(e^{-i\omega_{nm}t/2} - e^{i\omega_{nm}t/2}\right)\left(e^{i\omega_{nm}t/2} - e^{-i\omega_{nm}t/2}\right) = 4\sin^2\frac{\omega_{nm}t}{2} , \end{aligned} \tag{8.26}$$

this results in

$$P_{m\to n} = \frac{4}{\hbar^2} \frac{\sin^2(\omega_{nm}t/2)}{\omega_{nm}^2} |\langle n|V|m\rangle|^2 , \qquad (n \neq m) \tag{8.27}$$

Example 46: A periodic perturbation

Another important case in practice is when V has a periodic time dependence of the form

$$V(t) = \begin{cases} 0 & \text{for} \quad t \leq 0 \\ Ve^{-i\omega_0 t} + V^\dagger e^{i\omega_0 t} & \text{for} \quad t > 0 \end{cases}, \tag{8.28}$$

with ω_0 the frequency of the perturbation and V and $V^\dagger$ (generally not Hermitian) time-independent operators. This case applies, for instance, to the interaction of atoms with classical electromagnetic waves.
By working at first-order in time-dependent perturbation theory, we obtain

$$\begin{aligned} P_{m\to n}| &\simeq \left| -\frac{i}{\hbar}\int_{t_0}^{t} dt' e^{i\omega_{nm}t'/\hbar}\left(\langle n|V|m\rangle e^{-i\omega_0 t'} + \langle n|V^\dagger|m\rangle e^{i\omega_0 t'}\right)\right|^2 \\ &= \left| \frac{1-e^{-i(\omega_{nm}-\omega_0)t}}{\hbar(\omega_{nm}-\omega_0)}\langle n|V|m\rangle + \frac{1-e^{-i(\omega_{nm}+\omega_0)t}}{\hbar(\omega_{nm}+\omega_0)}\langle n|V^\dagger|m\rangle \right|^2 . \end{aligned} \tag{8.29}$$

Note that, given a final state n, this expression contains a priori cross or interference terms between the two contributions to the amplitude. In practical cases, however, will be mostly interested in final states for which one or the other of the two denominators in Eq. (8.29) is small,

$$\omega_{mn} \mp \omega_0 \approx 0 \quad \longrightarrow \quad E_n = E_m \pm \hbar\omega_0 \,. \tag{8.30}$$

The first of these two cases maximizes the first term amplitude in (8.29) and correspond to the absortion of a quantum of energy $\hbar\omega_0$ from the perturbation,

$$P_{m\to n}|_{\text{abs}} = \frac{4}{\hbar^2}\frac{\sin^2((\omega_{nm}-\omega_0)t/2)}{(\omega_{nm}-\omega_0)^2}|\langle n|V|m\rangle|^2 \,, \quad (\text{Absortion}, n \neq m). \tag{8.31}$$

The second case is resonant when the system emits a quantum of energy $\hbar\omega_0$,

$$P_{m\to n}|_{\text{emis}} = \frac{4}{\hbar^2}\frac{\sin^2((\omega_{nm}+\omega_0)t/2)}{(\omega_{nm}+\omega_0)^2}|\langle n|V|m\rangle|^2 \,, \quad (\text{Emission}, n \neq m) \,. \tag{8.32}$$

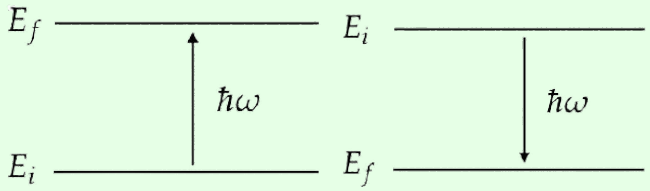


Figure 8.1: (Left) Absorption. (Right) Stimulated emission (if initial state is excited).

8.4 Temporal and energy dependence of the probability

Eqs. (8.27), (8.31) and (8.32) share a similar temporal and energy dependence. This is effectively controlled by a function

$$f(\omega) = \frac{\sin^2(\omega t/2)}{\omega^2}\,, \tag{8.33}$$

with $\omega = \omega_{ni}$ for the step-function case and $\omega = \omega_{ni} \pm \omega_0$ for the oscillatory one.

- *Temporal dependence*: As a function of time, (8.33) oscillates with frequency ω between 0 and a maximum amplitude proportional to $1/\omega^2$. The frequency ω measures the "off-resonance" of the system, i.e. how much the final and initial energies differ. Given a time t, the most favored transitions are those satisfying the *uncertainty relation*

$$\Delta t \lesssim \frac{2\pi}{\omega} = \frac{2\pi\hbar}{\Delta E}\,, \quad \longrightarrow \quad \Delta E \Delta t \lesssim h\,, \tag{8.34}$$

  with Δt the time over which the potential has been turned on, $\Delta E = |E_n - E_m|$ and $h = 2\pi\hbar$ the unreduced Planck constant. If ω is large, the transition probability oscillates rapidly between zero and a small maximum, but as the energies of the final and initial states approach each other, the period of oscillations becomes longer and the amplitude grows. In particular, if the final and initial state are degenerate in energy ($\omega = 0$), [1] the probability $P_{m\to n}$ grows without bound,

$$\lim_{t\to\infty} f(\omega) = \frac{t^2}{4}\,, \tag{8.35}$$

  signalling the eventual breaking of first order perturbation theory. Note, however, that this situation is unlikely to happens in systems with a discrete spectrum. In that case, first-order theory predicts that the transition probability to all final states just oscillates in time. On the other hand, if the final states exist within a continuum, numerous final states with energies infinitesimally close to that of the initial state could become available, making necessary to study the relation between transition probability and energy.

- Frequency dependence: As a function of frequency, (8.33) has a typical diffraction pattern. As illustrated in Fig. 8.2, it has a central peak at $\omega = 0$ with amplitude $t^2/4$ and width $\propto 1/t$ and displays a series of periodic minima,

$$f(\omega) = 0 \quad \longrightarrow \quad \frac{\omega t}{2} = k\pi \quad (k \neq 0)\,, \tag{8.36}$$

  and maxima,

$$f(\omega) = \frac{1}{\omega^2} \quad \longrightarrow \quad \frac{\omega t}{2} = \frac{\pi}{2} + k\pi\,. \tag{8.37}$$

  At time increases the central maximum grows in amplitude while becoming also narrower, a behavior reminiscent of functions approaching a delta function, although in this scenario, the result is not properly a delta function due to the non-conservation of area.

[1] Note, however, that the states cannot be the same, since we are assuming $i \neq j$

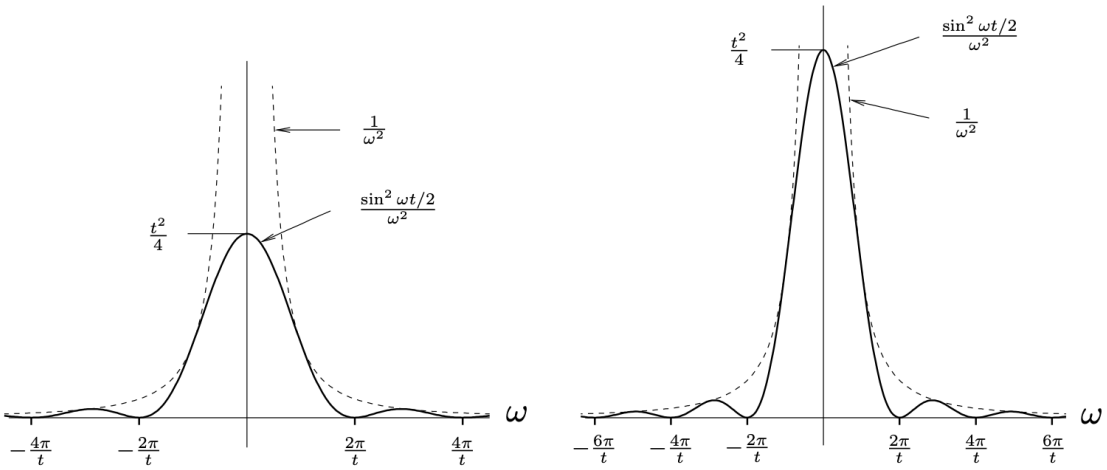


Figure 8.2: (Left) $f(\omega)$ as a function of ω. The dotted lines corresponds to the envelope $1/\omega^2$. (Right) Same function for a larger value of t. The area of the curve is dominated by the central lobe and grows linearly with t.

Lasers

In a system at thermal equilibrium, there is an equal likelihood of absorption and spontaneous emission. This results into a population of atoms or molecules distributed among different energy levels, with more of them in lower energy states. To achieve population inversion, with more particles in higher than lower energy states, it is necessary to supply energy to the system. This can be done through *optical pumping*, where photons are used to raise electrons to higher energy states. Population inversion sets the stage for stimulated emission. In this process, an already excited atom or molecule is stimulated by an incoming photon to release a second identical photon. The crucial point is that stimulated emission is proportional to the number of particles in the excited state. A laser cavity, typically consisting of mirrors at each end, helps trapping and amplifying the light produced through stimulated emission. The feedback provided by the mirrors allows for the amplification of specific wavelengths, resulting in coherent and directional laser light.

8.5 Transitions to the continuum: The Fermi's Golden Rule

Building on the previous examples, let us consider a transition from a discrete initial energy level E_m into a group of eigenstates with a very dense (essentially continuous) spectrum of final energy levels. This situation appears, for instance, in scattering problems, α-decays or optical emissions, where the momenta of the final particle state varies continuously. In these settings, the total transition probability at first in perturbation theory is given by the sum of the transition probabilities to all the final states with $E_n \approx E_m$, namely

$$P_{m\to n} = \sum_{E_n \approx E_m} |c_n^{(1)}(t)|^2 \qquad \longrightarrow P_{m\to n} \simeq \int dE_n \, \rho(E_n) \, |c_n^{(1)}(t)|^2 \,, \tag{8.38}$$

where in the last step we have introduced an energy density of final states $\rho(E_n)$ and replaced the sum by an integral over a group of states in an energy interval dE_n.

Applying (8.38) to the constant V scenario considered in the previous section we get

$$P_{m\to n} \simeq \int dE_n\, \rho(E_n) \frac{|\langle n|V|m\rangle|^2}{\hbar^2} 4f(\omega)\,. \tag{8.39}$$

For $t \to \infty$, we can make use of the fact that $f(\omega)$ is indeed related to a representation of the Dirac delta function,

$$\lim_{t\to\infty} 4f(\omega) = 2\pi t\delta(\omega) = 2\pi t\,\hbar\,\delta(E_n - E_m)\,. \tag{8.40}$$

where in the last step we have employed the property $\delta(x/a) = |a|\delta(x)$ to rewrite the result in terms on energies.

Validity of the approximation

In order for the function $f(\omega)$ to be substituted by a delta function, the width of the energy distribution of the final states ΔE must significantly exceed $h/\Delta t$. On top of that, the number of states within the resulting delta function must be large enough as to allow their characterization by a density of states. Denoting the energy level separation by $\delta\epsilon$, we must therefore have

$$\Delta E \gg \frac{h}{\Delta t} \gg \delta\epsilon \quad \longrightarrow \quad \frac{h}{\Delta E} \ll t \ll \frac{h}{\delta\epsilon} \tag{8.41}$$

Combining (8.39) and (8.40),

$$\lim_{t\to\infty} P_{m\to n} \simeq \frac{2\pi t}{\hbar} \int dE_n\, \rho(E_n) |\langle n|V|m\rangle|^2\, \delta(E_n - E_m)\,, \tag{8.42}$$

and *assuming approximately equal matrix elements for all the final states*,[2] we have

$$\lim_{t\to\infty} P_{m\to n} \simeq t \cdot \frac{2\pi}{\hbar} \rho(E_n) \overline{|\langle n|V|m\rangle|^2}\,, \tag{8.43}$$

with $\overline{|\langle n|V|m\rangle|^2}$ the average value of the matrix element for all final states with $E_n \approx E_m$. Since the expression is linear in t, we can immediately identify a transition rate or transition probability per unit of time,

$$\Gamma_{m\to n} = \frac{2\pi}{\hbar} \rho(E_n) \overline{|\langle n|V|m\rangle|^2}\,. \tag{8.44}$$

[2] If this would not be the case, one should rather consider *differential transition rates* to states with similar matrix elements. A typical example appears when considering elastic scattering, where the matrix elements depend strongly on the direction and magnitude of the momentum of the scattered particle.

The last two expressions were derived by Pauli in 1927 [P. A. M. Dirac, *The quantum theory of emission and absorption of radiation.* Proc. Roy. Soc. (London) A 114, 243265 (1927)], being coined much later as *the Golden Rule* by Fermi due to their multiple applications [(E. Fermi, *Nuclear Physics*, University of Chicago Press, 1950)]. According to these formulas, a given transition is only possible if $\langle n|V|m\rangle \neq 0$ (*selection rules*).

Example 47: The photoelectric effect

The photoelectric effect is a phenomenon where electrons are ejected from the surface of a material when it absorbs light or other electromagnetic radiation. It was first observed by Heinrich Hertz in 1887 and later explained by Albert Einstein in 1905, for which he won the Nobel Prize in Physics in 1921. To study this effect using time-dependent perturbation theory, let us consider the Hamiltonian of an electron in a Coulomb potential interacting with a weak external electromagnetic field,

$$H = \frac{1}{2m}(\vec{P} - e\vec{A})^2 - \frac{e^2}{r} = \underbrace{\frac{\vec{P}^2}{2m} - \frac{e^2}{r}}_{H_0} - \frac{e}{2m}(\vec{P}\cdot\vec{A} + \vec{A}\cdot\vec{P}) + \frac{e^2}{2m}\vec{A}^2\,, \tag{8.45}$$

with $\vec{A}$ a vector potential of the form

$$\vec{A} = \vec{A}_0 \cos\Big(\omega_0 t - \vec{k}\cdot\vec{r}\Big)\,. \tag{8.46}$$

Working in the Coulomb gauge $\nabla\cdot\vec{A} = 0$ where

$$(\vec{P}\cdot\vec{A} + \vec{A}\cdot\vec{P})\,\varphi = (-i\hbar\nabla\cdot\vec{A} + 2\vec{A}\cdot\vec{P})\,\varphi = (2\vec{A}\cdot\vec{P})\,\varphi\,, \tag{8.47}$$

and neglecting the quadratic term proportional to $\vec{A}^2$ at first order in time-dependent perturbation theory, the above Hamiltonian becomes $H = H_0 + H_1$ with

$$H_1 = -\frac{e}{m}\cos\Big(\omega_0 t - \vec{k}\cdot\vec{r}\Big)\vec{A}_0\cdot\vec{P} = -\frac{e}{2m}\left(e^{i(\omega_0 t - \vec{k}\cdot\vec{r})} + e^{-i(\omega_0 t - \vec{k}\cdot\vec{r})}\right)\vec{A}_0\cdot\vec{P}\,. \tag{8.48}$$

The first term gives a transition rate proportional to $\delta\left(E_n - E_m + \hbar\omega_0\right)$, while the second term gives a transition rate proportional to $\delta\left(E_n - E_m - \hbar\omega_0\right)$. Assuming the atom to be initially is in its ground state, only the second term can contribute. According to the computations in the previous section, we have

$$\Gamma_{0\to n} = \frac{2\pi}{\hbar}\sum_{\vec{k}_n}\left|\langle\vec{k}_n|\frac{e}{2m}\vec{A}_0\cdot\vec{P}\,e^{-i\vec{k}\cdot\vec{r}}|0\rangle\right|^2\delta\left(E_m - E_0 - \hbar\omega_0\right)\,, \tag{8.49}$$

with $|0\rangle$ the ground state of the hydrogen atom, with associated wave function

$$\varphi_0(\vec{x}) = \langle\vec{x}|0\rangle = \frac{1}{\left(\pi a_0^3\right)^{1/2}}e^{-r/a_0}\,, \tag{8.50}$$

and $|\vec{k}_n\rangle$ a plane wave state with wave vector $\vec{k}_n$ describing the electron kicked out of the atom by the time-dependent perturbation. The matrix element in (8.49) can be

more easily computed by temporally putting the system in a box of side L. Using this trick, we get

$$\begin{aligned}\frac{e}{2m}\vec{A}_0\cdot\int \mathrm{d}^3x\frac{e^{-i\vec{k}_n\cdot\vec{r}}}{L^{3/2}}(-i\hbar\nabla)\frac{e^{-r/a_0}}{\left(\pi a_0^3\right)^{1/2}} &= \frac{e}{2m}\vec{A}_0\cdot\int\frac{\mathrm{d}^3x}{L^{3/2}}\left(i\hbar\nabla\left(e^{-i\vec{k}_n\cdot\vec{r}}\right)\right)\frac{e^{-r/a_0}}{\left(\pi a_0^3\right)^{1/2}}\\ &= \frac{e\hbar}{2m}\frac{\vec{A}_0\cdot\vec{k}_n}{L^{3/2}\left(\pi a_0^3\right)^{1/2}}\int \mathrm{d}^3x\, e^{-i\vec{k}_n\cdot\vec{r}}e^{-r/a_0} = \frac{e\hbar}{2m}\frac{\vec{A}_0\cdot\vec{k}_n}{L^{3/2}\left(\pi a_0^3\right)^{1/2}}\frac{8\pi}{a_0\left(k_n^2+1/a_0^2\right)^2},\end{aligned} \tag{8.51}$$

where we have intentionally ignored the factor $e^{i\vec{k}\cdot\vec{r}}$ in (8.49) since for the typical wavelengths involved in photo-ionization, $|\vec{k}|\ll|\vec{k}_n|$. Making now the formal replacement

$$\sum_{\vec{k}_n}\to L^3\int\frac{\mathrm{d}^3k_n}{(2\pi)^3}, \tag{8.52}$$

we get

$$\begin{aligned}\Gamma_{0\to n} &= \frac{2\pi}{\hbar}\frac{L^3}{L^3}\int\frac{\mathrm{d}^3k_n}{(2\pi)^3}\left(\frac{e\hbar}{2m}\right)^2\frac{\left(\vec{A}_0\cdot\vec{k}_n\right)^2}{\pi a_0^3}\frac{(8\pi)^2}{a_0^2\left(k_n^2+1/a_0^2\right)^4}\delta\left(E_f-E_0-\hbar\omega\right)\\ &= \frac{e^2\hbar}{2m^2a_0^5}\int\frac{\mathrm{d}k_nk_n^2\,\mathrm{d}\Omega}{(2\pi)^3}\frac{\left(\vec{A}_0\cdot\vec{k}_n\right)^2(8\pi)^2}{\left(k_n^2+1/a_0^2\right)^4}\delta\left(\frac{\hbar^2k_n^2}{2m}-E_0-\hbar\omega\right)\\ &= \frac{e^2k_n}{2m^2a_0^5\hbar}\int\frac{\mathrm{d}\Omega}{(2\pi)^3}\frac{\left(\vec{A}_0\cdot\vec{k}_n\right)^2(8\pi)^2}{\left(k_n^2+1/a_0^2\right)^4},\end{aligned} \tag{8.53}$$

with k_n determined by energy conservation and the only angular dependence hidden in $\vec{A}_0\cdot\vec{k}_n$. Using

$$\int\mathrm{d}\Omega\cos^2\theta=\frac{4\pi}{3},$$

we obtain

$$\Gamma_{0\to n}=\frac{16e^2k_n^3A_0^2}{3m\hbar a_0^5}\frac{1}{\left(k_n^2+1/a_0^2\right)^4}. \tag{8.54}$$

