Undergraduate Lecture Notes in Physics

A. I. Lvovsky

Quantum Physics

An Introduction Based on Photons

Springer

Undergraduate Lecture Notes in Physics

Undergraduate Lecture Notes in Physics (ULNP) publishes authoritative texts covering topics throughout pure and applied physics. Each title in the series is suitable as a basis for undergraduate instruction, typically containing practice problems, worked examples, chapter summaries, and suggestions for further reading.

ULNP titles must provide at least one of the following:

- An exceptionally clear and concise treatment of a standard undergraduate subject.
- A solid undergraduate-level introduction to a graduate, advanced, or non-standard subject.
- A novel perspective or an unusual approach to teaching a subject.

ULNP especially encourages new, original, and idiosyncratic approaches to physics teaching at the undergraduate level.

The purpose of ULNP is to provide intriguing, absorbing books that will continue to be the reader's preferred reference throughout their academic career.

Series editors

More information about this series at http://www.springer.com/series/8917

A. I. Lvovsky

Quantum Physics

An Introduction Based on Photons

Springer

A. I. Lvovsky
University of Calgary and Russian
Quantum Center
Calgary, AB
Canada

A solutions manual for this book is available for download at https://www.springer.com/gp/book/9783662565827

ISSN 2192-4791 ISSN 2192-4805 (electronic)
Undergraduate Lecture Notes in Physics
ISBN 978-3-662-56582-7 ISBN 978-3-662-56584-1 (eBook)
https://doi.org/10.1007/978-3-662-56584-1

Library of Congress Control Number: 2018933016

Printed on acid-free paper

This Springer imprint is published by the registered company Springer-Verlag GmbH, DE part of Springer Nature
The registered company address is: Heidelberger Platz 3, 14197 Berlin, Germany

Preface

Why I wrote this book

The first rigorous formulation of quantum mechanics (QM) was proposed by Heisenberg and Schrödinger about 80 years ago. Since then, the field has experienced enormous evolution. Initially aimed at explaining atomic spectra, quantum mechanics has now entered the foundation of virtually all branches of physics. Accordingly, QM is an inseparable part of every physics student's training: whatever field future physicists choose after graduation, they will almost certainly need quantum mechanics in their work.

Yet our way of teaching QM to students has not changed much over the years. We begin with the notion of the wavefunction, and write the time-independent, and then the time-dependent Schrödinger equation in the position representation. We determine the energy spectra and the corresponding wavefunctions of simple potential wells, and evolution of wavepackets incident on potential barriers. Finally, we introduce the angular momentum operator and calculate the spectrum of the hydrogen atom. For the last 60 years, this has been, with minor variations, the first semester undergraduate quantum mechanics course program.

This tradition has many advantages. It works with a physical system that a student has already dealt with in classical physics classes, and it is one that they can easily imagine. It allows one to see differences between the behaviors of a classical and a quantum particle, and brings to light several fundamental phenomena that are characteristic of quantumness: tunneling, quantization and the uncertainty principle. It provides a student with the toolbox to solve experimentally relevant problems that cannot be solved classically: after calculating the hydrogen spectrum in the classroom, a student goes to a lab and measures it!

Yet this approach is imperfect. It gives a student an algorithm to analyze a specific physical system, but it does not reveal the inner workings of quantum physics and the logic behind it. We teach our students multiple facts and computational approaches related to wavefunctions, operators, and measurements, but we do not construct a rigorous logical connection among them and do not explain which of

these facts are the postulates and which their consequences, and in which logical sequence these consequences are derived.

As a result, the student — at least a *thinking* student — ends up being immensely confused. Why does simply placing hats on top of letters turn a classical expression into a valid quantum one? Why is the momentum operator's action on the wave function equivalent to taking the derivative? Why do we never see momentum eigenstates (and Schrödinger cats) in practice? Why do de Broglie waves have a normalization factor of $1/\sqrt{2\pi\hbar}$? Why do we observe atoms transitioning between energy eigenstates, but not other states? How is a projective measurement related to measuring an observable? Why are some states described by wavefunctions and some by columns of numbers? If all states have norm 1, why don't we normalize the de Broglie wave? If observables are matrices, what is the matrix of the momentum observable?

On top of that there is the most dreaded question. If quantum physics is supposed to be more general than classical, why must one resort to classical notions to understand the concept of measurement? Why is the measurement, in contrast to all other physical processes, not described by unitary evolution? If quantum systems do become classical at some point during the measurement, where exactly is that point?

The fundamental way of thinking we are trying so hard to instill in our students through the years of training in science is "Question everything!". In quantum classes our message seems to be just the opposite: "Shut up and calculate!"[1]

Having been a quantum mechanics student myself, I have eventually found answers to these questions, but in many cases not until long after my PhD. When I asked them as a student, there was no one around, not only to give me the answers, but even to help me state these questions properly.

My quest while writing this book is to address this issue. I attempted to build a clear logical structure, containing as few loopholes as possible. One that would allow the reader to trace each statement down the logical chain back to first principles. One that would *leave no question unanswered.*

So, in a sense, I wrote this book for *myself.* Not for today's myself, but for myself twenty-five years ago. A kind of book that I would have been grateful to have while a third-year student, and one that would have saved me years of agonizing search for the truth.

It is natural to ask: How realistic is my aspiration? Some of the questions I asked earlier sound quite advanced. Perhaps one does need a doctoral degree to answer them?

My answer is twofold. First, there is a pedagogical issue: mechanics, with its Hilbert space of infinite dimension, does not seem to be the best venue for conveying quantum principles. Many of the above questions can be addressed by exemplifying QM with a simpler physical system; I will further elaborate on this below. Second, most of the inconsistencies and paradoxes can be eliminated by properly introducing the notion of entanglement. This notion underlies two essential, mutually rela-

[1] More on this slogan, incorrectly ascribed to Feynman, in Sec. 2.4.

ted concepts: the *von Neumann measurement* and *decoherence*. The first provides a way to avoid making measurement an exception in the domain of quantum physics, thereby eliminating the Klein bottle logic characteristic of the Copenhagen interpretation. The second describes "inadvertent" measurements that occur naturally, making the quantum world appear to macroscopic observers such as ourselves in the way that is familiar to us under the name of classical physics.

These concepts are not excessively complicated. Mathematically, they are much simpler than many elements of the traditional quantum course, such as the aforementioned hydrogen atom or scattering theory. The main challenge in understanding entanglement is not the challenge to a student's mathematical skills; it rather concerns their imagination. And developing a strong imagination is inherent to becoming a good physicist; as Einstein said, imagination is in fact more important than knowledge.

Quantum mechanics or quantum optics?

The name of our discipline — quantum mechanics – implies that we are studying the applications of quantum principles to the laws of motion. But in fact the framework of quantum theory is not limited to mechanics; it actually applies to all fields of physics. If our aim is to study the general principles of quantum physics, is it wise to choose mechanics as the physical system for illustrating these principles?

Faced with this question, we are compelled to admit that the answer is negative. Mechanics is there mainly due to tradition, because historically, the first successful application of these principles in their modern form was in mechanics. But educationally speaking, using the example of mechanics to explain basic quantum principles is a recipe for disaster. The Hilbert space associated with this system is of infinite dimension; moreover, its basis has the cardinality of the continuum. The student must deal with the unfamiliar, enormously complicated, and not always rigorous mathematical background which includes generalized functions, Fourier transformations, and functional analysis. As a result, instead of concentrating our students' efforts on understanding the physical concepts, we force them to struggle with mathematics, and this often leads to confusion between the end and the means. It is unrealistic to expect any kind of deep understanding to result from this experience. The student simply won't see the forest behind the trees.

If we are to choose the physical system to illustrate quantum physics, we should pick one whose Hilbert space has the lowest nontrivial dimension: two. There is a variety of such systems that are currently studied in the context of quantum information technology as quantum bits. Among them, one stands out as the most thoroughly studied and intuitive: the polarization of the photon. Optical wave polarization would normally already have been studied by the student entering a quantum class. The Jones polarization vectors directly translate into photon polarization state vectors, and the matrices describing the transformation of these vectors by waveplates translate into operators. It is straightforward to argument the measurement

postulate from the classical polarization measurement picture, taking into account the discrete nature of the photon. In this way, the quantum fundamentals arise from classical polarization optics (and the students' laboratory experience with the same) in the most straightforward and natural fashion.

Photon polarization is of further benefit when we go on to study entanglement. A vast body of proof-of-principle experiments in quantum information have been performed using this system as the carrier of the quantum bit. Some of these experiments, such as those on quantum cryptography, teleportation and noonlocality, relate directly to the concepts covered in this book. By illustrating the theoretical material with these experiments, right from the start, this book will take students straight into the very heart of quantum physics using examples from today's hottest research topics. And what could make learning an academic discipline more exciting than fresh results from a research lab?

Talking about labs, the student's experience does not have to be limited by reading about experiments done by somebody else. A great advantage of the polarization qubit as the example system is that it is perfectly realistic to augment the course with a laboratory component. Almost all the material of the first chapter is illustrated by a classical polarization experiment containing a laser, a few waveplates, a polarizing beam splitter and two detectors. The material on entanglement can be visualized by a series of labs on single-photon interference remote state preparation and Bell nonlocality. Such experiments are more difficult to set up, but fully within the capabilities of an average physics department, as evidenced by the experience of many colleges around the world, including my home, the University of Calgary. More details on possible educational labs can be found at the book's web site.

The connection between quantum physics and optics in this book is not limited to using the photon to illustrate the discipline's primary concepts. It also manifests itself in the many optical examples scattered throughout the book, as well as the set of subjects chosen for more advanced sections (deep study of the harmonic oscillator, Heisenberg picture, squeezing, density matrices, two-level systems, quantum tomography). These subjects are particularly relevant for those who are interested in quantum optics in particular and quantum information in general.

Structure of the course

The book contains material that can be taught during a two-semester undergraduate quantum mechanics course. In the *first* chapter, the main principles and postulates of QM are introduced and illustrated by the photon polarization qubit. The reader may wish to study this chapter in parallel with Appendix A, which covers the basic linear algebra that is relevant to QM, as summarized in the following table.