8.5.1 Fermi's golden rule at second order in perturbation theory

While the first-order matrix element becomes zero due to symmetry considerations or selection rules, it becomes necessary to rely on second-order perturbation theory. Focusing once again in the constant V scenario and particularizing the first and second-order expressions

(8.19), and (8.20),

$$c_n^{(1)}(t) = -\frac{i}{\hbar}\langle n|V|m\rangle \int_{t_0}^{t} dt' e^{i\,\omega_{nm}(t'-t_0)}\,, \tag{8.55}$$

$$\begin{aligned} c_n^{(2)}(t) &= \left(-\frac{i}{\hbar}\right)^2 \sum_i \langle n|V|i\rangle\langle i|V|m\rangle \int_{t_0}^{t} dt' e^{i\,\omega_{ni}(t'-t_0)} \int_{t_0}^{t'} dt'' e^{i\,\omega_{im}(t''-t_0)} \\ &= \frac{i}{\hbar}\sum_i \frac{\langle n|V|i\rangle\langle i|V|m\rangle}{E_i - E_m} \int_{t_0}^{t} dt' \left(e^{i\,\omega_{nm}(t'-t_0)} - e^{i\,\omega_{ni}(t'-t_0)}\right)\,, \end{aligned} \tag{8.56}$$

we get $(n \neq i)$

$$\begin{aligned} c_n(t) &= c_n^{(0)}(t) + c_n^{(1)}(t) + c_n^{(2)}(t) \\ &= -\frac{i}{\hbar}\left[\langle n|V|m\rangle - \sum_i \frac{\langle n|V|i\rangle\langle i|V|m\rangle}{E_i - E_m}\right] \int_0^t dt' e^{i\,\omega_{nm}(t'-t_0)} \\ &\quad - \frac{i}{\hbar}\sum_i \frac{\langle n|V|i\rangle\langle i|V|m\rangle}{E_i - E_m} \int_{t_0}^{t} dt' e^{i\,\omega_{ni}(t'-t_0)}\,. \end{aligned} \tag{8.57}$$

When computing the associated transition probability at second order,

$$P_{m\to n} \simeq |c_n(t)|^2 = |c_n^{(0)}(t) + c_n^{(1)}(t) + c_n^{(2)}\,,(t)|^2\,, \tag{8.58}$$

one should a priori take into account interference terms between the amplitudes at different orders. Note, however, that the first term in the above expression has the same time-dependence than that at first order and gives therefore to a transition probability piece that grows linearly with time. On the contrary, the last term leads only to rapid oscillations in time and can be ignored in the $t \to \infty$ limit. Redoing the computation we performed for the first order term, we get finally

$$\Gamma_{m\to n} = \frac{2\pi}{\hbar}\rho(E_n)\overline{\left|\langle n|V|m\rangle - \sum_i \frac{\langle n|V|i\rangle\langle i|V|m\rangle}{E_i - E_m}\right|^2}\,.$$

8.5.2 Electric dipole selection rules

Having established the basics of time-dependent perturbation theory and derived the Fermi Golden Rule, we now have a framework to understand the probabilities of transitions between states under a time-dependent perturbing interaction, such as an oscillating electric field. In the context of atomic and molecular transitions, this oscillating field typically couples to the electric dipole moment $\vec{D}$ of the system of charges, defined as

$$\vec{D} = -e\vec{X}\,, \tag{8.59}$$

with e the elementary charge and $\vec{X} = (X, Y, Z)$ the position operator relative to a chosen origin, typically identified with the nucleus in atomic systems. The associated matrix elements between initial and final states determine the likelihood of a transition. To further refine our understanding of which transitions are possible or "allowed" in this framework, we consider the so-called electric dipole selection rules for the matrix elements

$$\langle n', l', m'|X|n, l, m\rangle\,, \qquad \langle n', l', m'|Y|n, l, m\rangle\,, \qquad \langle n', l', m'|Z|n, l, m\rangle\,. \tag{8.60}$$

Selection rules for m

The electric dipole matrix elements $\langle n', l', m'|Z|n, l, m\rangle$ in (8.60) were already considered in our treatment of the Stark effect, with the commutation relation $[L_z, Z] = 0$ enforcing that $m' = m$ or equivalently

$$\Delta m = m' - m = 0\,. \tag{8.61}$$

The other two matrix elements in (8.60) can be computed by taking into account the expectation value of the commutation relations between L_z and X, Y,

$$\begin{aligned} [L_z, X] &= [XP_y - YP_x, X] = 0 - [YP_x, X] = i\hbar Y\,, \\ [L_z, Y] &= [XP_y - YP_x, Y] = [XP_y, Y] - 0 = -i\hbar X\,, \end{aligned} \tag{8.62}$$

namely

$$\begin{aligned} i\hbar\,\langle n', l', m'|\,Y|n, l, m\rangle &= \hbar\,(m' - m)\,\langle n', l', m'|\,X|n, l, m\rangle\,, \\ -i\hbar\,\langle n', l', m'|\,X|n, l, m\rangle &= \hbar\,(m' - m)\,\langle n', l', m'|\,Y|n, l, m\rangle\,. \end{aligned} \tag{8.63}$$

Note that the solution $m' - m = 0$ found for $\langle n', l', m'|Z|n, l, m\rangle$ would lead to the trivial output $\langle n', l', m'|\,X|n, l, m\rangle = \langle n', l', m'|\,Y|n, l, m\rangle = 0$. Solving the above system of equations in the alternative case $m' \neq m$, we find

$$\hbar^2\,\big\langle n', l', m'\big|\,Y|n, l, m\big\rangle = \hbar^2\,\big(m' - m\big)^2\,\big\langle n', l', m'\big|\,Y|n, l, m\big\rangle\,, \tag{8.64}$$

which is true if and only if $(m' - m)^2 = 1$ or, equivalently,

$$\Delta m = m' - m = \pm 1\,. \tag{8.65}$$

Collecting the results (8.61) and (8.65), the conditions to have a non-zero expectation value for the dipole moment $\langle n', l', m'|\,\vec{D}\,|n, l, m\rangle$ are

$$\boxed{\Delta m = m' - m = 0, \pm 1\,.} \tag{8.66}$$

Selection rules for l

In our analysis of the Stark effect, we demonstrated explicitly that, due to the parity-odd nature of the Z operator, expectation values of the form $\langle n', l', m'|Z|n, l, m\rangle$ vanish identically for same parity states. An additional and more stringent selection rule for first-order electric dipole transitions can be obtained by considering the identity

$$[L^2, [L^2, \vec{X}]] = 2\hbar^2(\vec{X}L^2 + L^2\vec{X})\,, \tag{8.67}$$

whose proof is lengthy but completely straightforward.

Proof of the identity $[L^2, [L^2, \vec{X}]] = 2\hbar^2(\vec{X}L^2 + L^2\vec{X})$

Taking into account that

$$[L_i, X_j] = \epsilon_{inm}\,[X_n P_m, X_j] = \epsilon_{inm} X_n\,[P_m, X_j] = -i\hbar\epsilon_{inm} X_n \delta_{mj} = i\hbar\epsilon_{ijk} X_k\,,$$

we can write

$$\begin{aligned}\left[L^2, X_n\right] &= L_i\left[L_i, X_n\right] + \left[L_i, X_n\right] L_i = i\hbar\epsilon_{inj}\left(L_i X_j + X_j L_i\right) = i\hbar\epsilon_{inj}\left(\epsilon_{ijk}X_k + 2X_j L_i\right) = \\ &= 2i\hbar\left(\epsilon_{ijn}X_i L_j - i\hbar X_n\right)\,,\end{aligned}$$

or equivalently

$$\begin{aligned}\left[L^2, X\right] &= 2i\hbar\left(YL_z - ZL_y - i\hbar X\right),\\ \left[L^2, Y\right] &= 2i\hbar\left(ZL_x - XL_z - i\hbar Y\right),\\ \left[L^2, Z\right] &= 2i\hbar\left(XL_y - YL_x - i\hbar Z\right).\end{aligned}$$

Using the last of these expressions and the standard relation $\left[L^2, L_i\right] = 0$, the commutator $\left[L^2, \left[L^2, Z\right]\right]$ can be written as

$$[L^2, [L^2, Z]] = [L^2, 2i\hbar(XL_y - YL_x - i\hbar Z)]) = 2i\hbar([L^2, X]L_y - [L^2, Y]L_x - i\hbar[L^2, Z])\,.$$

Expanding the first two terms in this expression

$$\begin{aligned}\left[L^2, X\right] L_y &= 2i\hbar\left(YL_z - ZL_y - i\hbar X\right) L_y = 2i\hbar\left(YL_z - i\hbar X\right) L_y - 2i\hbar Z L_y^2\\ &= 2i\hbar\left(L_z Y L_y - Z L_y^2\right)\,,\\ -\left[L^2, Y\right] L_x &= -2i\hbar\left(ZL_z - XL_z - i\hbar Y\right) L_x = 2i\hbar\left(-ZL_x^2 + XL_z L_x + i\hbar Y L_x\right)\\ &= 2i\hbar\left(L_z X L_x - Z L_x^2\right)\,,\end{aligned}$$

and summing them, we obtain

$$\begin{aligned}\left[L^2, X\right] L_y - \left[L^2, Y\right] L_x &= 2i\hbar\left(L_z\left(XL_x + YL_y\right) - Z\left(L_x^2 + L_y^2\right)\right)\\ &= 2i\hbar\left(L_z\vec{X}\cdot\vec{L} - ZL^2\right) = -2i\hbar Z L^2\,,\end{aligned}$$

where in the last step we have taken into account that

$$\vec{X}\cdot\vec{L} = X_i L_i = \epsilon_{ijk}X_i X_j P_k = 0\,.$$

Combining this result with $-i\hbar\left[L^2, Z\right] = -i\hbar\left(L^2 Z - Z L^2\right)$, we get

$$\left[L^2, \left[L^2, Z\right]\right] = 2\hbar^2\left(ZL^2 + L^2 Z\right)\,.$$

Since there is nothing specific here to Z, we conclude that

$$[L^2, [L^2, \vec{X}]] = 2\hbar^2(\vec{X}L^2 + L^2\vec{X})\,.$$

Assuming the matrix elements $\langle n', l', m'|\,\vec{X}\,|n, l, m\rangle$ to be non-trivially vanishing and taking into account the action of L^2 on its eigenstates $|n, l, m\rangle$, $L^2|n, l, m\rangle = \hbar l(l+1)|n, l, m\rangle$, we get

$$\left\langle n', l', m'\right|L^2 L^2\vec{X} - 2L^2\vec{X}L^2 + \vec{X}L^2 L^2\left|n, l, m\right\rangle = 2\hbar^2\left\langle n', l', m'\right|\vec{X}L^2 + L^2\vec{X}\left|n, l, m\right\rangle\,, \quad (8.68)$$

$$\hbar^4\left[l'^2\left(l'+1\right)^2 - 2l'\left(l'+1\right)l(l+1) + l^2(l+1)^2 - 2l'\left(l'+1\right) - 2l(l+1)\right] = 0\,, \quad (8.69)$$

or, equivalently,

$$(l+l')(l+l'+2)((l'-l)^2-1)=0\,. \tag{8.70}$$

Given that the angular momentum quantum numbers are positive $l, l' > 0$, we must have

$$\Delta l = l' - l = \pm 1\,, \tag{8.71}$$

which, in the context of time-dependent electric fields, can be interpreted at the the absorption or emission of a spin-1 photon when transitioning between states.

The Wigner-Eckart theorem

A more systematic way of determining non-zero matrix elements in systems involving angular momentum states is to make use the representation theory of the rotation group. This approach leads to the so-called Wigner-Eckart theorem, which states that the matrix elements of spherical tensor operators[a] $T_q^{(k)}$ can be factorized into the product of a dynamical factor or reduced matrix element $\langle j'\|T^{(k)}\|j\rangle$, which depends only on the total angular momenta j and j' and some geometrical or non-dynamical Clebsch-Gordan coefficients $\langle j,k;m,q|j',m'\rangle$, which depend on the magnetic quantum numbers m and m' and the rank k of the operator, namely

$$\langle j',m'|T_q^{(k)}|j,m\rangle = \langle j,k;m,q|j',m'\rangle\langle j'\|T^{(k)}\|j\rangle\,. \tag{8.72}$$

[a] An irreducible spherical tensor operator of rank k (with k an integer) consists of a set of $2k+1$ operators $T_q^{(k)}$, with $q=-k,\ldots,k$, that transform as the angular momentum under a rotation of the coordinate axes. For instance, the case $k=1$ corresponds to a vector operator.

8.6 The adiabatic approximation

The so-called adiabatic approximation involves a slow temporal change in the Hamiltonian, typically associated with environmental effects such as the motion of the center of mass of a complex system through a media or gradually changing external fields.

Let us consider a Hamiltonian $H(\lambda)$ depending on some number of parameters λ^i assumed to vary adiabatically and displaying, for simplicity, a discrete spectrum

$$H|n(\lambda)\rangle = E_n(\lambda)|n(\lambda)\rangle\,. \tag{8.73}$$

Expanding the wave function in a basis of instantaneous energy eigenstates,

$$|\varphi(t)\rangle = \sum_m a_m(t)e^{i\xi_m(t)}|m(\lambda(t))\rangle\,, \tag{8.74}$$

with unknown coefficients $a_m(t)$ and energy-dependent phase factor

$$\xi_m(t) = -\frac{1}{\hbar}\int_0^t dt' E_m\left(t'\right)\,, \tag{8.75}$$

and substituting the result in the Schrödinger equation

$$i\hbar\frac{\partial|\varphi(t)\rangle}{\partial t} = H|\varphi(t)\rangle\,, \tag{8.76}$$

we obtain

$$\sum_m \left[\dot{a}_m e^{i\xi_m}|m(\lambda)\rangle + a_m e^{i\xi_m}\frac{\partial}{\partial\lambda^i}|m(\lambda)\rangle\dot{\lambda}^i\right] = 0\,, \tag{8.77}$$

where we have cancelled out the two terms depending on E_n. Multiplying this expression by $\langle n(\lambda)|$ we have

$$\dot{a}_n = -\sum_m a_m e^{i(\xi_m-\xi_n)}\Big\langle n(\lambda)\Big|\frac{\partial}{\partial\lambda^i}\Big|m(\lambda)\Big\rangle\dot{\lambda}^i = ia_n\mathcal{A}_i(\lambda)\dot{\lambda}^i - \sum_{m\neq n} a_m e^{i(\xi_m-\xi_n)}\Big\langle n(\lambda)\Big|\frac{\partial}{\partial\lambda^i}\Big|m(\lambda)\Big\rangle\dot{\lambda}^i\,, \tag{8.78}$$

where in the second line we have explicitly singled out the $m = n$ term and defined the so-called *Berry connection*

$$\mathcal{A}_i(\lambda) \equiv -i\Big\langle n(\lambda)\Big|\frac{\partial}{\partial\lambda^i}\Big|n(\lambda)\Big\rangle\,. \tag{8.79}$$

By differentiating now Eq. (8.73) with respect to λ,

$$\frac{\partial H}{\partial\lambda^i}|m(\lambda)\rangle + H\frac{\partial}{\partial\lambda^i}|m(\lambda)\rangle = \frac{\partial E_m}{\partial\lambda^i}|m(\lambda)\rangle + E_m\frac{\partial}{\partial\lambda^i}|m(\lambda)\rangle\,, \tag{8.80}$$

and multiplying the result by $\langle n|$ $(n \neq m)$, we get

$$(E_m - E_n)\,\langle n(\lambda)|\frac{\partial}{\partial\lambda^i}|m(\lambda)\rangle = \langle n(\lambda)|\frac{\partial H}{\partial\lambda^i}|m(\lambda)\rangle \tag{8.81}$$

This means that the second term in (2.34) is proportional to

$$\Big\langle n(\lambda)\Big|\frac{\partial}{\partial\lambda^i}\Big|m(\lambda)\Big\rangle\dot{\lambda}^i = \Big\langle n(\lambda)\Big|\frac{\partial H}{\partial\lambda^i}\Big|m(\lambda)\Big\rangle\frac{\dot{\lambda}^i}{E_m - E_n}\,, \tag{8.82}$$

allowing us to rewrite (8.78) as

$$\dot{a}_n = ia_n\mathcal{A}_i(\lambda)\dot{\lambda}^i - \sum_{m\neq n} a_m e^{i(\xi_m-\xi_n)}\frac{\dot{\lambda}^i}{E_m - E_n}\Big\langle n(\lambda)\Big|\frac{\partial H}{\partial\lambda^i}\Big|m(\lambda)\Big\rangle\,, \tag{8.83}$$

When the change of parameters is much smaller than the splitting of energy levels ($\hbar\dot{\lambda}^i \ll \Delta E$), we are left with

$$\dot{a}_n = ia_n\mathcal{A}_i\dot{\lambda}\,, \quad \longrightarrow \quad a_n = C_n\exp\left(i\int_0^t dt'\mathcal{A}_i\left(\lambda\left(t'\right)\right)\dot{\lambda}^i\right)\,, \tag{8.84}$$

with C_n some integration constants. This means, that if we start in a specific energy eigenstate $|n\rangle$ with $a_m = \delta_{mn}$, the system will remain in the state $|n(\lambda)\rangle$ for slowly varying λ, provided of course that we avoid level crossings leading to $E_m = E_n$.

Level crossings

Level crossings are relatively uncommon in quantum mechanics and typically require some level of finetuning. For instance, a two-level system characterized by a general Hermitian 2×2 matrix depends a priori on four parameters, but its eigenvalues only coincide if the matrix is proportional to the identity. Consequently, three of those parameters must be set to zero for the energies to become degenerate.

As explicit in Eq. (8.84), the variation of the λ^i parameters translates into the appearance of an additional phase beyond that in Eqs. (8.74) and (8.75). The true significance of this additional piece was first recognized by Michael Berry in 1984 and plays nowadays an important role in the description of various topological phenomena, like the quantum Hall effect or topological insulators. Unlike the usual phases associated to energy eigenstates, the *Berry phase* is intrinsically geometrical, meaning that, when considering a closed path C in parameter space,

$$e^{i\gamma} = \exp\left(i \oint_C d\lambda^i \mathcal{A}_i(\lambda) \right) , \tag{8.85}$$

the result does not depend on the time taken to make the journey, but only on the followed trajectory. Note also the information contained in the Berry connection (8.79) is somehow redundant, since we could have alternatively chosen another set of states differing by a phase for each choice of λ^i,

$$|n'(\lambda)\rangle = e^{i\omega(\lambda)} |n(\lambda)\rangle , \tag{8.86}$$

such that

$$\mathcal{A}'_i = -i \left\langle n' \left| \frac{\partial}{\partial \lambda^i} \right| n' \right\rangle = \mathcal{A}_i + \frac{\partial \omega}{\partial \lambda^i} . \tag{8.87}$$

This gauge-like transformation does not affect, however, the Berry phase (8.85) since $\oint \partial_i \omega d\lambda^i = 0$.

In analogy with electromagnetism, the physical information of the Berry connection is encoded in a *gauge invariant field strength* or *curvature of the connection*,

$$\mathcal{F}_{ij}(\lambda) = \frac{\partial \mathcal{A}_i}{\partial \lambda^j} - \frac{\partial \mathcal{A}_j}{\partial \lambda^i} . \tag{8.88}$$

In fact, using the higher-dimensional version of Stokes' theorem, we can write

$$e^{i\gamma} = \exp\left(-i \oint_C \mathcal{A}_i(\lambda) d\lambda^i \right) = \exp\left(-i \int_S \mathcal{F}_{ij} dS^{ij} \right) , \tag{8.89}$$

with S a two-dimensional surface in the parameter space bounded by the path C.

Example 48: Berry's phase for an electron in a precessing magnetic field

A common and straightforward illustration of the Berry phase involves a spin system with a Hilbert space comprising only two states, such as an electron in a slowly precessing magnetic field, whose components we will identify with the parameters λ^i,

namely $\lambda^i \equiv B^i$. The associated Hamiltonian takes the form

$$H = -\gamma \vec{S} \cdot \vec{B} + \tilde{E}\,, \tag{8.90}$$

with $\vec{S} = \hbar/2\,\vec{\sigma}$, $\vec{\sigma}$ the set of three Pauli matrices and $\tilde{E} \equiv \gamma\hbar B/2$ a constant offset included to ensure a zero-energy ground state, i.e. to eliminate *de facto* the usual ground state phase factor $e^{-iE_0 t/\hbar}$, which is irrelevant for our discussion. Writing the magnetic field $\vec{B}$ in spherical coordinates,

$$\vec{B} = \begin{pmatrix} B\sin\theta\cos\omega t \\ B\sin\theta\sin\omega t \\ B\cos\theta \end{pmatrix}\,, \tag{8.91}$$

with ω the precession frequency, we can rewrite Eq. (8.90) as

$$H = -\tilde{E}\begin{pmatrix} \cos\theta - 1 & e^{-i\omega t}\sin\theta \\ e^{+i\omega t}\sin\theta & -\cos\theta - 1 \end{pmatrix}\,. \tag{8.92}$$

For non-vanishing magnetic field, the diagonalization of this Hamiltonian provides two non-degenerate eigenvalues 0 and $+2B$,

$$H|-\rangle = 0 \quad \text{and} \quad H|+\rangle = 2\tilde{E}|+\rangle\,, \tag{8.93}$$

with normalized eigenvectors

$$|-\rangle = \begin{pmatrix} e^{-i\omega t}\sin(\theta/2) \\ -\cos(\theta/2) \end{pmatrix}\,, \qquad |+\rangle = \begin{pmatrix} e^{-i\omega t}\cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix}\,. \tag{8.94}$$

The Berry connections (8.79) arising from these states take the form

$$\mathcal{A}_\theta = -i\langle - |\frac{\partial}{\partial\theta}| - \rangle = 0\,, \qquad \mathcal{A}_\phi = -i\langle - |\frac{\partial}{\partial\phi}| - \rangle = -\sin^2\frac{\theta}{2}\,, \tag{8.95}$$

The resulting Berry curvature in polar and Cartesian coordinates is then given by

$$\mathcal{F}_{\theta\phi} = \frac{\partial\mathcal{A}_\phi}{\partial\theta} - \frac{\partial\mathcal{A}_\theta}{\partial\phi} = -\sin\theta\,, \quad \longrightarrow \quad \mathcal{F}_{ij}(\vec{B}) = -\epsilon_{ijk}\frac{B^k}{2|\vec{B}|^3}\,. \tag{8.96}$$

This is a magnetic monopole in the space of magnetic fields, sitting at the singular point $\vec{B} = 0$.

8.7 The sudden approximation

The sudden approximation involves a *fast* jump between two time-independent Hamiltonians, with *fast* understood as a timescale much shorter than any natural period in the system. In this approximation, perturbation theory is completely irrelevant: if the system is initially in an eigenstate $|n\rangle$ of the initial Hamiltonian H_0, the time evolution following the switch will be given by the final Hamiltonian H_0', meaning that the wave function must be simply

reexpressed in terms of the new eigenstates of the final Hamiltonian, which then determines the time evolution of the system,

$$H_0'|n\rangle = \sum_{n'} |n'\rangle \langle n'|n\rangle \ . \tag{8.97}$$

The only challenging task is estimating the actual time taken for the Hamiltonian to change and the periods of motion associated with the state $|n\rangle$ and with its transitions to neighboring states.

Example 49: Quantum quench of a harmonic oscillator

Consider a rapid transition ($\Delta t \ll \hbar/\omega_0$) in the frequency of a harmonic oscillator, $\omega_0 \to \omega$, at time $t = 0$. The corresponding Hamiltonian

$$H_0 = \frac{p^2}{2m} + \frac{1}{2}m^2\omega_0^2 x^2 = \hbar\omega_0 \left(a_0^\dagger a_0 + \frac{1}{2}\right) , \tag{8.98}$$

with

$$a_0 = \frac{1}{\sqrt{2m\omega_0}} \left(m\omega_0 x + ip\right) , \tag{8.99}$$

evolves then to

$$H_0' = \frac{p^2}{2m} + \frac{1}{2}\omega^2 x^2 = \hbar\omega \left(a^\dagger a + \frac{1}{2}\right) ,$$

with

$$a = \frac{1}{\sqrt{2m\omega}}(m\omega x + ip) = \frac{1}{2}\left(\sqrt{\frac{\omega}{\omega_0}} + \sqrt{\frac{\omega_0}{\omega}}\right) a_0 + \frac{1}{2}\left(\sqrt{\frac{\omega}{\omega_0}} - \sqrt{\frac{\omega_0}{\omega}}\right) a_0^\dagger$$

Expanded it in terms of the eigenstates $|n\rangle, n = 0, 1, \ldots$ of H_0', the ground state $|\emptyset\rangle$ of H_0,

$$a_0|\emptyset\rangle = 0 \, , \qquad \longrightarrow \qquad (\omega + \omega_0)\, a|\emptyset\rangle = (\omega - \omega_0)\, a^\dagger|\emptyset\rangle \, , \tag{8.100}$$

becomes a collection of parity-even excited states

$$|\emptyset\rangle = \sum_{n=0}^{\infty} \alpha_{2n'}|2n'\rangle \quad \text{with} \quad \alpha_{2n'+2} = \sqrt{\frac{2n'+1}{2n+2}} \left(\frac{\omega - \omega_0}{\omega + \omega_0}\right) \alpha_{2n'} \, . \tag{8.101}$$

In the Heisenberg picture, the position operator now evolves as dictated by the new Hamiltonian H,

$$x(t) = x(0)\cos(\omega t) + \frac{p(0)}{m\omega}\sin(\omega t) \, , \tag{8.102}$$

and

$$\langle\emptyset\,|x^2(0)|\,\emptyset\rangle = \frac{\hbar}{2m\omega_0} \quad \text{and} \quad \langle\emptyset\,|p^2(0)|\,\emptyset\rangle = \frac{\hbar m\omega}{2} \, . \tag{8.103}$$

For $t_2 > t_1$, this implies

$$\begin{aligned} &\langle\emptyset\,|x\,(t_2)\,x(t_1)|\,\emptyset\rangle \\ &= \frac{\hbar}{2m\omega}\left[e^{-i\omega(t_2-t_1)} + \frac{\left(\omega^2 - \omega_0^2\right)\cos\left(\omega\,(t_2+t_1)\right) + (\omega-\omega_0)^2\cos\left(\omega\,(t_2-t_1)\right)}{2\omega\omega_0}\right] . \end{aligned} \tag{8.104}$$

The first term in this expression accounts for the usual evolution of an energy eigenstate, i.e. the expected outcome in the absence of any quench. The absence of time translational invariance in the second term reflects the fact that the system retains a memory of the quench!

8.8 Some worked-out exercises

A particle in a evolving infinite square well

A particle of mass m is initially in the ground state of an infinite square well potential with width L. At time $t = 0$, the potential in the left half of the well begins to increase linearly from 0 to V over a time T. After reaching the value V, the potential starts decreasing back to zero at the same constant rate over the next T seconds. Assuming V to be small, determine the probability that, at time $2T$, the particle is found in the first excited state of the well.

The time dependent perturbation has the form

$$H'(t) = \begin{cases} \frac{V}{x} f(t) & 0 \leqslant x \leqslant \frac{L}{2}\,, \\ 0 & \text{elsewhere}\,, \end{cases}$$

with $f(t)$ displaying a triangular shape. The requested transition probability is given by

$$P_{1\to 2} = \left| \frac{1}{\hbar} \int_0^{2T} dt\, \Delta H_{21}(t) e^{i\omega_0 t} \right|^2 ,$$

with

$$\omega_0 = \frac{E_2 - E_i}{\hbar} = \frac{3\pi^2\hbar}{2mL^2}\,.$$

The matrix element $\Delta H(t)$ is given by

$$\Delta H(t) = \langle 2|\Delta H_{21}(t)|1\rangle = V f(t) \int_0^{\frac{L}{2}} \varphi_2^*(x)\varphi_1(x)dx = V f(t) \int_0^{\frac{L}{2}} \frac{2}{L} \sin\left(\frac{\pi x}{L}\right) \sin\left(\frac{2\pi x}{L}\right) dx$$
$$= \frac{2V}{L} f(t) \int_0^{\frac{L}{2}} 2\sin^2\left(\frac{\pi x}{L}\right) \cos\left(\frac{\pi x}{L}\right) dx = \frac{4V}{L} f(t) \frac{1}{3} \sin^3\left(\frac{\pi x}{L}\right) \cdot \frac{L}{\pi}\Big|_0^{\frac{L}{2}} = \frac{4V}{3\pi} f(t)\,.$$

Then, we have

$$P_{1\to 2} = \frac{16V^2}{9\pi^2\hbar^2} \left| \int_0^{2T} dt f(t) e^{i\omega t} \right|^2 = \frac{256V^2}{9\pi^2\hbar^2\omega^4 T^2} \sin^4\left(\frac{\omega T}{2}\right) ,$$

where we have used

$$\int_0^{2T} df\, f(t) e^{i\omega t} = \int_0^T dt \frac{t}{T} e^{i\omega t} + \int_T^{2T} \frac{(2T-t)}{T} e^{i\omega t} = \int_0^T dt \frac{t}{T} \left(e^{i\omega t} + e^{i\omega(2T-t)}\right)$$
$$= -\frac{T}{\omega^2 T^2} \left(1 - e^{i\omega T}\right)^2 = -\frac{e^{i\omega T}}{\omega^2 T^2} \left(2i \sin\left(\frac{\omega T}{z}\right)\right)^2 .$$

A 2D harmonic oscillator under a time-dependent angular perturbation

A two-dimensional harmonic oscillator with Hamiltonian

$$H_0 = \frac{1}{2m}(p_x^2 + p_y^2) + \frac{1}{2} m\,\omega_x^2\, x^2 + \frac{1}{2} m\,\omega_y^2\, y^2$$

is subject at $t \to -\infty$ to a time-dependent perturbation

$$\Delta H(t) = \lambda\, L_z\, e^{-|t|}\,,$$

with λ a small parameter. Assuming the oscillator to be initially in its ground state and working at first order in perturbation theory, determine the probability for it to transition to another energy level as $t \to +\infty$. Discuss the special case where $\omega_x = \omega_y$.

The angular momentum operator

$$L_z = x p_y - y p_x\,,$$

can be written in terms of creation ($a^\dagger$) and annihilation (a) operators through the relations

$$x = \sqrt{\frac{\hbar}{2m\omega_x}}\left(a_x + a_x^\dagger\right)\,, \quad p_x = i\sqrt{\frac{\hbar m\omega_x}{2}}\left(a_x^\dagger - a_x\right)\,, \quad y = \sqrt{\frac{\hbar}{2m\omega_y}}\left(a_y + a_y^\dagger\right)\,, \quad p_y = i\sqrt{\frac{\hbar m\omega_y}{2}}\left(a_y^\dagger - a_y\right)\,.$$

Substituting these expressions into the equation for L_z, we get:

$$L_z = \frac{\hbar}{2i}\sqrt{\frac{\omega_y}{\omega_x}}\left(a_x + a_x^\dagger\right)\left(a_y - a_y^\dagger\right) - \frac{\hbar}{2i}\sqrt{\frac{\omega_x}{\omega_y}}\left(a_y + a_y^\dagger\right)\left(a_x - a_x^\dagger\right)\,.$$

Applying this expression to the ground state $|0\,0\rangle$, we get

$$L_z|00\rangle = -\frac{\hbar}{2i}\left[\sqrt{\frac{\omega_y}{\omega_x}} - \sqrt{\frac{\omega_x}{\omega_y}}\right]|11\rangle\,.$$

The only non-vanishing matrix elements are therefore

$$\langle 11|\Delta H|00\rangle = -\frac{\lambda\hbar}{2i}\left[\sqrt{\frac{\omega_y}{\omega_x}} - \sqrt{\frac{\omega_x}{\omega_y}}\right]e^{-|t|}.$$

All other matrix elements are zero. The transition probability from the ground state to the state $|11\rangle$ is given by

$$P_{00\to 11} = \left|\frac{\lambda}{2}\left[\sqrt{\frac{\omega_y}{\omega_x}} - \sqrt{\frac{\omega_x}{\omega_y}}\right]\int_{-\infty}^{\infty} e^{-|t|}e^{i(\omega_x+\omega_y)t}dt\right|^2 = \frac{\lambda^2}{\left[1+(\omega_x+\omega_y)^2\right]^2}\frac{(\omega_x-\omega_y)^2}{\omega_x\omega_y}\,.$$

If $\omega_x = \omega_y$, this transition is also not allowed, since in this case, the ground state is invariant under rotation around the z-axis and is therefore an eigenstate of L_z (and of the entire Hamiltonian).

A hydrogen in a time-dependent Gaussian electric field

A hydrogen atom, initially in its ground state ($t < 0$), is placed at time $t = 0$ in a time-dependent electric field $\vec{\mathcal{E}}(t) = \mathcal{E}_0 e^{-t^2/\tau^2}(\vec{e}_y + \vec{e}_z)$, with $\mathcal{E}_0 > 0$ and $\tau > 0$ constants. Assuming $\mathcal{E}_0$ to be small, determine whether it is possible for the atom to transition to a state with $n = 2$ at $t > 0$, and if so, identify the possible final states.

The hydrogen atom interacts with the electric field via the electric dipole interaction

$$H_{\text{int}}(t) = -q\vec{\mathcal{E}}(t)\cdot\vec{r}\,.$$

The total Hamiltonian of the system is then given by

$$H = H_0 + V(t) = -\frac{\hbar^2}{2m}\nabla^2 - \frac{e^2}{r} + e\,\mathcal{E}(t)(y+z)\,,$$

with H_0 the Hamiltonian of the hydrogen atom, $V(t) = e\,\mathcal{E}(t)\,(y+z)$ and y and z the position operators in the y and z directions. At first order in perturbation theory, the transition probability from an initial state $|i\rangle \equiv |n,l,m\rangle$ at time $t=0$ to a final state $|f\rangle \equiv |n',l',m'\rangle$ at time t is given by

$$P_{i\to f} \simeq \left| -\frac{i}{\hbar}\int_0^t dt' e^{i\omega_{fi}t'}\langle f|V(t')|i\rangle \right|^2 , \qquad \text{with} \qquad \omega_{fi} = \frac{E_f - E_i}{\hbar}\,.$$

To have a nonzero transition probability one must necessarily have $\langle f|V(t')|i\rangle \propto \langle n',l',m'|y+z|n,l,m\rangle \neq 0$ Considering the transformation of y and z and the states under the parity operator Π,

$$\Pi\, y\, \Pi^\dagger = -y \qquad \Pi\, z\, \Pi^\dagger = -z\,, \qquad \Pi|n,l,m\rangle = (-1)^l|n,l,m\rangle\,,$$

we can write ($\Pi^\dagger\Pi = 1$)

$$\langle n',l',m'|(y+z)|n,l,m\rangle = \langle n',l',m'|\,\Pi^\dagger\Pi\,(y+z)\,\Pi^\dagger\Pi\,|n,l,m\rangle = (-1)^{l+l'+1}\langle n',l',m'|(y+z)|n,l,m\rangle\,,$$

concluding that the matrix elements are also zero for $l+l'$ equal to zero or even. In our case ($n=2$, $l=0,1$, $m=-l,\dots l$),

$$|i\rangle = |1,0,0\rangle\,, \qquad E_i = E_1\,; \qquad |f\rangle = |2,\ell,m\rangle\,, \qquad E_f = E_2\,; \qquad E_n = -\frac{Z^2e^2}{2a_0n^2}\,.$$

the only possible final states leading to $\langle 2,\ell,m|(y+z)|1,0,0\rangle \neq 0$ are of the form $|2,1,m\rangle$. Since

$$z = r\cos\theta \qquad\qquad y = r\sin\theta\sin\phi$$

are proportional to the spherical harmonics

$$Y_{10}(\theta,\phi) = \sqrt{\frac{3}{4\pi}}\cos\theta\,, \qquad Y_{1\pm1}(\theta,\phi) = \mp\sqrt{\frac{3}{8\pi}}\,\mathrm{sen}\,\theta e^{\pm i\phi}\,, \qquad \sin\theta\sin\phi \propto Y_1^{-1} - Y_1^1, \quad \cos\theta \propto Y_1^0.$$

and have therefore quantum numbers $\ell_2 = 1, m_2 = 0$ and $\ell_2 = 1, m_2 = \pm1$, respectively. The composition of these values with those of the ket $|1,0,0\rangle$, ($\ell_1 = 0, m_1 = 0$), provides a state $l_1 + l_2 = 1$, $m_1 + m_2 = 0, \pm1$. Therefore, the only non-vanishing overlaps with $n=2$ are those associated to $|2,1,0\rangle$ and $|2,1,\pm1\rangle$.

Of course, the same conclusion can be achieved by a brout-force computations.