Linear algebra concept (Appendix A)	**Quantum concept**	**Physical illustration**
Linear space, basis, dimension, inner product	Quantum state, Hilbert space	Polarization of the photon (Appendix C)
Orthonormal basis	Projective measurement, quantum tomography	Polarization measurements, polarization state tomography, quantum cryptography
Linear operator, Hermitian operator	Observable, uncertainty principle	Pauli matrices as observables in the polarization space
Unitary operator, functions of operators	Schrödinger evolution	Evolution of the photon in a birefringent medium

The *second* chapter is entirely dedicated to entanglement, its consequences and applications. I first introduce the tensor product space mathematically, then explain partial quantum measurements, remote state preparation, and the nonlocality paradox (both the Bell and Greenberger-Horne-Zeilinger forms of it), illustrating the theory with experiments on entangled photons. Nonlocality is arguably the primary paradox of quantum mechanics, and it is natural to follow up with a discussion of the mechanism of quantum measurements, their natural counterpart (decoherence), and the interpretations of quantum mechanics. This section (Sec. 2.4) is where we find out when and why a quantum system becomes classical during a measurement, and why we don't see Schrödinger cats walking around town. Subsequently, I talk in a fairly rigorous fashion about applications of entanglement, such as quantum computation, teleportation and repeaters. When this material is presented in a classroom setting, it is useful to ask two or three students to give presentations on recent experimental research on the subject.

The *third* and fourth chapters are, to some extent, a tribute to the "mainstream" undergraduate quantum mechanics of a particle in a potential field. Here we have to deal with the Hilbert space whose basis is a continuum, so the third chapter is accompanied by a tutorial on Dirac delta functions and the Fourier transform (Appendix D). It is my hope that at this point, when students have already internalized the primary tenets of QM, they will be able to face the technical issues associated with continuous-variable Hilbert spaces without losing sight of the physical principles. As an introduction to continuous-variable systems, I explain how and why some of the normalization rules are affected. Then I present the usual scenarios of potential wells, potential barriers, tunneling, and the harmonic oscillator. This is the point where I envision the first semester to be concluded.

The third chapter goes on to explain the Heisenberg picture and how it is consistent with the Schrödinger picture, illustrating with many examples relating to the physics of the harmonic oscillator (and demonstrated in quantum optics experiments): displacement, phase shift, as well as single- and dual-mode squeezing. With the help of the latter, I present the original version of the Einstein-Podolsky-Rosen paradox.

In the *fourth* chapter, I introduce the three-dimensional geometric space (as a tensor product of three one-dimensional spaces) and explain the angular momentum,

spin, and, finally, the hydrogen atom. Then I discuss the behavior of a spin in a magnetic field and magnetic resonance, covering the phenomena of spin echo and Ramsey spectroscopy.

The *fifth* chapter revisits the fundamental principles of QM, now presenting them in the language of density operators, which is of primary significance for all applications of quantum physics. To demonstrate the utility of this language, I apply it to give a formal description of decoherence and relaxation in nuclear magnetic resonance. I then cover the topics that are relevant to modern quantum information science: generalized measurements (POVMs) as well as quantum state, process and detector tomography.

How to use this book (a message to the student)

I have been involved in education, on both sides of the podium, for most of my conscious life. This experience taught me one truth: it is almost impossible to learn anything by passively listening to a lecturer or reading a book. Learning requires active engagement. In the case of theoretical physics, this means that you should perform the derivations yourself rather than observing them performed by somebody else on the blackboard or in the textbook.

With this in mind, I tried to write this text using the Socratic approach: the student comes to the truth by answering the teacher's questions. My own experience with this method comes from high school. I was fortunate enough to attend one of the best Russian science high schools, which has a unique approach to teaching mathematics. Instead of explanations, we were provided with assignments consisting of only of definitions, axioms, and problems. After solving them, we discussed our solution with a tutor, whose task was to ensure that we understood the material correctly.

This book works in the same fashion. You will notice that it contains an unusually large number of exercises. Some of them are conceptual theorems; others are there just for practice; quite a few are both. The idea is that, by solving them one-by-one, you yourself will construct QM with as little help from me as possible. Accordingly, it is not advisable to skip the exercises. This is equivalent to skipping a page or two from a regular textbook: you will not be able to follow what comes next.

Almost all of the exercises have solutions, which are available for download from my website, accessible via `https://www.springer.com/gp/book/9783662565827`. However, please do not look at the solution until you have at least tried to solve the exercise independently. Even if you fail to arrive at the answer yourself, you will be conscious of the point where you are stuck, so the solution will be there as an answer to your existing question. In this way, the seed will fall onto soil that is already fertilized.

On the other hand, I advise that you do have a quick look at the solution even if you have found your own. In this way you will become aware of the errors you (or I) may have made, or of a possible alternative approach to solving the same problem.

Those exercises I consider more difficult are marked with asterisks (*). Unfortunately, many of them contain statements that are essential for subsequent material. Therefore, while it may be acceptable to defer solving (or studying the solutions of) these exercises to the future, you should at least make sure you understand the statements contained therein.

Similarly, sections labeled with asterisks contain more advanced material. You can safely skip these sections without fear to compromise your ability to understand the subsequent content.

Some of the exercises (marked with the symbol §) are not provided with solutions. Usually this is the done when I believe that the problem is relatively simple; in these cases I generally provide an answer immediately after the exercise. Very rarely, there will be an exercise that is marked by both an asterisk and a paragraph sign. These constitute independent research projects that may be worth looking at in your free time.

What knowledge do I expect you to have before you open this textbook?

- You should be familiar with trigonometry (e.g. how to expand $\cos(\alpha+\beta)$ or $\cos\alpha\cos\beta$ into a sum of terms).
- You should be able to work with complex numbers, being familiar with the notions of conjugation, complex phase, and complex exponent (e.g. simplify $|1+\mathrm{e}^{\mathrm{i}\phi}|^2$).
- You should be reasonably comfortable with probability theory. To help you, Appendix B contains some rudiments of this field.
- The same applies to the physics of optical wave polarization: Appendix C briefly covers the necessary knowledge, but would not be a good replacement for a proper textbook.
- Calculus and ordinary differential equations are essential throughout book, especially Chap. 3 (quantum physics of continuous-variable systems); this requirement extends to multivariate calculus (Jacobian determinant, etc.) for Chap. 4. There is no appendix on calculus, but Appendix D covers the Dirac delta function as well as the direct and inverse Fourier transforms, so pre-existing knowledge of mathematical physics is not required.
- Of primary importance to quantum physics is linear algebra, including the notions of linear spaces, basis, dimension, inner product, orthonormal basis, linear operators and matrices, spectral theorem, functions of operators, and so on. These are covered in Appendix A. However, basic matrix manipulation techniques, such as matrix multiplication, finding eigenvectors and eigenvalues, do not feature in this appendix and should be familiar to you before you start the course.

Acknowledgements

It took me thirteen years to write this book, from January 2005 to December 2017. It is common to remember the end date precisely, because this is the publisher's deadline (in the present case, moved many times, and for a few years). The reason I also remember the start date is that it coincides with the semester I started to teach PHYS 443 Quantum Mechanics I at the University of Calgary. I had just joined the University of Calgary faculty at that time and was not even supposed to teach in that semester. However, when the department head Bart Hicks came to me one day and asked in a nice voice: "Alex, would you like to start teaching a bit earlier? I heard you were interested in Quantum, and we just have a slot", I (a naïve romantic rookie professor) said yes. This is when the first handwritten set of notes was generated.

But the true history of a fruit starts from the roots. And since this book is all about roots, I should also follow this principle in this section. I can trace the roots back to 1962, when my parents, Isay and Tatyana, just a few months before meeting each other in Moscow, saw *Nine Days in One Year* — a Soviet movie about physicists that was cult at that time (by the way, you should see it when you get a chance. It is easy to find online with English subtitles and still a pleasure to watch. And it does profess good values). The cult quickly subsided, but not with my parents. So my profession was decided upon eleven years before I was born. The only disagreement between my parents was whether I should become a member of the Academy of Sciences (the Soviet equivalent of a fellow of the Royal Society) or a Nobel Prize winner. My grandfather reconciled them by pointing out that one does not have to preclude the other.

Fortunately my natural inclinations turned out consistent with my parents' aspirations — in direction if not in magnitude (I sometimes ask myself who I might have become if I had been raised by a different family. I think, either a car mechanic or a software engineer. So experimental physicist seems like a nice compromise). So, skipping a few years, I found myself to be a student of the famous Moscow High School Number 57 (schools in the Soviet Union had numbers, not names) with enhanced training in math and physics. This is where I experienced the Socratic principle of teaching science that I described in the Preface and upon which this book is based. The method was invented by a Moscow teacher Nikolay Nikolaevich

Konstantinov, but the person who taught our class and introduced me to this method was Boris Mikhailovich Davidovich. The storyline of the first two sections from Appendix A and a few exercises from these sections come straight from my high school archives.

Then, college. The professor who introduced me to quantum physics, and fascinated me by it, was Yury Mikhailovich Belousov. He skillfully combined the rigor of the "old school" of Landau and Lifshitz with a colorful, insightful, and passionate teaching manner: "What is the state? An undefinable notion! Like in geometry: you don't define what a point or a straight line is, do you? State, same thing. What is your state? Drunk? Sober? Tired? That's the state. The set of states is called the state space. Again, why not? But then we say that this space is linear. Now that's quite a claim..."

Yet, as also mentioned previously, not all my questions were answered (nor even properly asked) at college, and I had to seek the answers myself for a long time after graduation. On this quest, I have been supported by many brilliant scientists. Just to name a few: Alain Aspect, Konrad Banaszek, Mauro D'Ariano, Hauke Hansen, Peter Marzlin, Philippe Grangier, Miklos Gyulassy, Paul Kwiat, Misha Lukin, Eugene Polzik, Mike Raymer, Barry Sanders, Christoph Simon, Aephraim Steinberg, Ian Walmsley, Xing Wei, and Anton Zeilinger. Two names I should mention separately: my undergraduate advisor, Anatoly Viktorovich Masalov, who introduced me to research, and my PhD advisor, Sven Hartmann, Mr. Photon Echo. In addition to a great deal of science, Sven taught me how to write scientific texts. If this book has any style, it is thanks to him.

While it is difficult to identify a single person who was most instrumental in forming my understanding of quantum physics, I can name a period in my life during which I made most progress. It is the time when I worked as a postdoc at the University of Konstanz, at the institute headed by Dr. Jürgen Mlynek. At that time, the institute was a "Mecca" for quantum physicists, visited by the top brains of the field. Sometimes I was able to steal a few minutes from their busy schedules to discuss those questions I was concerned about, including those of quantum fundamentals (whenever I could collect enough courage to overcome the fear of appearing foolish or ignorant).