A hydrogen atom in a time-dependent exponential electric field

An hydrogen atom initially ($t<0$) in the state $|n\,,l\,,m\rangle = |3,0,0\rangle$ is subject at time $t=0$ to an electric field $E(t) = E_0\,e^{-t/\tau}$ along the z-axis, with constants E_0 and τ.

a) Determine which of the following transitions are allowed.

$$|3,0,0\rangle \to |2,1,1\rangle\,, \qquad |3,0,0\rangle \to |2,1,-1\rangle\,, \qquad |300\rangle \to |210\rangle\,.$$

b) Working at first order in perturbation theory, determine the transition probability to the allowed final state(s) at time $t_s \gg \tau$, approximating the result of the radial

integration of the hydrogen atom wave functions $\varphi_{nlm}(r,\theta\phi) = R_{nl}(r)Y_{lm}(\theta\phi)$ as $\int_0^\infty R_{30}R_{21}r^3dr \approx a_0$, with a_0 the Bohr radius.

a) The total Hamiltonian of the system is given by

$$H = H_0 + V(t) = -\frac{\hbar^2}{2m}\nabla^2 - \frac{e^2}{r} + e\,E(t)z\,,$$

with H_0 the Hamiltonian of the hydrogen atom, $V(t) = e\,E(t)\,z$ and z the position operator in the z direction. At first order in perturbation theory, the transition probability from an initial state $|i\rangle \equiv |n,l,m\rangle$ at time $t=0$ to a final state $|f\rangle \equiv |n',l',m'\rangle$ at time t is given by

$$P_{i\to f}(t) \simeq \left| -\frac{i}{\hbar}\int_0^t dt' e^{i\omega_{fi}t'}\langle f|V(t')|i\rangle \right|^2 , \qquad \text{with} \qquad \omega_{fi} = \frac{E_f - E_i}{\hbar}\,.$$

To have a nonzero transition probability one must necessarily have $\langle f|V(t')|i\rangle \propto \langle n',l',m'|z|n,l,m\rangle \neq 0$. To determine the non-vanishing matrix elements one can proceed in 3 different ways:

- Way 1: Commutation relations. Using the fact that z commutes with the L_z component of the angular momentum,

  $$m\hbar\langle n',l',m'|z|n,l,m\rangle = \langle n',l',m'|zL_z|n,l,m\rangle = \langle n',l',m'|L_z z|n,l,m\rangle = m'\hbar\langle n',l',m'|z|n,l,m\rangle\,,$$

  we can easily conclude that matrix elements with $m \neq m'$ are zero. On the other hand, considering the transformation of z ($\Pi\, z\, \Pi^\dagger = -z$) and the states ($\Pi|n,l,m\rangle = (-1)^l|n,l,m\rangle$) under the parity operator Π, we can write ($\Pi^\dagger\Pi = 1$)

  $$\langle n',l',m'|z|n,l,m\rangle = \langle n',l',m'|\,\Pi^\dagger\Pi\, z\,\Pi^\dagger\Pi\,|n,l,m\rangle = (-1)^{l+l'+1}\langle n',l',m'|z|n,l,m\rangle\,,$$

  concluding that matrix elements are also zero unless $l+l'$ is odd. The only given non-vanishing transition is therefore the $|3,0,0\rangle \to |2,1,0\rangle$ one, since the transitions $|3,0,0\rangle \to |2,1,\pm1\rangle$ violate the above $m' = m$ rule.
- Way 2: Addition of angular momenta. In spherical coordinates the z operator, $z = r\cos\theta$, is proportional to the spherical harmonic $Y_{10}(\theta,\phi) \equiv \langle\theta,\phi|1,0\rangle$ (see e.g. the Table of Clebsch-Gordan coefficients) and has therefore quantum numbers $\ell_2 = 1, m_2 = 0$. The composition of these values with those of the bra $\langle 3,0,0|$ ($\ell_1 = 0, m_1 = 0$) provides a state with $l_1 + l_2 = 1$, $m_1 + m_2 = 0$. Therefore, the only non-vanishing overlap is again the one associated to $|2,1,0\rangle$.
- Way 3: Brout-force computation. The angular part of the integrals involving the orbitals $Y_{1\pm1}$ in $|2,1,\pm1\rangle$ is identically cero by symmetry considerations

  $$\int_0^\pi \cos\theta\sin^2\theta d\theta \int_0^{2\pi} e^{\pm i\phi}\,d\phi = 0\,.$$

  Therefore, the only non-vanishing overlap is again the one associated to $|2,1,0\rangle$.

b) The first-order transition probability for the process $|3,0,0\rangle \to |2,1,0\rangle$ is then given by

$$\begin{aligned}
P_{i\to f}(t_s) &= \frac{|\langle 2,1,0|z|3,0,0\rangle|^2}{\hbar^2}\left|\int_0^{t_s} dt e^{i(E_2-E_3)t/\hbar}E_0 e e^{-t/\tau}\right|^2 \\
&= \frac{E_0^2e^2\,|\langle 2,1,0|z|3,0,0\rangle|^2}{\hbar^2}\left|\int_0^{t_s} dt e^{t[i(E_2-E_3)/\hbar - 1/\tau]}\right|^2 \\
&= \frac{E_0^2e^2\,|\langle 2,1,0|z|3,0,0\rangle|^2}{\hbar^2}\left|\frac{e^{t[i(E_2-E_3)/\hbar-1/\tau]}}{i\,(E_2-E_3)/\hbar - 1/\tau}\right|_0^{t_s}\right|^2 \\
&\simeq \frac{E_0^2e^2\,|\langle 2,1,0|z|3,0,0\rangle|^2}{\hbar^2}\left|\frac{1}{1/\tau - i\,(E_2-E_3)/\hbar}\right|^2 ,
\end{aligned}$$

where in the last step we have used the condition $t_s \gg \tau$. Using now that the modulus of the inverse of a complex number equals the inverse of its modulus, we have

$$P_{i\to f}\left(t_s\right) = \frac{E_0^2 e^2 \left|\langle 2,1,0|z|3,0,0\rangle\right|^2}{\hbar^2}\left[\frac{1}{1/\tau^2 + \left(E_2 - E_3\right)^2/\hbar^2}\right].$$

The non-vanishing matrix element $\langle 210|z|300\rangle$ in this expression takes the form

$$\langle 210|z|300\rangle = \int R_{21}Y_{10}(r\cos\theta)R_{30}Y_{00}\, r^2\sin\theta dr d\theta d\phi = \int R_{21}R_{30}\, r^3 dr \int Y_{10}Y_{00}\, r^2\cos\theta\sin\theta d\theta d\phi \approx \frac{a_0}{\sqrt{3}},$$

where we have taken into account that

$$\int Y_{10}Y_{00}\cos\theta\sin\theta d\theta d\phi = \sqrt{\frac{3}{4\pi}}\sqrt{\frac{1}{4\pi}}\int_0^\pi \cos^2\theta\sin\theta d\theta \int_0^{2\pi} d\phi = \frac{\sqrt{3}}{4\pi}\left(-\frac{\cos^3\theta}{3}\right)\Bigg|_0^\pi (2\pi) = \frac{1}{\sqrt{3}},$$

and

$$\int_0^\infty R_{21}R_{30}\, r^3 dr \approx a_0\,.$$

A charged particle in a time-dependent damped oscillating electric field

A particle of charge q and mass m, initially in the ground state of a one-dimensional harmonic oscillator in the x direction and with angular frequency ω_0 is placed, at time $t = 0$, in a weak time-dependent and spatially-uniform electric field $E(t) = E_0\cos(\omega t)\exp\left(-t/\tau\right)$ along the x axis, with $E_0 > 0$ and $\tau > 0$ constants. Working at first order in perturbation theory, determine the probability of finding the system in an excited state n at time $t \gg \tau$.

The total Hamiltonian of the system is given by

$$H = H_0 + V(t)\,,$$

with H_0 the harmonic oscillator Hamiltonian, $V(t) = -q\,E(t)\,x$ and x the position operator in the x direction. At first order in perturbation theory, the transition probability from the initial state $|i\rangle \equiv |0\rangle$ at time $t = 0$ to a final state $|f\rangle \equiv |n\rangle$ at time t is given by $P_{0\to n} \simeq |c_n^{(1)}(t)|^2$ with

$$c_n^{(1)}(t) = \frac{i}{\hbar}\int_0^t \langle 0|V(t')|n\rangle e^{i\omega_{n0}t'}dt' = \frac{i}{\hbar}\left(-qE_0\right)\int_0^t \langle 0|x|n\rangle\cos\omega t' e^{-t'/\tau}e^{in\omega_0 t'}dt'\,,$$

and

$$\omega_{n0} = \frac{E_n - E_0}{\hbar} = n\omega_0\,, \qquad E_n = \hbar\omega_0\left(n + \frac{1}{2}\right), \qquad \langle 0|x|n\rangle = \sqrt{\frac{\hbar}{2m\omega_0}}\delta_{n,1}\,.$$

Expressing the cosine in the above expression for $c_1(t)$ in terms of exponentials, $\cos\omega t' = (e^{i\omega t'} + e^{-i\omega t'})/2$ and integrating, we obtain

$$\begin{aligned} c_1^{(1)}(t) &= -\frac{iqE_0}{\sqrt{2m\hbar\omega_0}}\frac{1}{2}\int_0^t \left(e^{i(\omega+\omega_0)t' - t'/\tau} + e^{i(-\omega+\omega_0)t' - t'/\tau}\right)dt' \\ &= \frac{iqE_0}{\sqrt{2m\hbar\omega_0}}\frac{1}{2}\left[\frac{1 - e^{i(\omega+\omega_0)t - t/\tau}}{i\left(\omega+\omega_0\right) - 1/\tau} + \frac{1 - e^{i(-\omega+\omega_0)t - t/\tau}}{i\left(-\omega+\omega_0\right) - 1/\tau}\right]. \end{aligned}$$

For $t \gg \tau$, this becomes

$$c_1^{(1)}(t) = \frac{iqE_0}{\sqrt{2m\hbar\omega_0}} \frac{1}{2} \left[\frac{1}{i\left(\omega + \omega_0\right) - 1/\tau} + \frac{1}{i\left(-\omega + \omega_0\right) - 1/\tau} \right] = \frac{iqE_0}{\sqrt{2m\hbar\omega_0}} \frac{\tau(i\omega_0\tau - 1)}{\left(\omega^2 - \omega_0^2\right)\tau^2 + 1 - 2i\omega_0\tau} \,.$$

The corresponding probability is then

$$P = |c_1^{(1)}|^2 = \frac{\left(qE_0\tau\right)^2}{2mh\omega_0} \frac{\left(\omega_0\tau\right)^2 + 1}{\left[\left(\omega - \omega_0\right)^2 \tau^2 + 1\right]^2 + 4\omega_0^2\tau^2} \,.$$

8.9 Exercises

1. **A charged particle in harmonic oscillator and time-dependent Gaussian electric field**

   A charged particle, initially in the ground state of a one-dimensional harmonic oscillator with angular frequency ω, is placed at time $t = 0$ in a time-dependent electric field

   $$E(t) = E_0\, e^{-t^2/\tau^2}\,,$$

   with E_0 and τ constants. Assuming E_0 to be small and working at first order in perturbation theory, determine the probability of finding the particle in a given excited state at $t = +\infty$. Analyze explicitly the limiting cases $\omega\tau \ll 1$ and $\omega\tau \gg 1$.

2. **A spatial and time-dependent perturbation to the harmonic oscillator**

   A particle, initially ($t < 0$) in the ground state of a one-dimensional harmonic oscillator, is subject at time $t = 0$ to a perturbation

   $$V(x,t) = V_0\, x^3 e^{-t/\tau}\,,$$

   with E_0 and τ constants. Assuming V_0 to be small and working at first order in perturbation theory, determine the probability of finding the particle in a given excited state at $t = +\infty$.

3. **A harmonic oscillator with moving center**

   The center of a one-dimensional harmonic oscillator, initially ($t < 0$) in its ground state, starts moving with constant velocity v_0 at $t = 0$ and do so until time T, where it stops completely. Assuming v_0 small and working at first order in perturbation theory, determine the probability of finding the particle in its first excited state at $t = +\infty$.

4. **A rotor under a time-dependent exponential perturbation**

   A certain type of rotor with orbital angular momentum $\ell = 2$ is governed by a Hamiltonian

   $$H = \frac{3a}{2\hbar} L_z - \frac{a}{\hbar^2}\left(L_x^2 + L_y^2\right)\,,$$

   with $a > 0$ a constant and L_i the i-th component of the angular momentum operator.

   a) Determine the energy spectrum of the system.

   b) Determine the average energy obtained in measurements if the system is in a state with wave function

   $$\varphi(\theta,\phi) = A\left(\sin\theta\cos\theta\cos\phi + \sin^2\theta\sin\phi\cos\phi\right)\,,$$

   with θ the polar angle, ϕ the azimuthal angle, and A a normalization constant,

   c) Assume the particle to be initially in the lowest energy state with $\ell = 2$ for $t < 0$. At $t = 0$, a time-dependent perturbation

   $$V(t) = \frac{\lambda}{\hbar}\, L_x e^{-t/\tau}$$

   is applied, with $\tau > 0$ and $\lambda > 0$ constants. Using first-order time-dependent perturbation theory, calculate the transition probabilities to possible excited states at $t \to \infty$.

5. **A hydrogen atom in an exponential electric field** ✐

   A hydrogen atom initially in its ground state ($t < 0$) is subject at time $t = 0$ to an electric field

   $$E(t) = E_0\, e^{-t/\tau}$$

   along the z-axis, with E_0 and τ constants. Working at first order in perturbation theory, calculate the probability that the atom transitions to the state $|n = 2, l = 1, m_l\rangle$ at time $t_s \gg \tau$. The energy of the state $|n, l, m_l\rangle$ is given by $E_n = -E_R/n^2$, with E_R the Rydberg energy. The matrix element of the position operator in the z-direction is given by

   $$\langle n = 1, l = 0, m_l = 0|z|n = 2, l = 1, m_l\rangle \simeq 0.745 a_0 \delta_{m_l,0}\,,$$

   with a_0 the Bohr radius of the atom.

6. **A hydrogen atom in a time-dependent Cauchy-like electric field** ✐

   A hydrogen atom, initially in its ground state ($t < 0$), is placed at time $t = 0$ in a time-dependent electric field

   $$E(t) = E_0\, \frac{\tau}{\tau^2 + t^2}$$

   along the z axis, with E_0 and τ constants. Assuming E_0 small and working at first order in perturbation theory, determine the probability of finding the particle in the first excited state at $t = +\infty$.
   Hint: You may want to use the results of Problem 2 in the Math basics exercise sheet.

7. **A spin-1/2 particle in time-varying magnetic field** ✐

   A spin-1/2 particle is in the presence of a time-varying magnetic field $\vec{B} = (0, 0, B(t))$, with $B(t) = B_0 + B_1 \cos(\omega t)$, and $B_1 \ll B_0$. The associated interaction Hamiltonian is given by

   $$H(t) = -\gamma S_z B(t)\,.$$

   Working at first-order perturbation theory, determine the possible transitions between the stationary states of the unperturbed Hamiltonian, $H_0 = -\gamma S_z B_0$. Can we conclude anything about the exact transition probabilities between these eigenstates?

8. **A time-dependent perturbation on a composite system** ✐

   Consider a composite quantum system of two spin-1/2 particles. For $t < 0$, the Hamiltonian is spin-independent and can be taken to be zero. At $t = 0$, an interaction is suddenly turned on, and for $t > 0$, the Hamiltonian becomes $H = \left(4\Delta/\hbar^2\right) \vec{S}_1 \cdot \vec{S}_2$, with $\Delta/\hbar^2 \ll 1$. Assuming the system is in the state $|+-\rangle$ at $t = 0$, determine the probability of finding the system in each of the basis states $|++\rangle\,, |+-\rangle\,, |-+\rangle\,, |--\rangle$ as a function of time in 2 ways: a) by solving the time-dependent Schrödinger equation exactly, b) using first-order time-dependent perturbation theory. Under what condition does this approximation reproduce the exact result?

9. **A time-dependent perturbation to a two-level system** ✐

   Consider a perturbation to a two-level system with matrix elements

   $$H'_{mn} = H'_{nm} = \frac{\alpha}{\sqrt{\pi}\tau} e^{-(t/\tau)^2}\,, \qquad H'_{mm} = H'_{nn} = 0\,,$$

   with τ and α positive constants with appropriate units.

a) According to first-order perturbation theory, if the system starts off in the state $|m\rangle$ at $t = -\infty$, what is the probability that it will be found in the state $|n\rangle$ at $t = \infty$?

b) Taking into account that in the limit $\tau \to 0, H'_{mn} = \alpha\,\delta(t)$, compute the $\tau \to 0$ limit of your expression from part a) and compare it to the exact result $P_{m\to n} = \sin^2\left(\frac{\alpha}{\hbar}\right)$.

10. **Selection rules at work**

An electron in the $n = 3, l = 0, m = 0$ quantum state of a hydrogen atom undergoes a sequence of electric dipole transitions to eventually reach the ground state.

a) Identify the available decay pathways. Represent the possible transitions as $|300\rangle \to |nlm\rangle \to |n'l'm'\rangle \to \ldots \to |100\rangle$.

b) Given an ensemble of atoms in this state, determine the proportion of them that decay via each possible route.

Hint: Once the electron makes the first transition, it is no longer in the $|300\rangle$ state, so only the initial transition step in each pathway affects the lifetime calculation. Useful integrals:

$$\int Y_{10} Y_{00} \cos\theta \sin\theta d\theta d\phi = \frac{1}{\sqrt{3}}\,, \qquad \int (Y_{1\pm1})^* Y_{00} \sin^2\theta \cos\phi d\theta d\phi = \frac{\mp 1}{\sqrt{6}}\,,$$

and

$$K \equiv \int_0^\infty R_{21} R_{30} r^3 dr = \frac{2^7 3^4}{5^6}\sqrt{2}a\,.$$

11. **Old Problem, New Tool: Fermi's Take on Scattering**

A particle of momentum $p = \hbar k > 0$ in one spatial dimension approaches a steep-sided potential well

$$V(x) = \begin{cases} -V_0\,, & \text{for } |x| < a\,,\ V_0 > 0 \\ 0\,, & \text{otherwise}\,, \end{cases}$$

from $x = -\infty$. Assuming periodic boundary conditions to ensure normalizable incoming and outgoing wavefunctions, and applying the Fermi golden rule, show that the probability for the particle to be reflected by the well is approximately

$$P \simeq \frac{2mV_0^2}{4p^2}\sin^2(2ka)\,.$$

Compare this result with the limit $p^2/(2m) \gg V_0$ of the exact reflection probability for this potential well.

12. **An infinite square well with slowly expanding walls**

Consider a particle confined in a one-dimensional infinite potential well, with the right boundary moving at constant velocity v. The width of the well at time t is given by $L(t) = L_0 + vt$, where L_0 is the initial width of the well at $t = 0$.

a) Show that this system admits an exact solution, with time-dependent wave functions

$$\varphi_n(x,t) = \sqrt{\frac{2}{L(t)}}\sin\left(\frac{n\pi x}{L(t)}\right)\exp\left[\frac{i}{2\hbar L(t)}\left(mvx^2 - 2E_n^0 L_0\, t\right)\right]\,,$$

with $E_n^0 = \frac{n^2\pi^2\hbar^2}{2mL_0^2}$ the energy levels at $t = 0$.

b) Assuming the particle starts in the ground state of the well at $t = 0$, express the initial wavefunction $\varphi(x, 0)$ as a linear combination of the exact time-dependent states $\varphi_n(x, t)$, $\varphi(x, 0) = \sum_{n=1}^{\infty} c_n \varphi_n(x, 0)$, determining the coefficients c_n.

c) Analyze the behavior of the system in the limit $v \to 0$. Show that in this limit, the system remains in the instantaneous ground state as expected from the adiabatic theorem. Discuss the relevant time scales and the condition for adiabaticity.

d) Compare the exact solution with the adiabatic approximation. Examine both the amplitude and phase of the wavefunction. Identify whether a geometric (Berry) phase is present, and comment on its physical meaning in this setting.

13. **A driven harmonic oscillator**

Consider a one-dimensional quantum harmonic oscillator with mass m and angular frequency ω, subjected to a time-dependent external force. The Hamiltonian of the system is

$$H(t) = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + \frac{1}{2}m\omega^2 x^2 - m\omega^2 x f(t)\,,$$

with $f(t)$ a real-valued function of time with dimensions of length, and $f(0) = 0$.

a) Determine the instantaneous eigenvalues $E_n(t)$ and eigenfunctions $\varphi_n(x, t)$ of the Hamiltonian $H(t)$. Express the eigenfunctions in terms of the known eigenfunctions of the unforced harmonic oscillator.

b) State the condition under which the time evolution of the system can be treated adiabatically. Express this condition using $f(t)$ and/or its time derivative $\dot{f}(t)$.
Show that, under this approximation, the function

$$x_{\mathrm{cl}}(t) = \int_0^t f(t')\omega \sin[\omega(t - t')]\, dt'$$

approximately equals $f(t)$.
Show furthermore that $x_{\mathrm{cl}}(t)$ satisfies the classical equation of motion for a driven harmonic oscillator, assuming appropriate initial conditions at $t = 0$.

c) Assume that the state of the system at time t can be expressed as $\varphi(x, t) = \sum_n c_n(t)\varphi_n(x, t)$. Derive the equations governing the time evolution of the coefficients $c_n(t)$. Simplify the expression using the properties of the annihilation and creation operators.

14. **Tritium decay in the sudden approximation**

A tritium atom ^{3}H can decay into a helium ion ^{3}He$^+$ via a rapid transition in which a neutron converts into a proton that remains in the nucleus, while an electron is ejected from the atom at high speed. Using the sudden approximation, determine the probability for an electron initially in the ground state of ^{3}H to be found in the ground state of the ^{3}He$^+$after the decay.

15. **Sudden quench in a two-body harmonic system**

Two particles of mass m move along the x-axis and interact via an ideal spring with force constant k. Assume the system is initially in its ground state with energy E_0, and the spring constant is suddenly halved. What is the probability that a subsequent measurement of the energy finds the system still in the ground state of the new potential?

CHAPTER 9

RELATIVISTIC CORRECTIONS TO THE HYDROGEN ATOM

A great deal of my work is just playing with equations and seeing what they give. I don't suppose that applies so much to other physicists; I think it's a peculiarity of myself that I like to play about with equations, just looking for beautiful mathematical relations which maybe don't have any physical meaning at all. Sometimes they do...

PAUL DIRAC, AHPQ INTERVIEW, 1902.

Almost all rules of spectroscopy follow from the symmetry of the problem.

EUGENE WIGNER IN HIS BOOK "GROUP THEORY AND APPLICATION", 1931.

While the physics encoded in the Pauli Hamiltonian (3.208) has played a crucial role in many of the examples considered in this course, the associated Schrödinger equation is unable to account for special relativity effects, expected to become increasingly important as particles approach the speed of light.

9.1 The Klein-Gordon equation

In order to incorporate relativistic effects, let us start by considering the relativistic dispersion relation

$$E^2 = \vec{p}^2 c^2 + m^2 c^4 \,, \tag{9.1}$$

and make use of the correspondence principle $E \to i\hbar\partial_t$ and $\vec{p} \to \frac{\hbar}{i}\vec{\nabla}^2$ to postulate a relativistic Schrödinger-like equation for the electron,

$$(i\hbar\partial_t)^2 \, \varphi(\vec{x},t) = \left(m^2c^4 - c^2\hbar^2\Delta\right) \varphi(\vec{x},t) \,, \tag{9.2}$$

or more compactly

$$\left(\Box + \frac{m^2c^2}{\hbar^2}\right) \varphi(\vec{x},t) = 0 \,, \tag{9.3}$$

with $\Box \equiv \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \Delta$ the so-called d'Alembertian or "box" operator. This wave equation, first discovered by Schrödinger itself, is called the *Klein-Gordon equation*, honoring the authors

first publishing it. As compared to its non-relativistic counterpart, it allows for the existence of negative energy solutions

$$\varphi(\vec{x},t) = e^{-\frac{i}{\hbar}Et+\frac{i}{\hbar}\vec{p}\vec{x}}\,, \qquad E = -\sqrt{mc^2+p^2c^4} \leq -mc^2\,. \tag{9.4}$$

9.2 The Dirac equation

To circumvent the existence of negative energy solutions in the Klein-Gordon, Dirac tried to rewrite the relativistic dispersion relation as the square of a linear function of the momentum,

$$c^2\vec{p}^2 + m^2c^4 = \left(c\,\vec{\alpha}\cdot\vec{p}+\beta mc^2\right)^2 = \left(c\,\alpha_1 p_1 + c\,\alpha_2 p_2 + c\,\alpha_3 p_3 + \beta mc^2\right)^2\,. \tag{9.5}$$

Expanding the right-hand side of this equation and equating coefficients in both sides, provides the constraint relations,

$$\beta^2 = 1\,, \qquad \{\alpha_i,\alpha_j\} = 2\delta_{ij}\,, \qquad \{\alpha_i,\beta\} = 0\,. \tag{9.6}$$

A simple inspection of these expressions reveals that the quantities α_i and β cannot be just numbers, since in that case they should be identically zero. Note also that, albeit the relations for α_i are satisfied by the Pauli matrices already familiar to us, the last anticommutation relation with β spoils any choice of a 2×2 matrix with $\beta^2 = 1$. The above relations can be satisfied, however, by four-by-four hermitian matrices,

$$\vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix}, \qquad \beta = \begin{pmatrix} \mathbb{I} & 0 \\ 0 & -\mathbb{I} \end{pmatrix}, \tag{9.7}$$

with $\vec{\sigma}$ the Pauli matrices. Roughly speaking, flipping signs requires increasing the square roots of unity by doubling the dimension. Note also that the cyclicity of the trace operator, $\mathrm{tr}(ABC) = \mathrm{tr}(CAB)$, requires α_i and β to be traceless,

$$\begin{aligned} \mathrm{tr}\,\alpha_i = \mathrm{tr}\,\alpha_i\beta^2 = \mathrm{tr}\,(\alpha_i\beta)\,\beta = \mathrm{tr}\,\beta\,(\alpha_i\beta) = -\,\mathrm{tr}\,\beta\,(\beta\alpha_i) = -\,\mathrm{tr}\,\alpha_i \quad &\longrightarrow \quad \mathrm{tr}\,\alpha_i = 0\,, \\ \mathrm{tr}\,\beta = \mathrm{tr}\,(\beta\alpha_1)\,\alpha_1 = \mathrm{tr}\,\alpha_1\,(\beta\alpha_1) = -\,\mathrm{tr}\,\alpha_1\,(\alpha_1\beta) = -\,\mathrm{tr}\,\beta \quad &\longrightarrow \quad \mathrm{tr}\,\beta = 0\,, \end{aligned} \tag{9.8}$$

which, combined with the conditions $\alpha_i^2 = \beta^2 = 1$, implies that half of the eigenvalues of α_i and β are $+1$ and the other half are -1. This is only possible for even-dimensional matrices.

Accepting the above paradigm shift, we obtain the so-called Dirac equation for a free-particle

$$i\hbar\frac{\partial\varphi}{\partial t} = H_{\mathrm{Dirac}}\varphi\,, \qquad H_{\mathrm{Dirac}} = c\,\vec{\alpha}\cdot\vec{p}+\beta mc^2\,, \tag{9.9}$$

where we have introduced a four-component column vector φ or *Dirac spinor* composed by two two-component Pauli spinors u and v,

$$\varphi = \begin{pmatrix} u \\ v \end{pmatrix}, \qquad u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \qquad v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}. \tag{9.10}$$

In order to deal with Lorentz invariance, it is convenient to define the so-called γ *matrices in the Dirac representation*,[1]

$$\gamma^j = \beta\alpha_j = \begin{pmatrix} 0 & \sigma_j \\ -\sigma_j & 0 \end{pmatrix}, \qquad \gamma^0 = \beta = \begin{pmatrix} \mathbb{I} & 0 \\ 0 & -\mathbb{I} \end{pmatrix}. \tag{9.12}$$

In terms of these, the Dirac equation (9.9) and the consistency relations (9.6), written in implicit Einstein summation convention, become explicitly covariant,

$$(i\hbar\gamma^\mu\partial_\mu - mc)\,\varphi = 0\,, \qquad \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2\eta^{\mu\nu}\mathbb{I}_{4\times4}\,, \tag{9.13}$$

with the Greek indexes taking values $0, 1, 2, 3$ and $\eta^{\mu\nu}$ the components of the Minkowski tensor.

The interpretation of Dirac equation

From a quantum mechanical perspective, the physical meaning of the two additional dimensions introduced by the 4×4 γ matrices remains unclear. Indeed, when constructing the Dirac equation, the original wave function becomes a vector or 2-spinor φ, encoding much more information than the initial spin values. As a matter of fact, this equation describes two particles, or, more precisely, a particle and an antiparticle, linked together into a unique and indivisible structure: a field. In spite of this, the procedure does not introduce new degrees of freedom. In particular, the two components in u are related by a symmetry to those in v, as becomes explicit when writing the Dirac equation in the Weyl representation (9.11).

9.2.1 Spin operator from consistency

Unlike the Schrödinger equation, the Dirac equation *predicts* the existence of the spin. The need for this operator arises from the lack of conservation of the orbital angular momentum $\vec{L} = \vec{X} \times \vec{P}$ in relativistic systems,

$$[H, L_i] = \left[c\vec{\alpha}\vec{P} + \beta mc^2, \epsilon_{ijk}X_jP_k\right] = \frac{\hbar}{i}c\epsilon_{ijk}\alpha_jP_k \neq 0\,, \tag{9.14}$$

where we have taken into account that $[P_l, X_j] = \frac{\hbar}{i}\delta_{lj}$. A conserved operator $\vec{J} = \vec{L} + \vec{S}$ can be, however, constructed by observing that the quantity

$$S_i = \frac{\hbar}{4i}\epsilon_{ijk}\alpha_j\alpha_k = \frac{\hbar}{2}\begin{pmatrix} \sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix} \tag{9.15}$$

[1] It can be shown that, up to unitary transformations $\alpha_i \to U\alpha_iU^{-1}, \beta \to U\beta U^{-1}$ with $UU^\dagger = \mathbb{I}$, (9.12) is the unique irreducible unitary representation of the Dirac algebra (9.6). In some contexts, it is convenient to define the γ *matrices in the related Weyl representation*

$$\gamma^j = \begin{pmatrix} 0 & \sigma_j \\ -\sigma_j & 0 \end{pmatrix}, \qquad \gamma^0 = \begin{pmatrix} 0 & \mathbb{I} \\ \mathbb{I} & 0 \end{pmatrix}. \tag{9.11}$$

satisfies

$$[H,S_i]=\frac{\hbar}{4i}\epsilon_{ijk}\left[c\vec{\alpha}\vec{P},\alpha_j\alpha_k\right]=\frac{\hbar}{4i}\epsilon_{ijk}cP_l\left(\{\alpha_l,\alpha_j\}\,\alpha_k-\alpha_j\,\{\alpha_l,\alpha_k\}\right)$$
$$=\frac{\hbar}{4i}\epsilon_{ijk}cP_l\left(2\delta_{lj}\alpha_k-2\delta_{lk}\alpha_j\right)=\frac{\hbar}{i}c\epsilon_{ijk}\alpha_kP_k=-\frac{\hbar}{i}c\epsilon_{ijk}P_j\alpha_k\,, \tag{9.16}$$

where we have made use of the property $[A,BC]=ABC+BAC-BAC-BCA=\{A,B\}C-B\{A,C\}$ and the antisymmetry of ϵ_{ijk}.

9.2.2 Interactions and Lorentz covariant form

To introduce interactions with the electromagnetic field, we make use of the minimal coupling prescription

$$\vec{p}\to\frac{\hbar}{i}\vec{\nabla}-\frac{e}{c}\vec{A}\,,\qquad\qquad E\to i\hbar\partial_t-V\,, \tag{9.17}$$

in Eq. (9.9), obtaining

$$\left(i\hbar\frac{\partial}{\partial t}-e\phi\right)\varphi=c\alpha_i\left(\frac{\hbar}{i}\partial_i-\frac{e}{c}A_i\right)\varphi+\beta mc^2\varphi\,. \tag{9.18}$$

Multiplying this expression by $\beta/(\hbar c)$ and collecting coordinates, energy-momenta, and gauge potentials into 4-vectors

$$x^\mu=(ct,\vec{x})\,,\qquad\partial_\mu=\left(\frac{1}{c}\partial_t,\vec{\nabla}\right)\,,\qquad p^\mu=\left(\frac{1}{c}E,\vec{p}\right)\,,\qquad A^\mu=(\phi,\vec{A})\,, \tag{9.19}$$

we obtain

$$\boxed{\left(\gamma^\mu\left(i\partial_\mu-\frac{e}{\hbar c}A_\mu\right)-\frac{c}{\hbar}m\right)\varphi=0\,,} \tag{9.20}$$

with

$$\gamma^\mu=(\beta,\beta\vec{\alpha})\,,\qquad\qquad\{\gamma^\mu,\gamma^\nu\}=2g^{\mu\nu}\mathbb{I}\,, \tag{9.21}$$

and

$$\eta^{\mu\nu}=\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix} \tag{9.22}$$

the *Minkowski metric*. Note that since $(\beta\alpha_i)^\dagger=\alpha_i\beta=-\beta\alpha_i$, the gamma matrices γ^μ in this expression are unitary, but not Hermitian for $\mu\neq 0$,

$$(\gamma^\mu)^\dagger=(\gamma^\mu)^{-1}\,. \tag{9.23}$$

This non-Hermiticity is related to the fact that Lorentz-boosts are given by non-unitary transformations, as required by consistency of probability density interpretations with Lorentz contractions.

9.2.3 The non-relativistic limit of the Dirac equation

Particularizing the above results for hydrogenic atoms and assuming the potentials $V = e\phi$ and $\vec{A}$ to be time-independent, we make the following stationary ansatz for energy eigenfunctions φ

$$\varphi(\vec{x},t) = \begin{pmatrix} u(\vec{x}) \\ v(\vec{x}) \end{pmatrix} e^{-\frac{i}{\hbar}Et} . \tag{9.24}$$

Introducing the gauge-invariant physical momentum

$$\vec{\Pi} \equiv \vec{P} - \frac{e}{c}\vec{A} = \frac{\hbar}{i}\vec{\nabla} - \frac{e}{c}\vec{A} , \tag{9.25}$$

and taking into account the representation (9.7), the stationary Dirac equation becomes a coupled system of differential equations,

$$\begin{pmatrix} E'-V & 0 \\ 0 & E'+mc^2-V \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 & c\vec{\sigma}\vec{\Pi} \\ c\vec{\sigma}\vec{\Pi} & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} , \tag{9.26}$$

with $E' \equiv E - mc^2 > 0$. While this system can be solved exactly for electrons in hydrogenic atoms, we will concentrate in what follows in the non-relativistic weak-field limit

$$V, \vec{A}, \vec{p}, E' \ll mc^2 . \tag{9.27}$$

Solving for v in the second equation in (9.26), we obtain

$$v = \frac{f(\vec{x})}{2mc}\vec{\sigma}\vec{\Pi}u , \tag{9.28}$$

with

$$f(\vec{x}) = \frac{1}{1+\frac{E'-V(\vec{x})}{2mc^2}} \simeq 1 - \frac{E'-V(\vec{x})}{2mc^2} + \dots . \tag{9.29}$$

Inserting this result into the first equation yields

$$(E'-V)u = \frac{1}{2m}(\vec{\sigma}\vec{\Pi}f(\vec{x})\vec{\sigma}\vec{\Pi})u , \tag{9.30}$$

which can be interpreted as a non-relativistic Schrödinger equation for an effective Hamiltonian

$$H_{\text{non-rel}} = \frac{1}{2m}(\vec{\sigma}\vec{\Pi}f(\vec{x})\vec{\sigma}\vec{\Pi}) + V \tag{9.31}$$

and non-relativistic energy E'. For a centrally symmetric potential like the one under consideration,

$$V(\vec{x}) = V(r) , \qquad \frac{\partial f(r)}{\partial x_i} = f'(r)\frac{x_i}{r} , \tag{9.32}$$

and taking into account that $\sigma_i\sigma_j = \delta_{ij} + i\epsilon_{ijk}\sigma_k$, we obtain

$$H_{\text{non-rel}} = \frac{f}{2m}\Pi_i\Pi_j\sigma_i\sigma_j + V = \frac{1}{2m}\left(f\Pi_i\Pi_j + \frac{\hbar}{i}\frac{\partial f}{\partial x_i}\Pi_j\right)(\delta_{ij} + i\epsilon_{ijk}\sigma_k) + V , \tag{9.33}$$

The decomposition of the first term in the right-hand side of the last expression according to the (anti)symmetry in ij yields four terms,

$$\frac{f}{2m}\Pi_i\Pi_j\sigma_i\sigma_j = \frac{f}{2m}\Pi_i\Pi_j\delta_{ij} + \frac{if}{2m}\epsilon_{ijk}\sigma_i\Pi_j\Pi_k - \frac{i\hbar}{2m}f'\frac{x_i}{r}\Pi_j\delta_{ij} + \frac{\hbar}{2m}f'\frac{x_i}{r}\Pi_j\epsilon_{ijk}\sigma_k\,, \quad (9.34)$$

which we now discuss in detail:

- *Angular momentum term*

$$\frac{f}{2m}\Pi_i\Pi_j\delta_{ij} = \frac{f}{2m}\left(\vec{p}^2 - \frac{e}{c}(\vec{p}\vec{A} + \vec{A}\vec{p}) + \frac{e^2}{c^2}\vec{A}^2\right)\,. \quad (9.35)$$

  For a constant magnetic field $\vec{B} = \nabla \times \vec{A}$, we choose a vector potential

$$A_i = -\frac{1}{2}\epsilon_{ijk}x_jB_k \quad (9.36)$$

  satisfying the Coulomb gauge condition

$$\vec{\nabla}\cdot\vec{A} = \partial_iA_i = 0\,, \quad (9.37)$$

  and neglect the last term in Eq. (9.35), since this is quadratic in $\vec{B}$. In this gauge, we have

$$\partial_iA_i\psi = (\partial_iA_i)\,\psi + A_i\partial_i\psi = A_i\partial_i\psi \quad (9.38)$$

  and

$$-2\frac{e}{c}\vec{A}\vec{p} = -2\frac{e}{c}\left(-\frac{1}{2}\epsilon_{ijk}x_jB_k\right)p_i = -\frac{e}{c}\vec{B}\vec{L}\,. \quad (9.39)$$

  Taking also into account that $E' - V = p^2/2m + \ldots$ and therefore $f \simeq 1 - p^4/(4m^2c^2) + \ldots$, the leading contribution in (9.35) yields the usual kinetic term, a $\vec{p}^4$ contribution and an angular momentum coupling,

$$\frac{f}{2m}\vec{\Pi}^2 \simeq \frac{\vec{p}^2}{2m} - \frac{\vec{p}^4}{8m^3c^2} - \frac{e}{2mc}\vec{L}\cdot\vec{B} + \mathcal{O}(B^2)\,. \quad (9.40)$$

- *Gyromagnetic ratio term*

$$\frac{if}{2m}\epsilon_{ijk}\sigma_i\Pi_j\Pi_k = -\frac{if}{2m}\frac{e}{c}\epsilon_{ijk}\sigma_k\,(p_iA_j + A_ip_j) = -\frac{ife}{2mc}\epsilon_{ijk}\sigma_k\,(p_iA_j - A_jp_i)\,. \quad (9.41)$$

  Since $(p_iA_j - A_jp_i)\,\psi = (p_iA_j)\,\psi$ this expression, at leading order $f \simeq 1$, becomes

$$-\frac{ife}{2mc}\epsilon_{ijk}\hbar\,(\partial_iA_j)\,\sigma_k \simeq -\frac{e\hbar}{2mc}(\vec{nabla}\times\vec{A})_k\sigma_k = -\frac{e\hbar}{2mc}\vec{B}\vec{\sigma} = -\frac{ge}{2mc}\vec{B}\vec{S}\,, \quad (9.42)$$

  where we have defined a g-factor $g = 2$.