Let me now turn back to the moment I started teaching Quantum I in Calgary and wrote my first set of notes. The notes have been rewritten and augmented dozens of times. Perhaps a turning point in converting them into a book was the addition of solutions to the exercises. Initially, solutions were not there; I simply presented them orally during the lectures (I still wonder how these students managed to pass the tests). But then, two important conversations happened. First, I spoke to Jeff Shapiro, an MIT professor who taught me quite a bit of quantum optics during our (alas, brief) meetings. I told Jeff about my idea to convert my lecture notes into a book and about the Socratic method. Jeff looked at me seriously and asked: "But the problems will have solutions... Right?" And again, almost by a miracle, around the same time, two of my undergraduates, Geoff Campbell and Dallas Hoffman, approached me. "Your notes could really benefit from solutions. We thought, perhaps we can write some of them." And they did — quite a few of the solutions for the

exercises from Chapters 1, 2 and Appendix A come from them, and I am immensely grateful to these guys.

The students' support was in fact paramount at all stages of this book's construction. Since 2005, I taught Quantum I six times to about 200 students, and many of them made important contributions: Russel Bate, Dante Bencivenga, Travis Brannan, Arthur Bury-Jones, Aveek Chandra, Jose da Costa, Ish Dhand, Stefan Donsa, Mark Girard, Chris Healey, Katanya Kuntz, Kimberley Owen, Adarsh Prasad, Mathew Richards, Steven Rogowski, Matthew Townley-Smith, Raju Valivarthi. Their help consisted not only in contributing solutions, but also in finding errors and in asking multiple questions that allowed me to understand which passages were unclear and needed to be improved. Again, I will not be able to name every student who helped, so I have to ask for the kind forgiveness of those who have not been mentioned.

As these notes have been inspired by my own high school experience, I wanted for a long time to try them in the same setting. I was able to realize this in 2013, when I took a sabbatical leave from the University of Calgary to help setting up the Russian Quantum Center in Moscow. At that time, I organized a voluntary quantum physics class — *Kruzhok* or "little circle" in Russian — for Moscow high school students. Together with a team of enthusiast tutors, led by Aleksey Fedorov, we met with the students weekly to hear how they solved the problems in these notes (no solution manual was provided), correct them, explain the subtleties, and — last but not least — discuss the notes themselves. The feedback received in these discussions has been instrumental in shaping up this text, and several members of the Kruzhok — including Aleksey — have now become professional scientists pursuing research on quantum technology on a full-time basis.

I would like to thank Stephen Lyle for thorough proof-reading of the book and providing many insightful comments.

But the greatest thanks of all goes to my wife, Bhavya Rawal. As I am writing these lines, she is driving to pick up our daughter, Sophie, from her grandfather. This is just one of many hundreds of occasions when I really should have been together with my family instead of hiding behind the monitor typing away strange wiggles. But now it seems that even Bhavya's infinite patience wears thin. Yesterday she showed me the movie *Paris Can Wait*, in which the wife of a guy who works too much lets herself get seduced by his French co-worker. Darling, I get the hint. Paris can wait no longer. This is the last sentence I will add to the book!

Calgary, December 10, 2017

Contents

Chapter 1
The quantum postulates

1.1 The scope of quantum mechanics

Perhaps the first thing to understand about quantum mechanics is that it has as much to do with mechanics as with, say, electrodynamics, optics, condensed-matter, or high-energy physics. Rather than describing a particular class of physical phenomena, quantum mechanics provides a *universal theoretical framework* that can be used in *all* fields of physics — akin to a computer's operating system that provides a foundation upon which other applications can run. The term "quantum mechanics" emerged historically, because the first successful applications of the quantum framework were in studies of the mechanical motion of electrons in an atom. A better term would be "quantum physics" or "quantum theory".

So the scope of quantum physics is global: it covers all physical phenomena in the universe. However, a quantum treatment is *practical* only in the case of very small (microscopic) physical systems. The behavior of larger systems is very well approximated by the laws of classical physics, which are much simpler and more intuitive, at least for beings that have evolved on that length scale.

Let me illustrate this by an example. You have probably heard of Heisenberg's uncertainty principle: $\Delta p \Delta x \gtrsim \hbar/2$. That is, a particle's position and momentum cannot be measured precisely and simultaneously: the product of the uncertainties is at least $\hbar/2 \approx 5 \times 10^{-35}$ kg·m^2/s. For a macroscopic object with a mass on a scale of a kilogram, reaching the quantum uncertainty limit would require measuring either the position with a precision on a scale of at least $\sim 10^{-17}$ m or the velocity with precision $\sim 10^{-17}$ m/s. This is, of course, unrealistic, so for all practical purposes we may as well forget about the uncertainty principle and treat the position and momentum as precise quantities. But for an electron of mass $\sim 10^{-30}$ kg, the product of the position and velocity uncertainties will be about 5×10^{-5} m^2/s, which is well within experimentally attainable measurement precision and must be taken into account.

So the predictions of quantum theory are different from classical ones only for relatively simple, microscopic objects. This explains why quantum mechanics was

A solutions manual for this chapter is available for download at https://www.springer.com/gp/book/9783662565827

A. I. Lvovsky, *Quantum Physics*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-662-56584-1_1

not discovered until the early 20th century. Before then, we (who ourselves are macroscopic entities) only dealt with macroscopic objects. But as soon as we developed tools to probe the microscopic world deeply enough, quantum phenomena became manifest.

This is an example of the *correspondence principle*: a philosophical maxim that states that any new, more modern, theory should reproduce the results of older well-established theories in those domains where the old theories have been tested. Here is another example of this principle. As long as we had to do with objects that move much more slowly than light, Newtonian mechanics was sufficient to describe the world around us. But as soon as we became able to observe bodies that move quickly (e.g., the Earth around the sun in the Michelson-Morley experiment), we began to see discrepancies and were compelled to develop the theory of relativity. This theory is distinctly different from Newtonian mechanics — yet it is consistent with the latter in the limiting case of low velocities. It would be unwise to use special relativity to describe, for example, a tractor transmission, because the classical approximation is in this case both sufficient and tremendously simpler. Similarly, using quantum physics to describe macroscopic phenomena would in most cases be overcomplicated and unnecessary.

In classical physics, we deal with *quantities*: a rock flying at a speed of 10 meters per second, a circuit carrying a current of 0.2 amperes, and so on. Even if we do not know a physical quantity exactly, we can work on improving our theory and experiment to predict and measure it with ever increasing precision. In other words, *the classical world is infinitely knowable*. In quantum physics, the situation is different: some knowledge (such as the simultaneous values of the position and momentum) is "sacred": it cannot be attained even in principle. And this situation can no longer be described in terms of quantities alone. Instead, we must use the concept of the *quantum state* of a physical system. As we shall see, this concept incorporates the boundary between the knowledge that is possible and the knowledge that is impossible to obtain. We can learn precisely what state the system is in, but each state is associated with fundamental limits on the precision with which physical quantities can be known.

Because quantum mechanics has this role as a general framework, we will study it in a fairly rigorous, mathematical fashion. I will introduce definitions and axioms, then predict phenomena that arise from them, and then illustrate these phenomena with examples from different fields of physics, primarily from optics.

The main mathematical tool of quantum mechanics is linear algebra. Appendix A of this book teaches the concepts of this discipline that are relevant to quantum physics. So if you feel comfortable with your linear algebra, please proceed to the next section. Otherwise I would recommend that you study the first four sections of Appendix A before moving on.

1.2 The Hilbert Space Postulate

Let me first give a succinct formulation of the Postulate[1], and then explain its meaning in more detail.

a) Possible states of a physical system form a Hilbert space over the field of complex numbers.
b) Incompatible quantum states correspond to orthogonal vectors.
c) All vectors that represent physical quantum states are normalized.

This Postulate contains two notions that have not been defined: quantum state and physical system. They are so basic that their rigorous definition is difficult[2]. So let me try to explain these notions intuitively, using examples.

A *physical system* is an object, or even one or several degrees of freedom of an object, that can be studied independently of other degrees of freedom and other objects. For example, if our object is an atom, quantum mechanics can study its motion as a whole (one physical system) or the motion of its electrons around the nucleus (another physical system). On the other hand, if we wish to study the formation of a molecule out of two atoms, motional states of both the atoms and the electrons therein affect each other, so we must consider all these degrees of freedom as one physical system. For a molecule itself, quantum mechanics can study its center of mass motion (one physical system), rotational motion (another physical system), vibration of its atoms (a third system), quantum states of its electrons (a fourth system), and so on.

To grasp the notion of a state, consider the following physical system: a massive particle that can move along the x coordinate axis. One can define its quantum state by saying "the particle's coordinate is exactly $x = 5$ meters". This is a valid definition; we would denote this state as $|x = 5\,\mathrm{m}\rangle$. Another valid state would be $|x = 3\,\mathrm{m}\rangle$. These states are orthogonal ($\langle x = 5\,\mathrm{m}|\ x = 3\,\mathrm{m}\rangle = 0$) because they are "incompatible": if a particle's coordinate is definitely known to be 5 meters, it cannot be detected at $x = 3$ meters. On the other hand, the particle can be in the state "moving at a speed $v = 4$ meters per second". This is also a valid quantum state. Because the momentum of the particle is certain in this state, the position is completely uncertain, which means that the particle in this state can, with some probability, be detected at $x = 5$ m. Hence the inner product $\langle x = 5\,\mathrm{m}|\ v = 4\,\mathrm{m}/s\rangle$ does not vanish; these states are not incompatible.

The Postulate also says that if $|x = 5\,\mathrm{m}\rangle$ and $|x = 3\,\mathrm{m}\rangle$ are valid quantum states, then $(|x = 5\,\mathrm{m}\rangle + |x = 3\,\mathrm{m}\rangle)/\sqrt{2}$ (where $1/\sqrt{2}$ is the normalization factor — see Ex. 1.1 for the explanation) is also a valid state. It is called a *superposition* state.

---

[1] There are no universally accepted postulates of quantum mechanics. If you say "This follows from Newton's Second Law", people will understand you, but if you say "This follows from the First Postulate of quantum mechanics", they won't. You should instead say, for example, "It follows from the linearity of the quantum Hilbert space".

[2] As in geometry, which is an extremely rigorous science, despite the fact that its primary notions such as the point, straight line, and plane are not defined.