- *Darwin term*

$$\frac{1}{2m}\frac{\hbar}{i}f'\frac{x_i}{r}\Pi_i \approx \frac{\hbar}{4im^2c^2}\frac{V'}{r}\left(\vec{x}\vec{p} - \frac{e}{c}\vec{x}\vec{A}\right) \approx -\frac{\hbar^2}{4m^2c^2}\partial_iV\partial_i\,. \quad (9.43)$$

where we used $\frac{x_i}{r}V' = \partial_i V$ and dropped the vector potential contribution, as this is quadratic in the electromagnetic fields. Since, for real eigenfunctions u, the expectation values of the Hamiltonians agree by partial integration,

$$\int u\partial_i V\partial_i u = \frac{1}{2}\int \partial_i V\partial_i\left(u^2\right) = -\frac{1}{2}\int (\Delta V)u^2\,, \tag{9.44}$$

the expression (9.43) is equivalent to the Darwin term

$$H_{\text{Darwin}} = \frac{\hbar^2}{8m^2c^2}\Delta V\,. \tag{9.45}$$

- *Spin-orbit coupling*

$$\frac{\hbar}{i}\frac{f'}{r}i\epsilon_{ijk}x_i p_j\sigma_k = \frac{f'}{r}\hbar L_k\sigma_k \approx \frac{dV}{dr}\frac{1}{mc^2 r}\vec{L}\vec{S}\,. \tag{9.46}$$

Collecting all relevant contributions, we recover the Pauli equation

$$H_{\text{Pauli}} = \frac{\vec{p}^2}{2m} + V - \frac{e}{2mc}\vec{B}(\vec{L}+2\vec{S}) + \frac{dV}{dr}\frac{\vec{L}\vec{S}}{2m^2c^2 r}\,, \tag{9.47}$$

which contains the spin-orbit coupling and explains the observed gyromagnetic ratio. In addition, we find the relativistic energy correction and the Darwin term, which together with the above expression explain the complete fine structure of the energy levels of the hydrogen atom.

9.3 Fine structure of hydrogen

In situations in which the above relativistic corrections are small, it is safe to use perturbation theory. For the particular case of the hydrogen atom, we have for instance

$$\Delta H_{\text{R}} = -\underbrace{\frac{\vec{p}^4}{8m^3c^2}}_{p\simeq\alpha mc} \sim -\alpha^4 mc^2\,, \tag{9.48}$$

$$\Delta H_{\text{SO}} = \underbrace{\frac{1}{2}\frac{e^2}{m^2c^2}\frac{1}{r^3}\vec{L}\cdot\vec{S}}_{\vec{L}\cdot\vec{S}\sim\hbar^2,\ r\sim a_0=\frac{\hbar}{mc\alpha}} \sim \frac{e^2}{m^2c^2}\frac{\hbar^2}{a_0^3} \sim \alpha^4 mc^2\,, \tag{9.49}$$

$$\Delta H_{\text{Darwin}} = \frac{\pi}{2}\frac{e^2\hbar^2}{m^2c^2}\delta(\vec{r}) \sim \frac{e^2}{m^2c^2}\frac{\hbar^2}{a_0^3} \sim \alpha^4 mc^2\,, \tag{9.50}$$

where in the last expression we have taken into account that the δ-function will introduce a factor $|\varphi(0)|^2 \sim a_0^{-3}$ when computing expectation values, with $a_0 = \hbar/(\alpha mc)$ the Bohr radius. The above energy corrections are therefore suppressed by a factor $\alpha^2 \simeq 1/137^2 \sim 5\times 10^{-5}$ as compared the zeroth-order energies $H_0 \sim \alpha^2 mc^2$. This suggests to consider the *fine structure constant* α^2 as the expansion parameter in perturbation theory. In what follows we study each of these terms separately, combining them at the end in order to obtain the fine structure of the hydrogen atom.

9.3.1 Relativistic corrections

The relativistic contribution

$$\Delta H_{\rm R} = -\frac{\vec{p}^4}{8m^3c^2} \tag{9.51}$$

is insensitive to both angular momentum and spin,

$$[\Delta H_{\rm R}, L^2] = 0\,, \qquad [\Delta H_{\rm R}, L_z] = 0\,, \qquad [\Delta H_{\rm R}, S_z] = 0\,, \tag{9.52}$$

The first condition ensures that the perturbation matrix is diagonal in ℓ, the second guarantees that is diagonal in m_ℓ, and the third that it is diagonal in m_s,

$$\langle n\ell m_\ell m_s|\Delta H_{\rm R}|nl'm'_\ell m'_s\rangle = \delta_{ll'}\delta_{m_l m'_l}\delta_{m_s m'_s}\,. \tag{9.53}$$

allowing us to make use of the non-degenerate perturbation theory developed in the previous chapter. Also due to rotational symmetry, we can anticipate that the remaining non-vanishing relativistic corrections,

$$\Delta E^{(1)}_{n,l}\Big|_{\rm R} = \langle \Delta H_{\rm R}\rangle_{n,l} \equiv \langle n\ell m_\ell m_s|\Delta H_{\rm R}|n\ell m_\ell m_s\rangle\,, \tag{9.54}$$

will not depend on the quantum number m. To calculate the right-hand side of this expression, it is convenient to rewrite the relativistic perturbation $\Delta H_{\rm R}$ in terms of the unperturbed Hamiltonian and the Coulomb potential $V(r) = Ze^2/r$,

$$\Delta H_{\rm R} = -\frac{1}{2mc^2}[H_0 - V(r)]^2\,. \tag{9.55}$$

This way, we obtain

$$\Delta E^{(1)}_{n,l}\Big|_{\rm R} = -\frac{1}{2mc^2}\left[(E^{(0)}_n)^2 - 2E^{(0)}_n\langle V(r)\rangle_{n,l} + \langle V(r)^2\rangle_{n,l}\right]\,, \tag{9.56}$$

and we reduce the problem to the computation of the quantities $\langle 1/r\rangle_{n,l}$ and $\langle 1/r^2\rangle_{n,l}$ within the terms $\langle V(r)\rangle_{n,l}$ and $\langle V(r)^2\rangle_{n,l}$. The first of these expressions follows directly from the virial theorem, $2\langle T\rangle = -\langle V\rangle$, which gives $\langle E\rangle = \langle T\rangle + \langle V\rangle = \langle V\rangle/2$ and

$$\left\langle \frac{1}{r}\right\rangle_{n,l} = -\frac{1}{Z\alpha\hbar c}\langle V\rangle_{n,l} = -\frac{1}{Z\alpha\hbar c}2E^{(0)}_n = \frac{Z}{a_0}\frac{1}{n^2}\,. \tag{9.57}$$

To calculate $\langle 1/r^2\rangle_{n,l}$, there's a trick. In particular, this perturbation can always be absorbed into the angular momentum term of H_0,

$$\frac{\hbar^2}{2m}\frac{l(l+1)}{r^2} + \frac{\lambda}{r^2} \equiv \frac{\hbar^2}{2m}\frac{l'\,(l'+1)}{r^2}\,. \tag{9.58}$$

This property allows us to reuse the exact result obtained in Quantum Physics I with the simple formal replacement $l \to l'(\lambda)$,

$$E\left(l'(\lambda)\right) = -\frac{mc^2(Z\alpha)^2}{2\left(k + l' + 1\right)^2}\,. \tag{9.59}$$

Expanding this result around $\lambda = 0$, we obtain

$$E\left(l'(\lambda)\right) = E_n^{(0)} + (Z\alpha)^2 mc^2 \left[\frac{1}{(k+l'+1)^3}\frac{dl'}{d\lambda}\right]\Bigg|_{\lambda=0} \lambda + \ldots = E_n^{(0)} + \frac{Z^2}{a_0^2}\frac{2\lambda}{n^3(2l+1)} + \ldots \tag{9.60}$$

from which we can directly read the correction

$$\left\langle \frac{1}{r^2} \right\rangle_{n,l} = \frac{Z^2}{a_0^2}\frac{1}{n^3(l+1/2)}\,. \tag{9.61}$$

Finally, by combining equations (9.56), (9.57), and (9.61), we arrive at the desired result

$$\Delta E_{n,l}^{(1)}\Big|_{\mathrm{R}} = -\frac{(Z\alpha)^2 mc^2}{2n^4}\left(\frac{n}{l+1/2} - \frac{3}{4}\right)\,, \tag{9.62}$$

whose dependence on l removes the degeneracy of H_0 in this quantum number. [2] As could be anticipated from our estimates, the result is proportional to the square of the fine-structure constant.

Note that in each degenerate subspace of fixed n and ℓ, ΔH_{rel} is a multiple of the identity, since the matrix elements are independent of the eigenvalues m and m_s of L_z and S_z. This allows to extend the computation in the uncoupled base (9.62) to the coupled basis $\{|n\ell j m_j\rangle\}$. Indeed, taking into account any state in the coupled basis is a superposition of orthonormal uncoupled basis states,

$$|n\ell j m_j\rangle = \sum_i c_i \left|n\ell m_\ell^i m_s^i\right\rangle\,, \tag{9.63}$$

with constant coefficients c_i satisfying $\sum_i |c_i|^2 = 1$, we have

$$\langle n\ell j m_j|\Delta H_{\mathrm{R}}|n\ell j m_j\rangle = \sum_{i,k} c_i^* c_k \langle n\ell m_\ell^i m_s^i|\Delta H_{\mathrm{R}}|n\ell m_\ell^k m_s^k\rangle = \underbrace{\sum_i |c_i|^2}_{=1} \left\langle n\ell m_\ell^i m_s^i\right|\Delta H_{\mathrm{R}}\left|n\ell m_\ell^i m_s^i\right\rangle\,,$$

or

$$\langle n\ell j m_j\,|\Delta H_{\mathrm{R}}|\,n\ell j m_j\rangle = \left\langle n\ell m_\ell^i m_s^i\right|\Delta H_{\mathrm{R}}\left|\,n\ell m_\ell^i m_s^i\right\rangle\,. \tag{9.64}$$

9.3.2 Spin-Orbit Coupling

The spin-orbit contribution to the Hamiltonian is

$$\Delta H_{\mathrm{SO}} = \frac{e}{2m^2c^2}\frac{1}{r^3}\vec{L}\cdot\vec{S}\,. \tag{9.65}$$

commutes with L^2 since this operator commutes with any L_i and S_i components. On top of that, as we showed explicitly in an example in Section 6.5, the interaction $\vec{L}\cdot\vec{S}$ commutes also with the total angular momentum J^2 and its thrid component J_z. It follows

[2] That degeneracy is due to the existence of a conserved Runge-Lenz vector, a symmetry that is broken in the presence of relativistic corrections.

then (cf. Eqs. (6.97) and (6.98)) that the spin-orbit contribution is diagonal in the level n degenerate subspace in the coupled basis $|n\ell j m_j\rangle$

$$\vec{L}\cdot\vec{S}|n,j,m_j;l\rangle = \frac{\hbar^2}{2}\begin{cases} -(l+1)|n,j,m_j;l\rangle\,, & j=l-\frac{1}{2} \quad (l\neq 0) \\ l|n,j,m_j;l\rangle\,, & j=l+\frac{1}{2} \end{cases} \tag{9.66}$$

and therefore

$$\Delta E^{(1)}_{n,j;l}\Big|_{\rm SO} = \frac{Ze^2}{2m^2c^2}\left\langle n\ell j m_j\left|\frac{1}{r^3}\vec{L}\cdot\vec{S}\right|n\ell j m_j\right\rangle = -\frac{Ze^2\hbar^2}{4m^2c^2}\left\{\begin{array}{c} -(l+1) \\ l \end{array}\right\}\left\langle n\ell j m_j\left|\frac{1}{r^3}\right|n\ell j m_j\right\rangle\,, \tag{9.67}$$

where, as in the previous expression, the upper and lower entries in $\{\cdot\}$ correspond respectively to $j=l-\frac{1}{2}$ with $l\neq 0$ and $j=l+\frac{1}{2}$. Note that when $l=0$, we have $\Delta E_2=0$ because there is no angular momentum for the spin to couple to.

To compute the remaining expectation value, we make use again of a nice trick. Introducing a new "radial momentum"

$$\tilde{p} = -i\hbar\left(\frac{\partial}{\partial r}+\frac{1}{r}\right)\,, \tag{9.68}$$

the radial part of the hydrogen atom Hamiltonian can be written us

$$H = -\frac{\hbar^2}{2m}\left(\frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r}\right)+\frac{\hbar^2 l(l+1)}{2mr^2}-\frac{Ze^2}{r} = \frac{\tilde{p}^2}{2m}+\frac{\hbar^2 l(l+1)}{2mr^2}-\frac{Ze^2}{r}\,. \tag{9.69}$$

Taking now into account that the resulting non-vanishing commutator among these two observables,

$$[\tilde{p},H] = -i\hbar\left(-\frac{\hbar^2 l(l+1)}{mr^3}+\frac{Ze^2}{r^2}\right)\,, \tag{9.70}$$

must vanish when evaluated on an energy eigenstate, $\langle[\tilde{p},H]\rangle_{n,j,l}=0$, we obtain

$$\left\langle n\ell j m_j\left|\frac{1}{r^3}\right|n\ell j m_j\right\rangle = \frac{Z}{a_0}\frac{1}{l(l+1)}\left\langle\frac{1}{r^2}\right\rangle_{n,j;l} = \left(\frac{Z}{a_0}\right)^3\frac{1}{l\left(l+\frac{1}{2}\right)(l+1)}\frac{1}{n^3} \quad (l\neq 0)\,. \tag{9.71}$$

Putting everything together and rewriting it in terms of j, we get

$$\Delta E^{(1)}_{n,j;l}\Big|_{\rm SO} = -\frac{(Z\alpha)^4 mc^2}{2}\left(\left\{\begin{array}{c} \frac{1}{j+1} \\ \frac{1}{j} \end{array}\right\}-\frac{3}{4n}\right)\frac{1}{n^3}\,. \tag{9.72}$$

which as anticipated is also proportional to the square of the fine-structure constant.

9.3.3 Darwin term

Finally, we determine the contribution of the Darwin term

$$\Delta H_{\rm Darwin} = \frac{\pi Ze^2\hbar^2}{2m^2c^2}\delta(\vec{r})\,, \tag{9.73}$$

which as anticipated affects only to $\ell=0$ states, since all wave functions with $l>0$ are vanishing at the origin. Note also, that, although for a given n there are two degenerate

orthogonal $\ell = 0$ states (one with electron spin up and one with electron spin down), there is no need to use degenerate perturbation theory since the Darwin term commutes with spin and is therefore diagonal in the two-dimensional subspace. Therefore, we have

$$\Delta E^{(1)}_{n00,\,\text{Darwin}} = \langle \varphi_{n00} \,|\Delta H_{\text{Darwin}}\,|\, \varphi_{n00}\rangle = \frac{\pi}{2}\frac{Ze^2\hbar^2}{m^2c^2}\,|\varphi_{n00}(0)|^2\,.$$

The normalised wave function $\varphi_{n00}(0)$ is of the form $\varphi_{n00}(\vec{r}) = R_{n0}(r)Y_{00}(\theta,\phi) = R_{n0}(r)/\sqrt{4\pi}$ with

$$R_{n0}(r) = -\sqrt{\left(\frac{2Z}{na_0}\right)^3 \frac{(n-1)!}{2n(n!)^3}}\, e^{-\frac{r}{na_0}} L^1_n\left(\frac{2r}{na_0}\right) \tag{9.74}$$

with L_n denoting the Laguerre polynomials, satisfying $L^1_n(x) = dL_n(x)/dx$ and $L_n(x) \approx n! - n!nx + \mathcal{O}\left(x^2\right)$, so that $L^1_n(0) = n!n$. The wave function at the origin then becomes

$$|\varphi_{n,l=0}(0)|^2 = \frac{Z^3}{\pi n^3 a_0^3}\,. \tag{9.75}$$

From this we get finally

$$\Delta E^{(1)}_{n00,\,\text{Darwin}} = \frac{e^2\hbar^2}{2m^2c^2}\frac{Z^4}{a_0^3 n^3} = \frac{(Z\alpha)^4 mc^2}{2n^3}\delta_{l0}\,. \tag{9.76}$$

9.3.4 Total effect

Collecting all the results in the previous section we have

$$\Delta E^{(1)}_{n,j} = \Delta E^{(1)}_{n,l}\Big|_{\text{R}} + \Delta E^{(1)}_{n,j;l}\Big|_{\text{SO}} + \Delta E^{(1)}_{n,l}\Big|_{\text{Darwin}} = \frac{(Z\alpha)^4 mc^2}{2}\left(\frac{3}{4n} - \frac{2}{2j+1}\right)\frac{1}{n^3}\,, \tag{9.77}$$

which can be more briefly written as

$$\Delta E^{(1)}_{nj} = -(Z\alpha)^4 mc^2 \cdot S_{n,j}, \quad \text{with} \quad S_{n,j} \equiv \frac{1}{2n^4}\left[\frac{n}{j+\frac{1}{2}} - \frac{3}{4}\right]. \tag{9.78}$$

At this point several comments are in order:

- The absence of a l-dependence in Eq. (9.78) reflects the fact that only the total angular momentum $\vec{J} = \vec{L} + \vec{S}$ commutes with the Dirac Hamiltonian, i.e. neither $\vec{L}$ nor $\vec{S}$ are separately conserved. With $\vec{J}$ a symmetry of the theory, the energy shifts are also m_j independent.
- Coincidentally, the final expression is also valid for $\ell = 0$ states since the contribution of the Darwin term present only when $l = 0$ coincides with formally with result of the spin-orbit coupling when $l = 0, j = 1/2$.
- For fixed n, multiplets with the same value of j remain degenerate, regardless of ℓ.

- Since

$$\frac{n}{j+\frac{1}{2}} \geq \frac{n}{j_{\max}+\frac{1}{2}} = \frac{n}{\ell_{\max}+\frac{1}{2}+\frac{1}{2}} = \frac{n}{n} = 1 \quad \rightarrow \quad \frac{n}{j+\frac{1}{2}} - \frac{3}{4} \geq \frac{1}{4} \quad \rightarrow \quad S_{n,j} > 0\,, \tag{9.79}$$

  all the energy shifts are down, being more sizable for lower values of j. As n increases splittings fall off like n^{-3}.

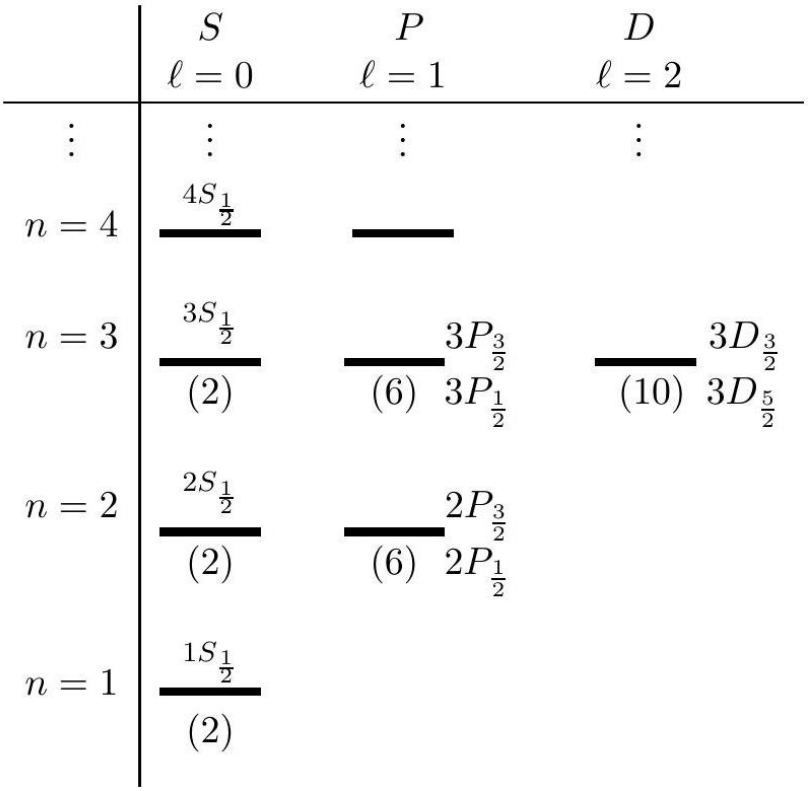
The levels of the hydrogen states resulting from the sum of the zeroth contribution and the fine structure contribution,

$$E_{n\ell j m_j} = -\frac{e^2}{2a_0}\frac{1}{n^2}\left[1 + \frac{\alpha^2}{n^2}\left(\frac{n}{j+\frac{1}{2}} - \frac{3}{4}\right)\right], \tag{9.80}$$

are customarily denoted as

$$nL_j$$

with specific capital letters $L = S, P, D, F, \ldots$ associated to the various values of the orbital angular momentum $l = 0, 1, 2, 3, \ldots$. For instance, the fundamental state is labeled as $1S_{1/2}$, the doublet P states corresponding to $l = 1$ are ${}^2P_{1/2}$ and ${}^2P_{3/2}$ states, and the doublet D states corresponding to $l = 2$ are ${}^2D_{3/2}$ and ${}^2D_{5/2}$.



The associated number of states is indicated in parenthesis.

9.4 Hyperfine structure and the 21 cm line

The spin of the electron interacts with the nuclear magnetic field through the hyperfine Hamiltonian

$$\Delta H = -\vec{m}\cdot\vec{B} = \frac{e}{m}\vec{S}\cdot\vec{B}\,, \tag{9.81}$$

where

$$\vec{B} = \frac{2\mu_0}{3}\vec{m}_N\delta^3(0) + \frac{\mu_0}{4\pi r^3}\left(3\left(\vec{m}_N\cdot\hat{\vec{r}}\right)\hat{\vec{r}} - \vec{m}_N\right) \tag{9.82}$$

is the magnetic field generated by a magnetic moment

$$\vec{m}_N = g_N\frac{Ze}{2M}\vec{I}\,. \tag{9.83}$$

at the origin, M stand for the mass of the nucleus and g_N is the nucleus gyromagnetic factor. Taking into account that for an s-wave the contribution of the second term in (9.82) vanishes, we get

$$\Delta H = \frac{2\mu_0 g_N Ze^2}{6Mm}\left|\varphi_{n,l=0}(0)\right|^2\vec{I}\cdot\vec{S} = \frac{4}{3}\frac{m}{M}(Z\alpha)^4mc^2\frac{1}{n^3}\frac{1}{\hbar^2}\vec{I}\cdot\vec{S}\,, \tag{9.84}$$

where we have taken into account Eq. (9.75) together with the usual definitions a_0 and α. While displaying the same parametric than the spin-orbit term in Section 9.3.2, this interaction is further suppressed by at least three orders of magnitudes due the ratio of masses $m/M \sim m/(Zm_p) \sim 1/(1836Z)$. Taking into the results in that section with the formal replacement $\vec{J}\to\vec{F}$ and $\vec{L}\to\vec{I}$ and concentrating in the hydrogen case $Z=1$ where both the electron and proton have spin 1/2, we have

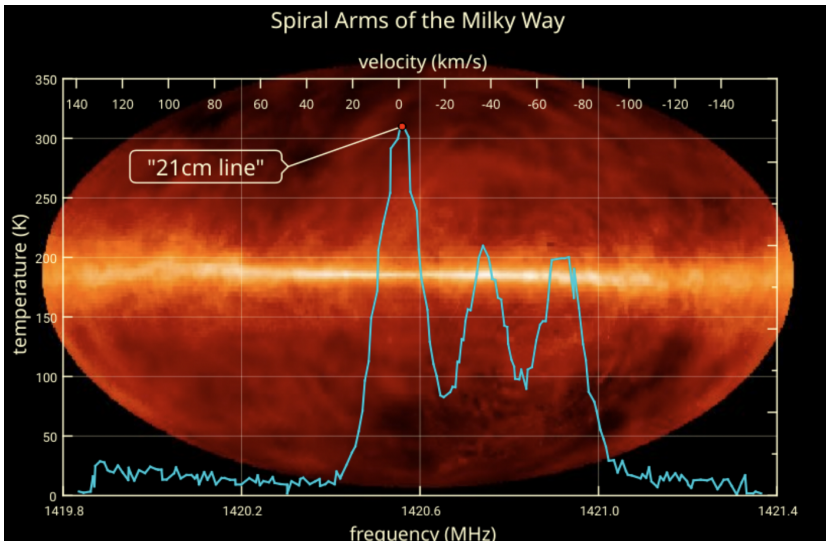
$$\frac{1}{\hbar^2}\vec{I}\cdot\vec{S} = \frac{1}{2\hbar^2}\left(\vec{F}^2 - \vec{S}^2 - \vec{I}^2\right) = \frac{1}{2}\left(F(F+1) - \frac{3}{2}\right) = \frac{1}{2}\begin{cases} -\frac{3}{2} & F=0\,, \\ \frac{1}{2} & F=1\,. \end{cases} \tag{9.85}$$

This gives rise to the splitting between the spin up and spin down elecotrnic states or, equivalently, between the total spin $F=0$ and $F=1$ of the atom.

The most important application of hyperfine structure is the splitting of the $1s_{1/2}$ ground state of hydrogen. When an electron drops from the $F=1$ state to the $F=0$ state, it emits a photon of energy

$$\Delta E_{1s_{1/2}} = \frac{4\alpha^4m^2c^2}{3M} \approx 9.39\times10^{-25}J\,, \tag{9.86}$$

corresponding to a frequency $\sim$ 1400 MHz and wavelength $\sim$ 21 cm. Contrary to visible light, the above wavelength exceeds significantly the size of that dust particles blocking our view in space. This makes the 21 cm emission line of hydrogen invaluable in astronomy and cosmology, as it could allow us to test the reionization epoch. This emission is extremely weak due to the long lifetime of the triplet state $F=1$, which is approximately $\tau \sim 3.5\times10^{14}$ s ($\sim 10^7$ years). However, the vast abundance of hydrogen in the interstellar medium ensures that the signals detected on Earth remain significant. The emission across the sky is displayed in the following figure, which appears naturally brighter along the Milky Way's plane due to the higher concentration and density of matter (Photo credit: J.Dickey/NASA SkyView).


Spiral Arms of the Milky Way
velocity (km/s)
140 120 100 80 60 40 20 0 -20 -40 -60 -80 -100 -120 -140
"21cm line"
temperature (K)
350 300 250 200 150 100 50 0
1419.8 1420.2 1420.6 1421.0 1421.4
frequency (MHz)

APPENDIX A

ALGEBRAIC SOLUTION OF THE HARMONIC OSCILLATOR

The career of a young theoretical physicist consists of treating the harmonic oscillator in ever-increasing levels of abstraction.

SIDNEY COLEMAN.

The Hamiltonian for a one-dimensional quantum harmonic oscillator is given by

$$H = \frac{P^2}{2m} + \frac{1}{2}m\omega^2 X^2 \,, \tag{A.1}$$

with the first term accounting for the kinetic energy associated with the momentum operator P and the second one describing the potential energy of the particle in a harmonic potential with angular frequency ω. As shown in Chapter 1, the corresponding time-independent Schrödinger equation in the position representation,

$$-\frac{\hbar^2}{2m}\frac{d^2\varphi(x)}{dx^2} + \frac{m\omega^2}{2}x^2\varphi(x) = E\varphi(x) \,, \tag{A.2}$$

can be solved using standard methods, such as the Frobenius series approach. Alternatively, one can consider a more elegant procedure based on operator methods. More precisely, we define the so-called annihilation operator a and its Hermitian conjugate, the creation operator $a^\dagger$, as

$$a = \sqrt{\frac{m\omega}{2\hbar}}X + \frac{i}{\sqrt{2m\omega\hbar}}P \,, \qquad a^\dagger = \sqrt{\frac{m\omega}{2\hbar}}X - \frac{i}{\sqrt{2m\omega\hbar}}P \,. \tag{A.3}$$

These new operators satisfy the commutation relation

$$[a, a^\dagger] = -\frac{i}{2\hbar}[X, P] + \frac{i}{2\hbar}[P, X] = 1 \quad \longrightarrow \quad [a, a^\dagger] = 1 \,, \tag{A.4}$$

following directly from the commutation relations for X and P. In terms of them, the Hamiltonian (A.1) becomes

$$H = \hbar\omega\left(N + \frac{1}{2}\right) \,, \tag{A.5}$$

with

$$N = a^\dagger a \tag{A.6}$$

a *number operator* counting the number of quanta or excitations in the system and satisfying the commutation relations

$$[a, N] = a\,, \qquad [a^\dagger, N] = -a^\dagger\,. \tag{A.7}$$

Note that, given an energy eigenstate $|E\rangle$ of the Hamiltonian, $H|E\rangle = E|E\rangle$, both $a^\dagger|E\rangle$ and $a|E\rangle$ are also eigenstates of H with eigenvalues $E \pm \hbar\omega$,

$$Ha^\dagger|E\rangle = \Big([H, a^\dagger] + a^\dagger H\Big)\,|E\rangle = (\hbar\omega a^\dagger + a^\dagger E)|E\rangle = (E + \hbar\omega)a^\dagger|E\rangle\,, \tag{A.8}$$

$$Ha|E\rangle = ([H, a] + aH)\,|E\rangle = (-\hbar\omega a + aE)|E\rangle = (E - \hbar\omega)a|E\rangle\,. \tag{A.9}$$

In other words, the creation and annihilation operators increase and decrease the energy of the system by a quantum $\hbar\omega$. Obviously, the energy of the system cannot decrease indefinitely since

$$0 \le \langle E|a^\dagger a|E\rangle = \frac{E}{\hbar\omega} - \frac{1}{2} \quad \longrightarrow \quad E \ge \frac{\hbar\omega}{2}\,. \tag{A.10}$$

The existence of a lower non-vanishing value is a direct consequence of the Heisenberg uncertainty principle. While in classical mechanics the ground state corresponds to a configuration with zero potential energy (i.e. localized exactly at $x = 0$) and zero kinetic energy (i.e. having no momentum), these two conditions cannot be simultaneously satisfied in quantum mechanics, preventing the particle from having zero total energy. For the ground state $|0\rangle$, we have $a|0\rangle = 0$ and

$$H|0\rangle = \frac{\hbar\omega}{2}|0\rangle\,. \tag{A.11}$$

Higher energy eigenstates can be constructed by repeatedly applying the creation operator $a^\dagger$. In particular, for a general n-th excited state we have

$$|n\rangle = \frac{1}{\sqrt{n!}}(a^\dagger)^n|0\rangle\,, \qquad E_n = \hbar\omega\left(n + \frac{1}{2}\right). \tag{A.12}$$

The action of the annihilation operator on these excited states,

$$a|n\rangle = \alpha_n|n-1\rangle\,, \tag{A.13}$$

follows directly from the commutation relation

$$[a, a^{\dagger\,n}] = aa^{\dagger\,n} - a^{\dagger\,n}a = [a, a^\dagger]a^{\dagger\,n-1} + a^\dagger[a, a^\dagger]a^{\dagger\,n-2} + a^{\dagger\,2}[a, a^\dagger]a^{\dagger\,n-3} + \ldots = na^{\dagger\,n-1}\,,$$

with α_n determined by the normalization condition

$$\langle n|a^\dagger a|n\rangle = |\alpha_n|^2 = \langle n|N|n\rangle = n\,. \tag{A.14}$$

Choosing this constant to be real, we have

$$a|n\rangle = \sqrt{n}|n-1\rangle\,. \tag{A.15}$$

The analogous expression for the creation operator,

$$a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle\,, \tag{A.16}$$

follows from noting that

$$a^\dagger|n\rangle = \beta_n|n+1\rangle\,, \qquad \langle n|aa^\dagger|n\rangle = |\beta_n|^2 = \langle n|N + [a,a^\dagger]|n\rangle = n+1\,, \tag{A.17}$$

and applying repeatedly $a^\dagger$

$$a^{\dagger\,n}|0\rangle = \sqrt{n!}|n\rangle\,. \tag{A.18}$$

To determine the wave functions associated with the previous states in Hilbert space, we project the relation

$$a|0\rangle = \left[\sqrt{\frac{m\omega}{2\hbar}}X + \frac{i}{\sqrt{2m\omega\hbar}}P\right]|0\rangle = 0 \tag{A.19}$$

on the position eigenstates $|x\rangle$,

$$0 = \langle x|a|0\rangle = \langle x|\left[\sqrt{\frac{m\omega}{2\hbar}}X + \frac{i}{\sqrt{2m\omega\hbar}}P\right]|0\rangle = \int dx'\langle x|\left[\sqrt{\frac{m\omega}{2\hbar}}X + \frac{i}{\sqrt{2m\omega\hbar}}X\right]|x'\rangle\langle x'|0\rangle\,, \tag{A.20}$$

and evaluate the expectation values of the position X and momentum P operators,

$$\langle x|X|x'\rangle = x\delta(x-x')\,, \qquad \langle x|P|x'\rangle = -i\hbar\frac{d}{dx}\delta(x-x')\,. \tag{A.21}$$

obtaining

$$0 = \left[\sqrt{\frac{m\omega}{2\hbar}}x + \frac{\hbar}{\sqrt{2m\omega\hbar}}\frac{d}{dx}\right]\langle x|0\rangle\,. \tag{A.22}$$

The normalized solution of this first-order differential equation for the ground state wave function $\varphi_0(x) \equiv \langle x|0\rangle$ is a Gaussian centered at $x = 0$,

$$\varphi_0(x) = \frac{1}{(2\pi\ell^2)^{1/4}}e^{-\frac{x^2}{4\ell^2}}\,, \qquad \ell = \sqrt{\frac{\hbar}{2m\omega}}\,, \tag{A.23}$$

The remaining wave functions for excited states can be obtained from this expression by an iterative procedure,

$$\varphi_n(x) = \langle x|n\rangle = \frac{1}{\sqrt{n}}\langle x|a^\dagger|n-1\rangle = \frac{1}{\sqrt{n}}\left[\frac{x}{2\ell} - \ell\frac{d}{dx}\right]\underbrace{\langle x|n-1\rangle}_{\varphi_{n-1}(x)}\,, \tag{A.24}$$

obtaining

$$\varphi_n(x) = \frac{1}{\sqrt{n!}}\left[\frac{x}{2\ell} - \ell\frac{d}{dx}\right]^n\varphi_0(x) \equiv \frac{1}{\sqrt{n!}}f_n(x)\varphi_0(x)\,, \tag{A.25}$$

with $f_n(x)$ a polynomial satisfying the recurrence relation

$$f_n(x) = \frac{x}{\ell}f_{n-1}(x) - \ell f'_{n-1}(x), \quad f_0(x) = 1\,. \tag{A.26}$$

Comparing this expression with the recurrence relation for the Hermite polynomials $H_n(z)$,

$$H_n(z) = 2zH_{n-1}(z) - H'_{n-1}(z)\,, \qquad H_0(z) = 1\,, \tag{A.27}$$

we conclude that

$$f_n(x) = \frac{1}{2^{n/2}} H_n\left(\frac{x}{\sqrt{2}\ell}\right) . \tag{A.28}$$

Thus, the wave functions of the excited states are finally given by

$$\varphi_n(x) = \frac{1}{\sqrt{2^n n!}} H_n\left(\frac{x}{\sqrt{2}\ell}\right) \varphi_0(x) . \tag{A.29}$$

Note that:

- $\varphi_{2n}(x)$ are even functions, i.e. $\varphi_{2n}(-x) = \varphi_{2n}(x)$,
- $\varphi_{2n+1}(x)$ are odd functions, i.e. $\varphi_{2n+1}(-x) = -\varphi_{2n+1}(x)$,
- $\varphi_n(x)$ has n nodes (zeros).

A list of the first few Hermite polynomials can be found in Appendix C.