More spectacularly, if $|\text{living cat}\rangle$ and $|\text{dead cat}\rangle$ are valid states of the physical system "cat", so is the superposition of these states[3].

Are superposition states a mathematical abstraction or do they manifest themselves in their physical behavior? The answer is, certainly, the latter. As we shall see shortly, if we subject, e.g., a cat in states $(|\text{living cat}\rangle + |\text{dead cat}\rangle)/\sqrt{2}$, $(|\text{living cat}\rangle - |\text{dead cat}\rangle)/\sqrt{2}$ and just a probabilistic mixture of either $|\text{living cat}\rangle$ or $|\text{dead cat}\rangle$ to a *quantum measurement*, we will observe distinctly different results.

Another natural question to ask here is the following. We don't see superposition states in everyday life — and yet they are fully compatible with the canons of quantum mechanics. Why is that so? As we shall see in the next chapter, this is because superpositions of macroscopically distinct states are extremely fragile and quickly transform into one of their components — in the case of Schrödinger's cat, into either the dead state or the alive state. In the microscopic world, on the other hand, superposition states are relatively robust and are necessary for its physical description. The need to deal with entities whose very existence is in conflict with our everyday experience is one of the reasons why quantum mechanics is so difficult to comprehend.

Exercise 1.1. What is the normalization factor $\mathcal{N}$ of the state of the Schrödinger cat $|\psi\rangle = \mathcal{N}[2|\text{alive}\rangle + \mathrm{i}|\text{dead}\rangle]$ that ensures that $|\psi\rangle$ is a physical state?

Exercise 1.2. What is the dimension of the Hilbert space associated with one motional degree of freedom of a massive particle?
Hint: If you think the answer is obvious, check the solution.

1.3 Polarization of the photon

We will begin studying quantum mechanics with one of the simplest physical systems: the polarization of the photon[4]. The dimension of its Hilbert space is just two, yet it is quite sufficient to show how amazing the world of quantum mechanics can be.

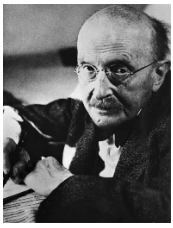
Suppose we can isolate a single particle of light, a photon, from a polarized wave. The photon is a microscopic object and must be treated quantum-mechanically. We begin this treatment by defining the associated Hilbert space. We first notice that the horizontally polarized photon state, which we denote by $|H\rangle$, is incompatible with its vertical counterpart, $|V\rangle$: an $|H\rangle$ photon can never be detected in a $|V\rangle$ state. That is, if we prepare a horizontally polarized photon and send it through a polarizing beam splitter (with the properties described in Sec. C.2), it will always be transmitted and never reflected. This means that states $|H\rangle$ and $|V\rangle$ are orthogonal.

[3] This state is sometimes called Schrödinger's cat, after one of the founding fathers of quantum physics, Erwin Schrödinger. But in fact, Schrödinger discussed a more complex entity, see Box 2.5.

[4] If you are not familiar with the polarization of an electromagnetic wave, this is a good place to read the first two sections of Appendix C.

Box 1.1 Discovery of the photon

In 1900, *Max Planck* explained the experimentally observed spectrum of blackbody radiation by introducing the quantum of light, now known as the photon*. He found that a good agreement between theory and experiment can be obtained if one assumes that the energy of the photon is proportional to the frequency ω of the light wave. The proportionality coefficient, $\hbar = 1.054\,571\,48 \times 10^{-34}$, became known as Planck's constant.



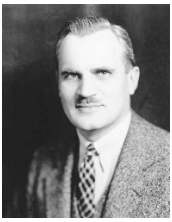
Max Planck

In 1905, *Albert Einstein* reconfirmed the validity of Planck's formula

$$E = \hbar\omega$$

by using it to explain quantitatively the experimental results on the photoelectric effect (see Box 4.6 for more details)**. Later, in 1916, Einstein argued that, since it is known from classical electrodynamics*** that an electromagnetic wavepacket carrying energy E also carries momentum $p = E/c$, the same must be true for photons. From Planck's formula he found† $p = \hbar\omega/c$. Expressing the frequency of the wave in terms of its wavelength, $\omega = 2\pi c/\lambda$, he then wrote

$$p = 2\pi\hbar/\lambda\,.$$



Arthur Compton

Arthur Holly Compton used Einstein's findings in 1923 to provide a theoretical explanation for his own experiments in which he studied the scattering of X rays on free electrons††. By treating X ray photons as high-energy particles, he applied the laws of momentum and energy conservation to the collision between a photon and an electron to calculate the scattered photon energy as a function of the scattering angle. He then related that energy to the wavelength, thereby obtaining a theoretical fit to his experimental data. The excellent agreement he observed serves as an explicit proof of the photon's existence.

Curiously, the term "photon" did not exist at that time. It was introduced later, in 1926, by the physical chemist Gilbert Lewis.

*M. Planck, *Über das Gesetz der Energieverteilung im Normalspectrum*, Annalen der Physik **4**, 553 (1901).

A. Einstein, *Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt*, Annalen der Physik **17, 132 (1905).

***This phenomenon manifests itself, in particular, through the effect of radiation pressure, which was observed experimentally by Peter Lebedev in 1900.

† The expression for the photon momentum can also be obtained as follows. Using Einstein's famous relation $E = Mc^2$ together with Planck's formula, we can calculate the mass of the photon, $M = \hbar\omega/c^2$. The photon moves with the speed of light, and hence its momentum is $p = Mc = \hbar\omega/c$.

†† A. H. Compton, *A Quantum Theory of the Scattering of X-Rays by Light Elements*, Physical Review **21** 483 (1923).

†††G. N. Lewis, *The conservation of photons*, Nature **118**, 874 (1926).

A light wave whose electric field is given as a function of space and time by [see Eq. (C.2)]

$$\vec{E}(z,t) = \mathrm{Re}[(A_H e^{i\varphi_H}\hat{i} + A_V e^{i\varphi_V}\hat{j})e^{ikz-i\omega t}] \tag{1.1}$$

(with real $A_{H,V}$ and $\varphi_{H,V}$) consists of photons in the state[5]

$$|\psi\rangle = \frac{1}{\sqrt{A_H^2 + A_V^2}}\left(A_H e^{i\varphi_H}|H\rangle + A_V e^{i\varphi_V}|V\rangle\right)e^{-i\omega t}. \tag{1.2}$$

For example, if $A_H = A_V$ and $\varphi_H = \varphi_V = 0$, the associated classical wave is $\vec{E} = \mathrm{Re}[A_H(\hat{i}+\hat{j})e^{ikz-i\omega t}]$, i.e., linearly polarized at $+45^\circ$. Accordingly, the state $(|H\rangle + |V\rangle)/\sqrt{2}$ (where the factor of $\sqrt{2}$ is due to normalization) denotes a single photon with $+45^\circ$ linear polarization. Some further examples are listed in Table 1.1[6].

It follows that states $|H\rangle$ and $|V\rangle$ form an orthonormal basis in the Hilbert space of photon polarization states — so this space is two-dimensional. To begin with, these states are orthogonal and thus linearly independent (Ex. A.17). Furthermore, any polarized classical wave can be written in the form (1.1), so any polarization state of the photon can be written in a similar way to (1.2), i.e., as a linear combination of the states $|H\rangle$ and $|V\rangle$. We will call the basis $\{|H\rangle, |V\rangle\}$ the *canonical* basis of our Hilbert space.

Table 1.1 Important polarization states.

state	matrix	description	notation
$\|H\rangle$	$\begin{pmatrix}1\\0\end{pmatrix}$	horizontal	$\|H\rangle$
$\|V\rangle$	$\begin{pmatrix}0\\1\end{pmatrix}$	vertical	$\|V\rangle$
$\cos\theta\|H\rangle + \sin\theta\|V\rangle$	$\begin{pmatrix}\cos\theta\\\sin\theta\end{pmatrix}$	linear polarization at angle θ to horizontal	$\|\theta\rangle$
$\frac{1}{\sqrt{2}}(\|H\rangle + \|V\rangle)$	$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$	diagonal, $+45^\circ$ polarization	$\|+45^\circ\rangle$ or $\|+\rangle$
$\frac{1}{\sqrt{2}}(\|H\rangle - \|V\rangle)$	$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}$	(anti-)diagonal, -45° polarization	$\|-45^\circ\rangle$ or $\|-\rangle$
$\frac{1}{\sqrt{2}}(\|H\rangle + i\|V\rangle)$	$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\i\end{pmatrix}$	right circular polarization	$\|R\rangle$
$\frac{1}{\sqrt{2}}(\|H\rangle - i\|V\rangle)$	$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\-i\end{pmatrix}$	left circular polarization	$\|L\rangle$

[5] It may appear surprising that Eq. (1.2) carries no information about the position of the photon along the z axis. The reason is that the photon, as a quantum particle, is smeared across space and time, potentially to a large extent. Among the factors affecting the spread are the properties of the source, as well as the "quantization volume" chosen for the theoretical analysis. In the case of a coherent laser beam, the photon length is limited by the coherence length of the laser, which can be many kilometers. In this book, we will usually assume that the photons are spread over a distance that is much larger than the size of any apparatus, and can therefore be treated as infinitely long.

[6] See footnote 1 on page 289 for a discussion of conventions for circularly polarized states.

Exercise 1.3. Show that

a) $\pm 45^\circ$ polarization states form an orthonormal basis;
b) right and left circular polarization states form an orthonormal basis.

Exercise 1.4. Decompose $|H\rangle$ and $|V\rangle$ into the $\{|+\rangle, |-\rangle\}$ and the $\{|R\rangle, |L\rangle\}$ bases.

Exercise 1.5. Decompose $|a\rangle = |+30^\circ\rangle$ and $|b\rangle = |-30^\circ\rangle$ in the $\{|H\rangle, |V\rangle\}$, $\{|+\rangle, |-\rangle\}$, and the $\{|R\rangle, |L\rangle\}$ bases. Find the inner product $\langle a|\, b\rangle$ using matrix multiplication in all three bases. Do they come out the same?

Let me clarify a possible confusion. For linearly polarized photons, there is a continuum of polarization angles. But in the case of one-dimensional particle motion, discussed in the previous section, there is also a continuum of position states. Why do we say that one of these Hilbert spaces has dimension two and the other infinity?

The difference is that linearly polarized states can be written in the form (1.2), i.e., as a superposition of other linearly polarized states. If we place a polarizing beam splitter (Sec. C.2), which transmits only horizontally polarized photons, in the way of a diagonally polarized wave, a part of it will be transmitted. This means that a diagonally polarized photon can be detected in the horizontal polarization state.