APPENDIX B

ALGEBRAIC SOLUTION OF THE HYDROGEN ATOM

Beauty in bound up with symmetry

HERMANN WEYL, SYMMETRY, 1952

Let us consider the reduced Hamiltonian describing the motion of the electron in the static electric potential of the proton,

$$H = \frac{\vec{P}^2}{2m} - \frac{e^2}{r}\,, \tag{B.1}$$

with $m \simeq m_e$ the reduced mass of the system, e the magnitude of the electric charge and $r = \sqrt{x^2+y^2+z^2}$ the magnitude of the position vector $\vec{r}$ of the electron relative to the nucleus. Performing a rescaling

$$\vec{r} \to \frac{\vec{r}}{me^2}\,, \qquad \vec{P} \to me^2\vec{P}\,, \qquad H \to me^4 H\,, \tag{B.2}$$

consistent with the usual commutation relations for the position and momentum operators, we obtain a dimensionless Hamiltonian

$$H = \frac{\vec{P}^2}{2} - \frac{1}{r}\,. \tag{B.3}$$

Building on the classical description of the system, we can identify two conserved quantities: the usual angular momentum operator $\vec{L}$ and the so-called Runge-Lenz operator

$$\vec{A} = \frac{1}{2}(\vec{P}\times\vec{L} - \vec{L}\times\vec{P}) - \frac{\vec{r}}{r} = \frac{1}{2}\epsilon_{abc}\,(p_b L_c - L_b p_c) - \frac{r_a}{r}\,, \tag{B.4}$$

generalizing the classical Runge-Lenz vector $\vec{A} = \vec{P}\times\vec{L} - \vec{r}/r$ responsible for the conservation of the eccentricity and the position of the periapsis in the Kepler problem. These operators satisfy the commutation relations

$$[H, L_a] = 0\,, \qquad [H, A_a] = 0 \tag{B.5}$$

$$[L_a, L_b] = i\epsilon_{abc}L_c\,, \qquad [L_a, A_b] = i\epsilon_{abc}A_c\,, \qquad [A_a, A_b] = -2i\epsilon_{abc}L_c H\,, \tag{B.6}$$

and the operator equalities

$$\vec{A}\cdot\vec{L} = \vec{L}\cdot\vec{A} = 0\,, \qquad 2H(\vec{L}^2+1) - \vec{A}^2 + 1 = 0\,. \tag{B.7}$$

Note that the algebra in (B.5) and (B.6) is generically not closed since the commutator $[\vec{A}, \vec{A}]$ leads to new operators like $\vec{L}H$ which cannot be written as linear combinations of $\vec{L}$ and $\vec{A}$. This is not the case, however, if we restrict ourselves to subspaces spanned by eigenvectors of H with the same eigenvalue, since the Hamiltonian there is proportional to the identity. Working in these subspaces ($H \to E$) and assuming the existence of a bound states with $E < 0$, we have

$$[T_a, T_b] = i\epsilon_{abc} T_c \,, \qquad [K_a, K_b] = i\epsilon_{abc} K_c \,, \qquad [T_a, K_b] = 0 \,, \tag{B.8}$$

and

$$T^2 = K^2 \,, \qquad 2E\left(4T^2 + 1\right) + 1 = 0 \,, \tag{B.9}$$

with

$$\vec{T} = \frac{1}{2}(\vec{L} + \vec{N}) \,, \qquad \vec{K} = \frac{1}{2}(\vec{L} - \vec{N}) \,, \qquad \vec{N} \equiv \frac{\vec{A}}{\sqrt{-2E}} \,. \tag{B.10}$$

According to these expressions, $\vec{T}$ and $\vec{K}$ are essentially two angular-momentum-like operators with an additional constraint $T^2 = K^2$. The resulting algebra is indeed the algebra of the special orthogonal group $SO(4) = SO(3) \oplus SO(3)$, i.e. the group of all 4×4 orthogonal matrices with determinant one. [1] A complete basis of generators of this group is, for instance,

$$T_1 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & -\vec{i} \\ 0 & 0 & \vec{i} & 0 \\ 0 & -\vec{i} & 0 & 0 \\ \vec{i} & 0 & 0 & 0 \end{pmatrix} , \quad T_2 = \frac{1}{2}\begin{pmatrix} 0 & 0 & -\vec{i} & 0 \\ 0 & 0 & 0 & -i \\ \vec{i} & 0 & 0 & 0 \\ 0 & \vec{i} & 0 & 0 \end{pmatrix} , \quad T_3 = \frac{1}{2}\begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & -i & 0 \end{pmatrix} ,$$

$$K_1 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & -\vec{i} \\ 0 & 0 & -\vec{i} & 0 \\ 0 & \vec{i} & 0 & 0 \\ \vec{i} & 0 & 0 & 0 \end{pmatrix} , \quad K_2 = \frac{1}{2}\begin{pmatrix} 0 & -i & 0 & 0 \\ \vec{i} & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & \vec{i} & 0 \end{pmatrix} , \quad K_3 = \frac{1}{2}\begin{pmatrix} 0 & 0 & -\vec{i} & 0 \\ 0 & 0 & 0 & i \\ \vec{i} & 0 & 0 & 0 \\ 0 & -\vec{i} & 0 & 0 \end{pmatrix} . \tag{B.11}$$

Since the operators $\vec{T}$ and $\vec{K}$ commute, we can construct a common basis of eigenstates, labeled by the associated quantum numbers t, m_t and m_k, namely[2]

$$\begin{aligned} T^2 \left|t, m_t, m_k\right\rangle &= K^2 \left|t, m_t, m_k\right\rangle = t(t+1) \left|t, m_t, m_k\right\rangle \\ T_3 \left|t, m_t, m_k\right\rangle &= m_t \left|t, m_t, m_k\right\rangle \\ K_3 \left|t, m_t, m_k\right\rangle &= m_k \left|t, m_t, m_k\right\rangle \end{aligned} \tag{B.12}$$

with $t = 0, 1/2, 1, 3/2, 2, \ldots$, $-t \leq m_t, m_k \leq t$ and degeneation $(2t+1)^2$ for each t. Acting with the second equation in (B.9) on the subspace of eigenstates of energy E, we get

$$-\frac{1}{4}\left(\frac{1}{2E} + 1\right) = t(t+1) \,, \tag{B.13}$$

which has a solution

$$E = -\frac{1}{2(4t(t+1)+1)} = -\frac{1}{2(2t+1)^2} \equiv -\frac{1}{2n^2} \,, \tag{B.14}$$

[1] More precisely, $\vec{T}$ and $\vec{K}$ generate two independent $SU(2)$ groups that combine into the $SO(4)$ group.
[2] Note that due to the constraint $T^2 = K^2$, the label s is redundant and can be omitted.

with

$$n \equiv 2t + 1 = 1, 2, 3, \dots \,, \tag{B.15}$$

the so-called *principal quantum number*. Restoring the energy units, the energy levels and eigenstates of the hydrogen atom are given by

$$E_n = -\frac{me^2}{2n^2} \,, \qquad |t, m_t, m_k\rangle \,, \tag{B.16}$$

with level degeneracy $(2t + 1)^2 = n^2$.

The energy in (B.16) is minimized for $n = 1$ (or equivalently $t = m_t = m_s = 0$). This defines the ground state of the system, $|0, 0, 0\rangle$. The associated wave function can be obtained by solving the the differential equation $T^2\varphi_{000}(r) = 0$, or equivalently

$$H\varphi_{000}(r) = E_1\varphi_{000}(r) \qquad \longrightarrow \qquad \left(\frac{P^2}{2m} - \frac{e^2}{r}\right)\varphi_{000}(r) = -\frac{me^4}{2}\varphi_{000}(r) \,. \tag{B.17}$$

Considering a Gaussian ansatz

$$\varphi_{000}(\vec{r}) \propto e^{-r/a} \,, \tag{B.18}$$

we get $ame^2 - 1 = 0$ or $a_0 = 1/(me^2)$, i.e. the Bohr radius. The normalized wave function for the ground state reads therefore

$$\varphi_{000}(\vec{r}) = \frac{2}{a_0^{3/2}} e^{-r/a_0} \,. \tag{B.19}$$

APPENDIX C

PHYSICAL CONSTANTS AND USEFUL FORMULAS

Physical Constants

Electron rest mass	m_e	9.109×10^{-31} kg
Proton rest mass	M_p	1.6726×10^{-27} kg
Electronic charge	e	1.6022×10^{-19}C
Speed of light in free space	c	2.9979×10^{8} m s^{-1}
Permeability of free space	μ_0	$4\pi \times 10^{-7}$Hm^{-1}
Permittivity of free space	ϵ_0	8.854×10^{-12} F m^{-1}
Planck's constant	h	6.626×10^{-34} J s
Reduced Planck's constant	$\hbar = h/2\pi$	1.0546×10^{-34} J s
	$\hbar c$	197.33 MeV fm
Boltzmann's constant	k_B	1.3807×10^{-23}JK^{-1}
Gas constant	$\mathcal{R} = k_B/m_H$	8.250×10^{3} J kg^{-1} K^{-1}
Molar gas constant	R	8.315 J mol^{-1} K^{-1}
Avogadro's number	N_A	6.022×10^{23} mol^{-1}
Standard molar volume		22.414×10^{-3} m^3 mol^{-1}
Unified atomic mass unit (^{12}C scale)	u	931.5MeV/c^2 = 1.660538×10^{-27} kg
Mass of hydrogen atom	m_H	$1.0078u = 1.6735 \times 10^{-27}$ kg
Bohr magneton	μ_B	9.274×10^{-24} A m^2 or JT^{-1}
Nuclear magneton	μ_N	5.051×10^{-27} A m^2 or JT^{-1}
Proton magnetic moment	μ_p	$2.7928\mu_N$
Neutron magnetic moment	μ_n	$-1.9130\mu_N$
Bohr radius	a_0	5.292×10^{-11} m
Fine structure constant	$\alpha = e^2/(4\pi\epsilon_0\hbar c)$	$(137.04)^{-1}$
Compton wavelength of electron	$\lambda_C = h/(m_e c)$	2.4263×10^{-12} m
Rydberg's constant	R_∞	1.0974×10^{7} m^{-1}
	$R_\infty hc$	13.606 eV
Stefan-Boltzmann constant	σ	5.671×10^{-8} W m^{-2} K^{-4}
Radiation density constant	$a = 4\sigma/c$	7.561×10^{-16} J m^{-3} K^{-4}
Gravitational constant	G	6.673×10^{-11} N m^2 kg^{-2}

Useful Gaussian and exponential integrals

$$\int_0^\infty e^{-ax^2}dx = \frac{1}{2}\left(\frac{\pi}{a}\right)^{1/2} \qquad \int_0^\infty xe^{-ax^2}dx = \frac{1}{2a}$$
$$\int_0^\infty x^2e^{-ax^2}dx = \frac{1}{4a}\left(\frac{\pi}{a}\right)^{1/2} \qquad \int_0^\infty x^3e^{-ax^2}dx = \frac{1}{2a^2}$$
$$\int_0^\infty x^{2n}e^{-ax^2}dx = \frac{1\cdot3\cdot5\cdots(2n-1)}{2^{n+1}a^n}\left(\frac{\pi}{a}\right)^{1/2} \qquad \int_0^\infty x^{2n+1}e^{-ax^2}dx = \frac{n!}{2a^{n+1}}$$
$$\int_0^\infty x^ne^{-ax}dx = \frac{n!}{a^{n+1}}$$

Useful trigonometric integrals

$$\int \sin(ax)dx = -\frac{1}{a}\cos(ax)$$
$$\int \sin^2(ax)dx = \frac{x}{2} - \frac{\sin(2ax)}{4a}$$
$$\int \sin^3(ax)dx = -\frac{1}{a}\cos(ax) + \frac{1}{3a}\cos^3(ax)$$
$$\int \sin^4(ax)dx = \frac{3x}{8} - \frac{3\sin(2ax)}{16a} - \frac{\sin^3(ax)\cos(ax)}{4a}$$
$$\int \sin(ax)\sin(bx)dx = \frac{\sin[(a-b)x]}{2(a-b)} - \frac{\sin[(a+b)x]}{2(a+b)} \text{ where } a^2 \neq b^2$$
$$\int \cos(ax)\cos(bx)dx = \frac{\sin[(a-b)x]}{2(a-b)} + \frac{\sin[(a+b)x]}{2(a+b)}$$
$$\int \sin(ax)\cos(bx)dx = \frac{-\cos[(a-b)x]}{2(a-b)} - \frac{\cos[(a+b)x]}{2(a+b)}$$
$$\int x\sin^2(ax)dx = \frac{x^2}{4} - \frac{x\sin(2ax)}{4a} - \frac{\cos(2ax)}{8a^2}$$
$$\int x^2\sin^2(ax)dx = \frac{x^3}{6} - \left(\frac{x^2}{4a} - \frac{1}{8a^3}\right)\sin(2ax) - \frac{x\cos(2ax)}{4a^2}$$
$$\int x\sin(ax)\sin(bx)dx = \frac{\cos[(a-b)x]}{2(a-b)^2} - \frac{\cos[(a+b)x]}{2(a+b)^2} + \frac{x\sin[(a-b)x]}{2(a-b)^2} - \frac{x\sin[(a+b)x]}{2(a+b)^2}$$
$$\int x\sin(ax)dx = \frac{\sin(ax)}{a^2} - \frac{x\cos(ax)}{a}$$
$$\int x\cos(ax)dx = \frac{x\sin(ax)}{a} + \frac{\cos(ax)}{a^2}$$
$$\int \cos^2(ax)dx = \frac{\sin(ax)}{a}$$
$$\int \cos^2(ax)dx = \frac{x}{2} + \frac{\sin(2ax)}{4a}$$
$$\int x^2\cos^2(ax)dx = \frac{x^3}{6} + \left(\frac{x^2}{4a} - \frac{1}{8a^3}\right)\sin(2ax) + \frac{x\cos(2ax)}{4a^2}$$
$$\int \cos^2(bx)e^{-ax^2}dx = \frac{e^{ax}}{(a^2+b^2)}[a\cos(bx) + b\sin(bx)]$$

First few radial functions of the hydrogen atom

$$R_{10} = 2a^{-3/2}\exp(-r/a)$$
$$R_{20} = \frac{1}{\sqrt{2}}a^{-3/2}\left(1-\frac{1}{2}\frac{r}{a}\right)\exp(-r/2a)$$
$$R_{21} = \frac{1}{\sqrt{24}}a^{-3/2}\frac{r}{a}\exp(-r/2a)$$
$$R_{30} = \frac{2}{\sqrt{27}}a^{-3/2}\left(1-\frac{2}{3}\frac{r}{a}+\frac{2}{27}\left(\frac{r}{a}\right)^2\right)\exp(-r/3a)$$
$$R_{31} = \frac{8}{27\sqrt{6}}a^{-3/2}\left(1-\frac{1}{6}\frac{r}{a}\right)\left(\frac{r}{a}\right)\exp(-r/3a)$$
$$R_{32} = \frac{4}{81\sqrt{30}}a^{-3/2}\left(\frac{r}{a}\right)^2\exp(-r/3a)$$
$$R_{40} = \frac{1}{4}a^{-3/2}\left(1-\frac{3}{4}\frac{r}{a}+\frac{1}{8}\left(\frac{r}{a}\right)^2-\frac{1}{192}\left(\frac{r}{a}\right)^3\right)\exp(-r/4a)$$
$$R_{41} = \frac{\sqrt{5}}{16\sqrt{3}}a^{-3/2}\left(1-\frac{1}{4}\frac{r}{a}+\frac{1}{80}\left(\frac{r}{a}\right)^2\right)\frac{r}{a}\exp(-r/4a)$$
$$R_{42} = \frac{1}{64\sqrt{5}}a^{-3/2}\left(1-\frac{1}{12}\frac{r}{a}\right)\left(\frac{r}{a}\right)^2\exp(-r/4a)$$
$$R_{43} = \frac{1}{768\sqrt{35}}a^{-3/2}\left(\frac{r}{a}\right)^3\exp(-r/4a)$$

Laplacian and momentum operators in spherical coordinates

$$\nabla^2 = \frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r}+\frac{1}{r^2}\left(\frac{\partial^2}{\partial\theta^2}+\cot\theta\frac{\partial}{\partial\theta}+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right)$$
$$L^2 = -\hbar^2\left(\frac{\partial^2}{\partial\theta^2}+\cot\theta\frac{\partial}{\partial\theta}+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right)$$
$$L_z = \frac{\hbar}{i}\frac{\partial}{\partial\phi};\quad L_\pm = \hbar e^{\pm i\phi}\left(\pm\frac{\partial}{\partial\theta}+i\cot\theta\frac{\partial}{\partial\phi}\right)$$

First few spherical harmonics $Y_{l,m}(\theta,\phi) \equiv \langle \theta,\phi \mid \ell,m\rangle$

$$
\begin{aligned}
Y_{0,0} &= \left(\tfrac{1}{4\pi}\right)^{1/2} & Y_{2,\pm2} &= \left(\tfrac{15}{32\pi}\right)^{1/2} \sin^2\theta e^{\pm 2i\phi} \\
Y_{1,0} &= \left(\tfrac{3}{4\pi}\right)^{1/2} \cos\theta & Y_{3,0} &= \left(\tfrac{7}{16\pi}\right)^{1/2} \left(5\cos^3\theta - 3\cos\theta\right) \\
Y_{1,\pm1} &= \mp\left(\tfrac{3}{8\pi}\right)^{1/2} \sin\theta e^{\pm i\phi} & Y_{3,\pm1} &= \mp\left(\tfrac{21}{64\pi}\right)^{1/2} \sin\theta\left(5\cos^2\theta - 1\right) e^{\pm i\phi} \\
Y_{2,0} &= \left(\tfrac{5}{16\pi}\right)^{1/2} \left(3\cos^2\theta - 1\right) & Y_{3,\pm2} &= \left(\tfrac{105}{32\pi}\right)^{1/2} \sin^2\theta\cos\theta e^{\pm 2i\phi} \\
Y_{2,\pm1} &= \mp\left(\tfrac{15}{8\pi}\right)^{1/2} \sin\theta\cos\theta e^{\pm i\phi} & Y_{3,\pm3} &= \mp\left(\tfrac{35}{64\pi}\right)^{1/2} \sin^3\theta e^{\pm 3i\phi}
\end{aligned}
$$

$$
Y_{l,m}(\theta,\phi) = \epsilon\sqrt{\frac{(2l+1)}{4\pi}\frac{(l-|m|)!}{(l+|m|)!}} e^{im\phi} P_{l,m}(\cos\theta)
$$

First few Legendre functions $P_{l,m}(\cos\theta)$

$$
\begin{aligned}
P_{1,1} &= \sin\theta & P_{3,3} &= 15\sin\theta\left(1-\cos^2\theta\right) \\
P_{1,0} &= \cos\theta & P_{3,2} &= 15\sin^2\theta\cos\theta \\
P_{2,2} &= 3\sin^2\theta & P_{3,1} &= \tfrac{3}{2}\sin\theta\left(5\cos^2\theta - 1\right) \\
P_{2,1} &= 3\sin\theta\cos\theta & P_{3,0} &= \tfrac{1}{2}\left(5\cos^3\theta - 3\cos\theta\right) \\
P_{2,0} &= \tfrac{1}{2}\left(3\cos^2\theta - 1\right)
\end{aligned}
$$

First few spherical Bessel and Neumann functions $j_l(x)$ **and** $n_l(x)$

$$
\begin{aligned}
j_0 &= \tfrac{\sin x}{x} & n_0 &= -\tfrac{\cos x}{x} \\
j_1 &= \tfrac{\sin x}{x^2} - \tfrac{\cos x}{x} & n_1 &= -\tfrac{\cos x}{x^2} - \tfrac{\sin x}{x} \\
j_2 &= \left(\tfrac{3}{x^3} - \tfrac{1}{x}\right)\sin x - \tfrac{3}{x^2}\cos x & n_2 &= -\left(\tfrac{3}{x^3} - \tfrac{1}{x}\right)\cos x - \tfrac{3}{x^2}\sin x \\
j_l &\to \tfrac{x^l}{(2l+1)!!} & n_l &\to -\tfrac{(2l-1)!!}{x^{l+1}}, \quad \text{for} \quad x \ll 1.
\end{aligned}
$$

First few Laguerre polynomials $L_q(x)$

$$
\begin{aligned}
L_0 &= 1 \\
L_1 &= -x + 1 \\
L_2 &= x^2 - 4x + 2 \\
L_3 &= -x^3 + 9x^2 - 18x + 6 \\
L_4 &= x^4 - 16x^3 + 72x^2 - 96x + 24 \\
L_5 &= -x^5 + 25x^4 - 200x^3 + 600x^2 - 600x + 120 \\
L_6 &= x^6 - 36x^5 + 450x^4 - 2400x^3 + 5400x^2 - 4320x + 720
\end{aligned}
$$

First few Hermite polynomials, $H_n(x)$

$$
\begin{aligned}
H_0 &= 1 \\
H_1 &= 2x, \\
H_2 &= 4x^2 - 2 \\
H_3 &= 8x^3 - 12x \\
H_4 &= 16x^4 - 48x^2 + 12 \\
H_5 &= 32x^5 - 160x^3 + 120x.
\end{aligned}
$$

Commutator identities

$$
\begin{aligned}
&[A, BC] = [A, B]C + B[A, C] \\
&e^A B e^{-A} = B + [A, B] + \frac{1}{2}[A, [A, B]] + \frac{1}{3!}[A, [A, [A, B]]] + \ldots, \\
&e^A B e^{-A} = B + [A, B], \quad \text{if} \quad [[A, B], A] = 0 \\
&\left[B, e^A\right] = [B, A]e^A, \quad \text{if} \quad [[A, B], A] = 0 \\
&e^{A+B} = e^A e^B e^{-\frac{1}{2}[A,B]} = e^B e^A e^{\frac{1}{2}[A,B]}
\end{aligned}
$$