The states associated with different positions, in contrast, are all orthogonal: a particle prepared in the state $|x = 3\text{ m}\rangle$ cannot be observed at $x = 4$ m. Nor can a position state be written as a superposition of other position states. Accordingly, the corresponding Hilbert space would need to have a much larger basis than the Hilbert space of polarization states.

For a classical wave (1.1), shifting the phases of both the horizontal and the vertical component by the same amount (i.e., $\varphi_H \to \varphi_H + \varphi_0,\ \varphi_V \to \varphi_V + \varphi_0$, which is equivalent to multiplying the right-hand side by $e^{i\varphi_0}$) does not change the polarization of the wave.

A similar rule applies to quantum states. Multiplying a state vector by $e^{i\phi}$ does not change the physical nature of a state. For example, $|V\rangle$, $i|V\rangle$, and $-|V\rangle$ represent the same physical object, as do, say, $|R\rangle = (|H\rangle + i|V\rangle)/\sqrt{2}$ and $e^{-i\pi/2}|R\rangle = (|V\rangle - i|H\rangle)/\sqrt{2}$. For this reason, we will neglect the factor $e^{-i\omega t}$ in Eq. (1.2) for the time being.

We call the complex quantity $e^{i\varphi}$ with a real φ a *phase factor*. Multiplying a quantum state by a phase factor is called *applying a phase shift* by φ. In the language of this definition, we say that applying a phase shift to a quantum state does not change its physical properties. As we shall see in the next section, this rule turns out to be very general: it works for all physical systems, not only electromagnetic waves. Of course, the phase shift must be of an *overall* nature: if we apply it to only a part of a state, that state will change. For example, if we apply a $\pi/2$ phase shift to the vertical component of the $+45^\circ$ polarized photon, $|+\rangle = (|H\rangle + |V\rangle)/\sqrt{2}$, we will obtain $(|H\rangle + i|V\rangle)/\sqrt{2} = |R\rangle$ — a right circularly polarized photon, i.e., a physically different object.

The photon polarization is a realization of the *quantum bit (qubit)*. This term is used to denote any physical system whose Hilbert space is two-dimensional in the context of viewing it as an information carrier. The qubit is the basic unit of quantum information — in analogy to the bit being the unit of information in classical computers. In contrast to the latter, a quantum bit can be, not only in one of the two basis states, but also in their superposition. This enables a number of new technological opportunities that we will discuss throughout this book.

1.4 Quantum measurements

1.4.1 The Measurement Postulate

The second Postulate deals with *quantum measurements*, i.e., experiments whose aim is to obtain information about the quantum state of a system. In classical, macroscopic physics, measurements are more a matter of technology than fundamental science. This is because we can precisely measure the state and the evolution of a system without disturbing it. For example, a soccer ball will not fly differently depending on whether the stadium is empty or full of cheering spectators — so we don't need to know what technique is used to observe its trajectory in order to study the laws of its motion.

In the quantum world, the situation is different: we are big and the things we want to measure are small. Therefore, any measurement will most likely change the quantum state of our system. More generally, quantum measurements are events in which the state of a microscopic quantum object affects that of a macroscopic apparatus. As such, measurements involve crossing the boundary between the quantum and classical domains of physics. As we know, the laws governing these two realms are very different. In order to have a unified picture of the world, we need to understand when and how the transition between these two "jurisdictions" occurs.

Furthermore, phenomena in which a quantum state of something microscopic affects something macroscopic are not limited to laboratories. They range from thermodynamic phase transitions and lasing to hurricanes, the birth of black holes, and perhaps the emergence of the universe itself. The physics of such phenomena is quite similar to that of quantum measurements. Understanding this physics is hence essential for learning the nature of the world around us.

The main principles of the Measurement Postulate are quite intuitive. Suppose a photon in state (1.2) hits a polarizing beam splitter (PBS) — an optical element that transmits horizontally polarized light but reflects vertically polarized [Fig. 1.2(a)]. What will happen to that photon? If we were dealing with a classical wave (1.1), we would expect it to split: a part would be transmitted through the PBS, and the remainder reflected. The fractions of energy going into the transmitted and reflected channels would be proportional to A_H^2 and A_V^2, respectively. But the photon is the smallest energy portion of light, and cannot be divided into parts.

We have come to an apparent contradiction. On the one hand, we know that a classical wave, which consists of photons, divides. On the other hand, every individual photon is indivisible. How can these two imperatives be upheld at the same time?

It seems that the only way to solve the conundrum is to postulate that the outcome will be *random*: the photon will be transmitted through the PBS with probability $\mathrm{pr}_H = A_H^2/(A_H^2 + A_V^2) = |\langle H|\,\psi\rangle|^2$, and reflected with probability $\mathrm{pr}_V = A_V^2/(A_H^2 + A_V^2) = |\langle V|\,\psi\rangle|^2$. In this way, if a large number N of photons are incident on the PBS, the number ratio of the transmitted and reflected energies will be A_H^2/A_V^2, as expected classically (see Sec. C.2). And yet, no individual photon is divided.

As we know, the part of the classical wave that is transmitted through the PBS is horizontally polarized — that is, all photons making up the wave are of horizontal polarization. The same is true for the reflected wave: all its photons are vertically polarized. But then, the same must be true if the photons are sent to the PBS one-by-one. Not only will the photon randomly choose its path, but also, in a quite Orwellian fashion, it will *change its state* to conform with the path chosen. After the PBS, the photon state in the transmitted channel will become $|H\rangle$, and in the reflected channel $|V\rangle$. If we place a series of additional PBS's in the transmitted channel of the first PBS, the photon will be transmitted through all of these PBS's — there will be no further randomness.

The process I just described constitutes the *polarization state measurement* of a photon. To complete it, we place single-photon detectors (Box 1.2) into both output channels of the PBS. Of these two detectors, only one will click, thereby providing us with the information about the photon's polarization [Fig. 1.2(a)].

The above measurement apparatus is designed to distinguish between the horizontal and vertical polarizations. One can think of other designs as well. For example, by tilting the PBS by 45°, we can have it transmit $|+\rangle$ and reflect $|-\rangle$, so if we send an arbitrary state $|\psi\rangle$, it will transmit or reflect with probabilities $\mathrm{pr}_+ = |\langle +|\,\psi\rangle|^2$ and $\mathrm{pr}_- = |\langle -|\,\psi\rangle|^2$, respectively. More generally, we can construct a measurement apparatus that would distinguish between any two polarization states, as long as these states are orthogonal to each other.

We are now ready to formulate our Postulate.

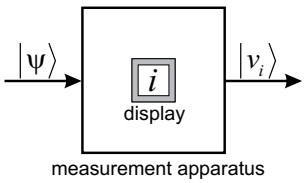


Fig. 1.1 A theoretician's picture of a quantum measurement.

Measurement Postulate. An idealized measurement apparatus is associated with some orthonormal basis $\{|v_i\rangle\}$. After the measurement, the apparatus will randomly,

with probability

$$\mathrm{pr}_i = |\langle v_i|\ \psi\rangle|^2, \tag{1.3}$$

where $|\psi\rangle$ is the initial state of the system, point to one of the states $|v_i\rangle$. The system, if not destroyed, will then be converted (*projected*) onto state $|v_i\rangle$ (Fig. 1.1).

A quantum measurement that proceeds in accordance with the above Postulate is called a *projective measurement*. The projection of the state measured onto one of the basis elements is also called *collapse* of the quantum state. Equation (1.3) is called *Born's rule*.

The probabilistic behavior of quantum objects led to a lot of controversy at the time quantum mechanics was founded. This is because, by the end of the 19th century, the principle of *determinism* was universally accepted: physicists believed that, if the initial conditions of a given quantum system are known precisely enough, its future evolution can be predicted arbitrarily well. Quantum physics breached this fundamental belief, and many physicists found it extremely difficult to accept. For example, Albert Einstein made a famous statement that "God does not play dice" and came up with a brilliant *Gedankenexperiment*[7] showing that the postulates of quantum mechanics are in contradiction with common sense. We will study this Gedankenexperiment in the next chapter and see that quantum randomness can be attributed to observers themselves being quantum objects, but not being able to verify their own quantum nature experimentally. For now, however, let us accept quantum randomness as a postulate corroborated by vast experimental evidence.

Exercise 1.6. Show mathematically that, for a state $|\psi\rangle$, the sum of detection probabilities (1.3) for all basis elements is $\langle\psi|\ \psi\rangle$, i.e., it equals 1 if the state is physical.

Exercise 1.7. Show that applying an overall phase factor to a quantum state will not change the probabilities of its measurement results — in agreement with the fact that this phase has no influence on the physics of a state, as discussed in the previous section.

1.4.2 Polarization measurements

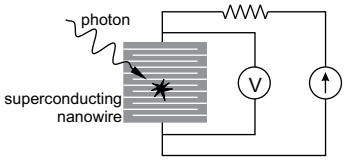
Above, we discussed the fact that one can rotate the PBS to modify the apparatus of Fig. 1.2(a) so that it can measure the polarization in a non-canonical, linearly polarized basis. However, the photon reflected from the PBS will not propagate in the horizontal direction, and this is not convenient in a practical tabletop experiment (Box 1.3). Therefore most experimentalists take advantage of the optical element called a waveplate[8] which interconverts polarization states of a photon from one to another. Here are some examples.

Exercise 1.8. Show that:

[7] "Gedankenexperiment" is the German for "thought experiment".

[8] This is a good place to read the third section of Appendix C.

Box 1.2 How to detect a photon?



A photon detector is a device that converts a photon into a "click" — a macroscopic pulse of electric current or voltage. Making such an extremely sensitive device is a challenging technological task. This figure sketches one of the modern ways of addressing this challenge: the superconducting single-photon detector.

The sensitive area of the detector is a nanowire that is cooled down to a superconducting state, with a small constant current flowing in it. The nanowire is so thin that, when it absorbs even a single photon, it warms up enough to become resistive in part of its length. The current will then heat up this area as predicted by Joule's law, further destroying superconductivity around it. In this way, a kind of avalanche process develops, in such a way that the entire nanowire becomes resistive for some time. This resistance leads to a pulse in the voltage across the nanowire that is easily detectable.

This detector suffers from a few imperfections that are typical of practical photon detectors. First, the detector is *non-discriminating*: its response to a pulse containing multiple photons is the same as its response to a single photon. This is because the entire nanowire will lose superconductivity and acquire the same resistance no matter how many photons are absorbed. Second, a photon incident on the detector may get reflected, thereby generating no click. The probability that a click will occur in response to a photon is known as the *quantum efficiency* of the detector. In some modern detectors, this parameter exceeds 99%. Finally, a detector may produce a click even in the absence of a photon. The frequency of such *dark counts* is another important technical characteristic of this device.

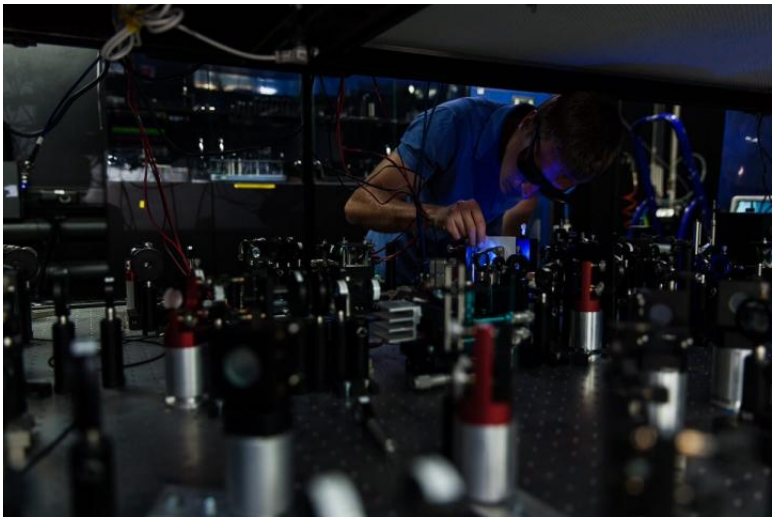
a) the setup in Fig. 1.2(b) performs the photon polarization measurement in the diagonal ($|\pm 45^\circ\rangle$) basis;
b) the setup in Fig. 1.2(c) performs the measurement in the circular ($\{|R\rangle, |L\rangle\}$) basis.

Hint: When a piece of apparatus described in the Measurement Postulate is measuring one of its own basis states $|v_i\rangle$, the measurement will point to that state with probability 1. Conversely, if the apparatus can distinguish a particular orthonormal set of states with certainty, we can conclude that this set is the measurement basis of the apparatus. Therefore, to solve this exercise, it is enough to show that the basis states [i.e., $|\pm 45^\circ\rangle$ in (b) and $|R\rangle, |L\rangle$ in (c)], when sent onto the PBS, will generate clicks in different photon detectors.

Exercise 1.9[§] Each of the states $|H\rangle, |V\rangle, |+\rangle, |-\rangle, |R\rangle, |L\rangle$ is measured in

a) canonical,
b) diagonal,
c) circular

Box 1.3 Optical table



This photograph shows a typical quantum optical experiment. It is performed on an *optical table* — a massive metal plate upon which one mounts various optical elements, such as lenses, mirrors, lasers, crystals, and detectors. The beams typically run horizontally, at the same level throughout the entire table.

bases. Find the probabilities of the possible outcomes for each case.
Answer: For each state, when the measurement is performed in the basis to which the state belongs, the probabilities are 0 and 1. If the state does not belong to the measurement basis, the probabilities of both outcomes are $\frac{1}{2}$.

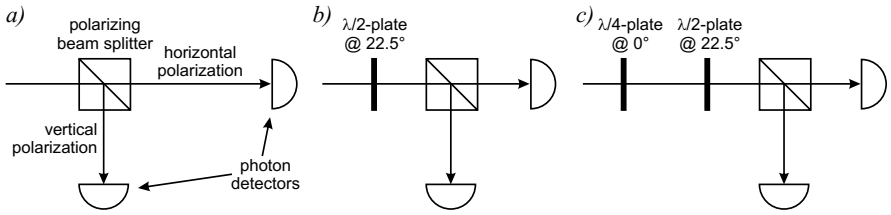


Fig. 1.2 Photon polarization measurements in the canonical $\{|H\rangle, |V\rangle\}$ (a), diagonal $\{|+\rangle, |-\rangle\}$ (b), and circular $\{|R\rangle, |L\rangle\}$ (c) bases.

Exercise 1.10. Propose a scheme for a quantum measurement in the basis $\{|\theta\rangle, |\frac{\pi}{2}+\theta\rangle\}$.

Exercise 1.11. Propose a scheme for a quantum measurement in the basis $\{|R\rangle, |L\rangle\}$ that would use just one waveplate.

Exercise 1.12. Consider a photon that is not in a superposition state, but in a random *statistical mixture*, or *ensemble*[9]: *either* $|H\rangle$ with probability $1/2$ *or* $|V\rangle$ with probability $1/2$. The polarization of this photon is measured in

a) canonical,
b) diagonal,
c) circular

bases. Find the probabilities of the possible outcomes for each case.

Exercise 1.13. A photon is prepared with a linear polarization 30° to horizontal. Find the probability of each outcome if its polarization is measured in (a) the canonical, (b) the diagonal, and (c) the circular basis.

Exercise 1.14. A photon in state $|\psi\rangle = (|H\rangle + e^{i\varphi}|V\rangle)/\sqrt{2}$ is measured in the diagonal basis. Find the probability of each outcome as a function of φ.

This exercise, along with Ex. 1.7, shows once again the important difference between a phase factor applied to a part of a quantum state or applied to the whole. In the former case the added phase has an effect on the measurable properties of the object; in the latter, it doesn't.

Although a single measurement provides us with some information about the initial state of a quantum system, this information is very limited. For example, suppose we have measured a photon in the canonical basis and found that it has been transmitted through the PBS. Does this tell us that the initial photon was in the state $|H\rangle$? No. It could have been in any state $\psi_H|H\rangle + \psi_V|V\rangle$; as long as $\psi_H \neq 0$, there is some probability of getting a click in the transmitted channel. So the only thing we learn from this measurement is that the photon was not vertically polarized.

Suppose now we have performed the same measurement many times, every time preparing our photon in the same state[10]. Now we know much more! We know how many clicks we obtained from the "horizontal" detector, and how many from the "vertical" one — that is, we have *measurement statistics*. From these, we can calculate, with some error, $\mathrm{pr}_H = |\psi_H|^2$ and $\mathrm{pr}_V = |\psi_V|^2$, i.e., learn about the absolute values of the state components. But both ψ_H and ψ_V are complex numbers, and their arguments are still unknown. For example, if we observe $\mathrm{pr}_H = \mathrm{pr}_V = 1/2$, the state $|\psi\rangle$ could be $|R\rangle$ or $|L\rangle$ or $|+\rangle$ or $|-\rangle$, or many other options. What can we do about this?

As you will see in the following exercise, it is helpful to perform additional sets of measurements in other bases. From the statistics acquired, we obtain additional equations, which can be solved to find ψ_H and ψ_V up to an uncertainty associated with a common phase factor.

[9] Such *mixed states* are not elements of the quantum Hilbert space. More detail on this in Sec. 2.2.4.

[10] Although we don't know what the state is, we can make sure we can repeatedly prepare the photon in the same state by setting up identical experimental conditions.

Exercise 1.15. Suppose multiple polarization measurements of photons identically prepared in the state $|\psi\rangle$ are carried out in the canonical, diagonal, and circular bases, and all six respective probabilities (pr_H, pr_V, pr_+, pr_-, pr_R, pr_L) are determined. Show that this information is sufficient to fully determine $|\psi\rangle$ and express its decomposition in the canonical basis through pr_H, pr_+, and pr_R. Give an example to show that measuring just in the canonical and diagonal bases would be insufficient — that is, find two different states that would yield the same pr_H and pr_+.

This method of obtaining complete information about the quantum state by performing series of measurements in several different bases on the state's multiple identical copies is called *quantum state tomography*. It can be generalized to other quantum systems, including those of higher dimension. We discuss quantum tomography in more detail at the end of the book (Sec. 5.7).

Exercise 1.16. Suppose you are given a *single* copy of a quantum system that is in one of the two non-orthogonal states $|a\rangle$ and $|b\rangle$. You know what these states are, but you don't know which one the system is in.

a) Show that it is not possible to construct a piece of measurement apparatus that would always reveal the system's state with certainty.

b)* Show that it is possible to construct a measurement device that would produce, with some probability, outcomes of three types: "definitely $|a\rangle$", "definitely $|b\rangle$", and "not sure", and the outcomes of the first two types would always be correct. **Hint:** It may be helpful to use a non-polarizing beam splitter — an optical element that randomly reflects or transmits a photon independently of its polarization.

1.5 Quantum interference and complementarity

Consider the experiment displayed in Fig. 1.3. A single photon, initially in the diagonal polarization state $|+\rangle = (|H\rangle + |V\rangle)/\sqrt{2}$, enters an arrangement known as an *interferometer*[11]. First, a PBS transmits the horizontal component of the state and reflects the vertical one. The reflected component then propagates through a variable delay line,[12] after which the two components are recombined by means of another PBS. Subsequently, the interferometer output state is subjected to a measurement in the diagonal basis.

The delay line introduces a difference between the optical path lengths of the vertical and horizontal components. If the length of the delay line is l, the vertical component will acquire a phase shift of $\varphi = kl$ with respect to the horizontal one, where $k = 2\pi/\lambda$ is the wavenumber. As a result, the photon, when exiting the interferometer, will be in the state $|\psi\rangle = (|H\rangle + \mathrm{e}^{\mathrm{i}\varphi}|V\rangle)/\sqrt{2}$.

[11] More specifically, the Mach-Zehnder interferometer.

[12] The delay line is assumed much shorter than the length of the photon pulse, so the interference visibility remains constant when the delay is varied.

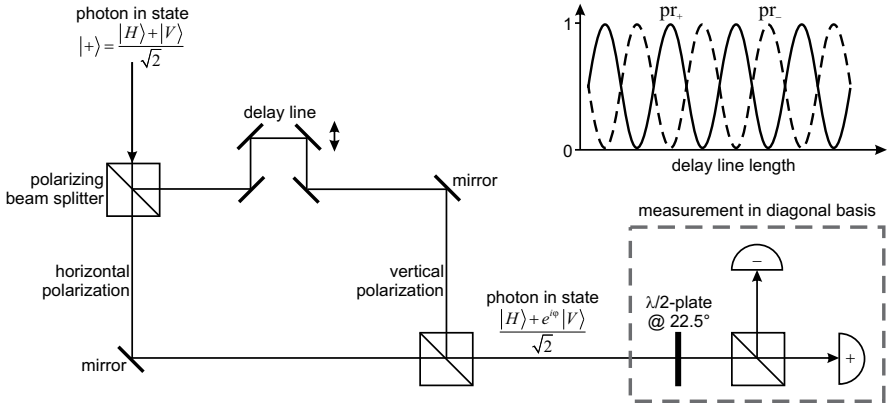


Fig. 1.3 A single-photon interference experiment. Inset: event probabilities for the two detectors as a function of the interferometer path length difference.

We studied the measurement of this state in Ex. 1.14 and found that the probabilities for the detectors "+" and "−" to click are $\mathrm{pr}_{\pm} = \frac{1}{2}(1 \pm \cos\varphi)$, respectively. When we vary the length of the delay line, the probabilities will vary sinusoidally. In other words, we will observe *interference fringes* — of the same kind that a macroscopic wave would exhibit in this optical arrangement.

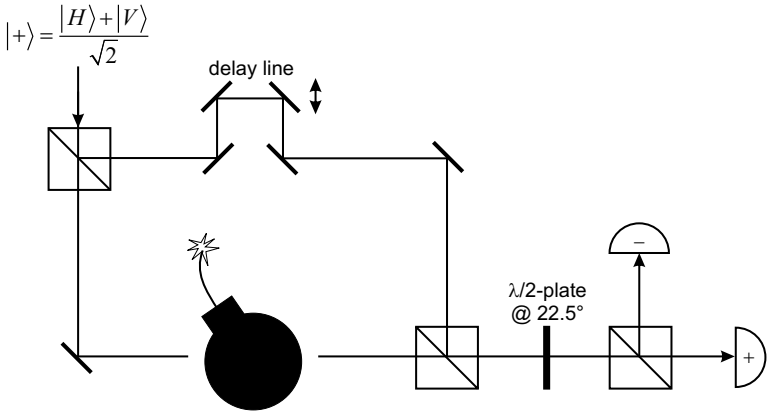
What is truly remarkable about this conclusion (and it has, of course, been thoroughly confirmed experimentally) is that the interference fringes are generated by a single photon. This is in drastic conflict with intuition. Indeed, in a classical experiment, interference would occur because the two waves, traveling along the two paths of the interferometer, acquire different phases and then add coherently on the photodetectors. But in our experiment only one photon is present. A photon is an indivisible elementary particle of light — and hence it cannot split[13] in the interferometer to generate the two waves required to produce interference fringes. It must be traveling along *either* the upper *or* the lower interferometer path — but never the two at the same time.

However reasonable and intuitive, this "devil's advocate" argument turns out to be inconsistent with both our calculation and the experimental observation. How can this be explained?

The photon entering the interferometer is in a superposition of the vertical and horizontal polarization states. After the first PBS, it remains in a superposition — but now, it is also a superposition of the upper and lower interferometer paths. Once the paths are recombined, the superposition is again transformed into that of polarization states — albeit with a phase shift on one of its components. It is these two

[13] We will see later that a photon may in fact split into two photons of lower energy through a nonlinear optical phenomenon known as parametric down-conversion. However, this effect is quite exotic, occurring with low probability and only under special conditions. Our interferometer contains no nonlinear optical elements, so parametric down-conversion is irrelevant here.

Box 1.4 Quantum weapon inspection



Here is an exciting paradox associated with the single-photon interference experiment discussed in Sec. 1.5*. Suppose there is a "bomb" equipped with a photon sensor, so that it will explode if even a single photon interacts with it. Can we detect the presence of the bomb in one of the arms of our interferometer without detonating it?

Let us set up the delay line in our single-photon interferometer (Fig. 1.3) such that $\varphi = 0$. Then, if the bomb is *absent*, every incoming photon will leave the interferometer polarized at $+45^\circ$ and cause an event in detector "+". Detector "−", on the other hand, will never click.

Now if the bomb is *present*, as shown in the figure above, it may explode or not, depending on which way the photon goes. In this way, the bomb implements a Welcher-Weg measurement. Accordingly, the photon will behave like a particle that goes randomly into either the lower or the upper part of the interferometer. If it goes into the lower path, the bomb will explode. But if it goes into the upper part, the bomb will remain intact and the photon will exit the interferometer in the vertical polarization state. When measured in a diagonal basis, this photon will be equally likely to generate an event in either of the two detectors.

Hence, if the bomb is present, there will be a nonzero probability of hearing a click in detector "−". Moreover, this detector can click *only* in the presence of the bomb. If this detector does click, we know for certain that the bomb is present — without having interacted with it!

The above setup is not a perfect tool for weapon inspection, as it does not guarantee a conclusive result, nor that the bomb will not detonate (Ex. 1.17). However, if one places the bomb in a high-finesse Fabry-Perot interferometer rather than a Mach-Zehnder interferometer, one can achieve an efficiency close to 100%. In this case, the photon will likely pass through the interferometer when the bomb is absent, but reflect if the bomb is present.

*A. C. Elitzur, L. Vaidman, *Quantum mechanical interaction-free measurements*, Foundations of Physics **23**, 987 (1993).

components of the superposition that play the role of the two waves in the classical experiment, interfering with each other. This is known as the *wave-particle duality* of quantum particles[14].

So, in a sense, the photon does get divided between the two interferometer channels. However, this wavelike behavior is only possible if the components remain in the superposition state. To illustrate this, let us suppose that we place non-destructive detectors in both interferometer arms, able to register the presence of the photon without destroying it. Every time a photon is sent into the interferometer, one of these detectors will "click", indicating whether the photon went through the upper or the lower path. In this way, as the founding fathers of quantum mechanics would say, we obtain *Welcher-Weg* (which-way) information about the photon.

Obtaining Welcher-Weg information means *measuring* the location of the photon. As we learned in the previous section, such a measurement will collapse the superposition state onto the photon being either in the upper or lower path of the interferometer. By looking at the Welcher-Weg detector, the observer is able to tell with certainty whether the photon will leave the interferometer in the horizontal or vertical state. In either case, a subsequent measurement of that photon in the diagonal basis will yield either outcome with probability $1/2$, with no dependence on the path-length difference. The Welcher-Weg measurement destroys the wavelike property of the photon and makes it behave like a particle.

This is, of course, the case even if the observer does not look at the Welcher-Weg detectors. The photon is then in a mixed state of being *either* in the upper or lower path of the interferometer with probability $1/2$, but no longer in the superposition state. That is, we are now in the situation of Ex. 1.12 rather than 1.14. The photon state has lost its *quantum coherence* — a well-defined phase relation between the superposition terms. Hence it is no longer able to exhibit interference.

This Gedankenexperiment demonstrates *quantum complementarity* — a general principle of quantum physics stating that objects may have complementary properties which cannot be observed or measured at the same time. We can have either the Welcher-Weg information or interference, but not the two together.

Exercise 1.17. In the setting of Box 1.4, what are the probabilities of

a) detecting a bomb without detonating it,
b) detonating the bomb,
c) obtaining an inconclusive result without detonating the bomb?

1.6 Quantum cryptography

We can now discuss the first application of quantum physics in this course. This application is to *cryptography* — the art of exchanging secret messages over insecure channels.

[14] This is probably why popular quantum books like to describe superposition states as ones in which "an object is in two different places at the same time".

Box 1.5 Classical cryptography

Cryptography is easily implemented if the communication parties, which we call Alice and Bob, share a prearranged, secret data set (a sequence of 0's and 1's) known as *secret key* or *one-time pad*. With this resource available, a cryptographic protocol can proceed as follows. Alice chooses a piece of the secret key which has the same length (i.e., the same number of bits) as the message she wishes to send to Bob. She then applies an XOR (exclusive OR, or bitwise sum modulo 2) operation to every bit of her message and the corresponding bit of her secret key:

original message `01110011...`
XOR
secret key `10011010...`
encrypted message `11101001...`

In this way she obtains an *encrypted message* which can be safely transmitted over an insecure channel, as it cannot be decrypted by anyone who is not privy to the secret key. Bob, on the other hand, can easily decrypt the message. To this end, he applies XOR to every bit of the encrypted message he receives and the corresponding bit of the secret key, thereby recovering the original message.

encrypted message `11101001...`
XOR
secret key `10011010...`
recovered original message `01110011...`

This protocol, known as *private-key cryptography*, is very secure and simple; it has been known for hundreds of years. The trouble is, it is not easy for Alice and Bob to arrange sharing random data that would be secret to everyone else. As a rule, the only safe way to do this would be to send a courier carrying a briefcase loaded with random data. This is, of course, very expensive. For this reason, private-key cryptography is only used in the most sensitive government and commercial communications.

For other applications, such as e-commerce, a family of protocols known as *public-key cryptography* is used. Without going into details, these protocols rely on the existence of "one-way" functions that are easy to compute, but very difficult to invert. For example, multiplying two prime numbers containing a few dozen digits will take microseconds on a modern computer, but factoring a number of similar length will take months or years. Public-key cryptography protocols rely on one-way functions to enable secure communication between parties who have never had an opportunity to exchange a secret key.

While public-key protocols are convenient and inexpensive, they are not perfectly secure. The computational power available to us doubles every year or two, so a calculation that takes years at present may take only hours a few years in the future. Furthermore, *quantum computers* (Sec. 2.5) are potentially capable of cracking the security of public-key protocols almost instantly.

Known since ancient times, cryptography is now a major branch of the telecommunications industry, aimed at protecting the privacy and information security of individuals, businesses and government entities. Box 1.5 reviews classical approaches to cryptography. To sum it up, within the classical domain we are compelled to choose between private-key cryptography, which is secure but expensive, and public-key cryptography, which is cheap, but not perfectly secure.

Quantum mechanics offers us a solution that takes "the best from both worlds". On the one hand, its security is guaranteed by fundamental laws of nature. On the other hand, it does not require random information to have been shared previously between the parties.

1.6.1 The BB84 protocol

Quantum cryptography, or, more precisely, *quantum key distribution*, relies on the property of measurements to alter the quantum state they are used on. The idea is that the sending party (Alice) sends secret data to the receiving party (Bob) by means of single photons, encoding the data in their quantum states. Anyone who tries to eavesdrop on this transmission will either destroy or alter these photons, thereby revealing themselves.

The best known quantum cryptography protocol is named "BB84" after its inventors C.H. Bennett and G. Brassard[15]. To implement it, Alice and Bob perform the following operations.

1. Alice tosses a coin to randomly choose the value of a bit, either 0 or 1, to be sent.
2. Alice tosses a coin again to choose the encoding basis, either canonical or diagonal.
3. Alice generates a photon and encodes the bit in that photon's polarization:

$$\begin{cases} 0 \to |H\rangle \\ 1 \to |V\rangle \end{cases} \text{ or } \begin{cases} 0 \to |+45^\circ\rangle \\ 1 \to |-45^\circ\rangle \end{cases}$$

   She then sends the photon to Bob.
4. Bob tosses a coin to choose the measurement basis, either canonical or diagonal.
5. Bob measures the arriving photon in the chosen basis:
   - if he chooses the same basis as Alice, he will detect the same bit value as the one Alice sent;
   - if he chooses the other basis, he will detect a random bit value.

This procedure is repeated many times. Of course, both Alice and Bob must keep record of the bases they used, states sent or detected, and the exact time when the

[15] C. H. Bennett, G. Brassard, "Quantum Cryptography: Public Key Distribution and Coin Tossing", Int. Conf. on Computers, Systems and Signal Processing, Bangalore, India (IEEE, New York, 1984), p. 175.

photons were sent or received. After many thousands of such records have been collected, Alice and Bob inform each other (via a classical, insecure channel) of their choice of bases for each photon, but *not* the bit values they sent or measured. Bob also informs Alice of those instances when he did not detect a photon, e.g., if it has been absorbed in the transmission line (this requires, of course, that the timing of Alice's transmissions be known to Bob, but this information need not be secret). Subsequently, Alice and Bob discard the data for those events in which different bases were used or the photon has been lost.

Alice and Bob now share a string of identical bits, which they can use as the one-time pad in a private-key protocol. To see why this string is guaranteed to be secret, let us suppose an eavesdropper (Eve) cuts the transmission line, intercepts Alice's photons, measures their polarization, and re-sends them to Bob (Fig. 1.4). Will she be able to obtain a copy of the secret key?

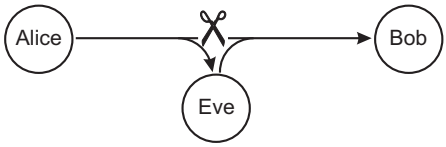


Fig. 1.4 Eavesdropping in quantum cryptography.

The answer is negative. Eve's problem is that, according to the Measurement Postulate, she must measure in a particular basis, and does not know which basis to choose. No matter how she chooses that basis, it will sometimes happen that Alice and Bob work in the same basis and Eve in a different one. But in this case Eve's measurement will alter the photon's state and Bob may not receive the same bit value as the one Alice sent him. The secret keys that Alice and Bob record will end up being different, and this will alert them to the eavesdropping.

Suppose, for example, that Alice and Bob both work in the canonical basis, but Eve in the diagonal basis. Alice sends a horizontally polarized photon, encoding bit value 0. But Eve uses the diagonal basis, so she will detect $|+\rangle$ or $|-\rangle$ with equal probabilities. If she detects and resends either of these states, Bob (who detects in the canonical basis) is equally likely to observe $|H\rangle$ or $|V\rangle$. Bob's observation of $|V\rangle$ will cause him to record a different bit value compared to the one Alice has sent.

In order to check whether the eavesdropper was present, Alice and Bob exchange, via an insecure channel, a part of the secret bit string they obtained. If there are no (or very few) errors, they can use the remainder of that string as the one-time pad.

Exercise 1.18. Suppose Eve intercepts Alice's photons and measures them in either the canonical or diagonal basis (she chooses at random). She then encodes the bit she measured in the same basis and re-sends it to Bob. What *error rate* will Alice and Bob register, i.e., what fraction of bits in the secret key they created will come out differently on average?

This exercise implies that, if Alice and Bob see that the secret key they obtain contains a certain fraction of non-identical bits, they can no longer be sure they are not being eavesdropped. However, the error rate obtained in Ex. 1.18 applies only to one specific attack by Eve. By choosing more sophisticated attack strategies, Eve can obtain a copy of the secret string while creating an even lower error rate in Alice and Bob's records.

So how low does the error rate need to be for Alice and Bob to be sure their communication is secure? It has been proven[16] that the boundary lies at about 11%. No matter what Eve's strategy is, if the error rate is below that value, Alice and Bob can use a procedure called *privacy amplification* to "distill" a perfectly secure, identical secret key from the partially non-identical bit strings they obtained via the quantum protocol.

Exercise 1.19. As discussed, a large fraction of the photons sent by Alice do not reach Bob. But Alice and Bob do not know whether these photons were in fact lost due to the absorption in the line or "stolen" by an eavesdropper. Does this consideration affect the security of quantum key distribution?

1.6.2 Practical matters in quantum cryptography

Quantum cryptography is no science fiction. The protocol described above is fully within reach of present day technology. In fact, there exist commercial quantum cryptography servers that can be connected to commercial fiber optical communication lines and implement the BB84 protocol. Many cities have constructed metropolitan quantum communication networks. Quantum cryptography was used for communications during the 2007 Swiss federal elections and the 2010 FIFA world cup in South Africa. Further examples will emerge as this text is being written.

Still, we do not see quantum key distribution universally replacing classical cryptographic protocols. Is there still a technical obstacle or is this just a matter of mental inertia?

Unfortunately, there do exist unsolved practical issues, the main one being loss in communication lines. This loss follows *Beer's law* $n(L) = n_0 \mathrm{e}^{-\beta L}$, where $n(L)$ is the number of non-absorbed photons at distance L from Alice, and β is the loss coefficient. The best fibers used in telecommunications today involve a loss of about 5% per kilometer. This does not sound like much; yet, when transmitted through an intercity communication line, only a tiny fraction of all photons will reach Bob; the remainder will be lost.

Exercise 1.20. Alice sends a photon to Bob, who is 300 km away, via a fiber line. The fiber has a loss rate of 5% per kilometer.

a) Find the loss coefficient β in that fiber.

[16] P. W. Shor and J. Preskill, *Simple Proof of Security of the BB84 Quantum Key Distribution Protocol*, Physical Review Letters **85**, 441 (2000).

b) What fraction of the photons sent by Alice will reach Bob?

In addition to the loss, there is a problem associated with dark counts (see Box 1.2). It may happen that, e.g., an $|H\rangle$ photon sent by Alice is lost, and, at the same time, Bob's detector in the vertical polarization channel generates a dark event. Bob will then interpret this event as an observation of a $|V\rangle$ photon coming from Alice, and make a record of it. As a result, Alice and Bob will observe an error, and will no longer be confident that their communication is secure.

As long as the transmission line is not too long, there are enough photons reaching Bob so the fraction of errors due to dark counts is small. But the photon rate decreases exponentially with the distance, whereas the rate of dark counts remains constant. So at some point secure transmission will no longer be possible.

This is illustrated in Fig. 1.5. When the communication distance is short, the rate of secure bits distilled by Alice and Bob (dotted lines) follows the rate of photons reaching Bob (solid lines) multiplied by a constant factor. But when this rate decreases, so that the fraction of errors due to dark counts becomes significant, the secure key starts to fall off faster with the privacy amplification protocol becoming less and less efficient. When the number of photons reaching Bob falls below the critical level corresponding to the error fraction of 11%, the transmission is no longer secure.

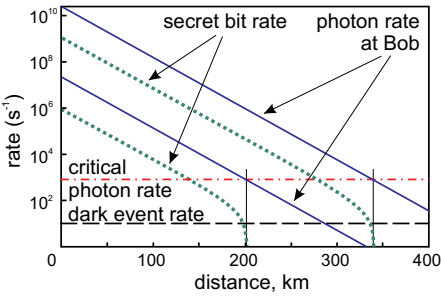


Fig. 1.5 Secure communication rate as a function of distance in the setting of Ex. 1.21.

Exercise 1.21. Assuming that Alice has a perfect single-photon source, sketch the photon transfer rate and the distilled secret bit rate as functions of the distance and estimate the maximum possible secure communication distance given the following parameters:

- photon loss in the fiber communication line: $\beta = 0.05$ km^{-1};
- emission rates of Alice's source: $n_0 = 2 \times 10^7$ and 2×10^{10} photons per second;
- quantum efficiency of the photon detectors: $\eta = 0.1$;

- frequency of dark events that are synchronized with Alice's photons[17] in each of Bob's detectors: $f_d = 10\ \mathrm{s}^{-1}$;

Answer: see Fig. 1.5.

The range of secure quantum communication can be improved by increasing Alice's photon emission rate or reducing the detector dark counts. However, this will not lead to dramatic results: the exponential nature of Beer's law prevents quantum communication at distances beyond a few hundred kilometers. In the setting of Ex. 1.21, increasing the emission rate by three orders of magnitude increases the communication distance by only a factor of 1.7 (Fig. 1.5).

To overcome this limit — and create the "quantum internet" that would cross oceans and eventually cover the entire planet — we need a fundamentally different technology. This technology, known as the *quantum repeater*, is discussed at the end of Chapter 2.

1.7 Operators in quantum mechanics

We now proceed to discussing linear operators, which are a key element of quantum physics[18]. They play a dual role. First, they describe evolution: as time passes, quantum states change, and this change is described mathematically by operators. A second, less obvious application of linear operators is the formal description of quantum measurements. We shall start with the first role in this section.

Exercise 1.22. Find the matrix of the operator $|+\rangle\langle -|$ in the canonical and the $(|R\rangle, |L\rangle)$ bases.

Exercise 1.23. Find, in the canonical basis, the matrix of the linear operator $\hat{A}$ that maps

a) $|H\rangle$ onto $|R\rangle$ and $|V\rangle$ onto $2|H\rangle$;
b) $|+\rangle$ onto $|R\rangle$ and $|-\rangle$ onto $|H\rangle$.

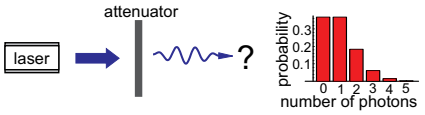
The waveplate, which transforms photon polarization states, is an example of a physical operation that can be associated with a quantum operator. In order to calculate this operator, we need to adopt a convention. As discussed in Sec. C.3, the waveplate changes the relative phase of the extraordinary (parallel to the optic axis) and ordinary (orthogonal to the optical axis) polarization states by an angle $\Delta\varphi$, which is equal to π for a half-wave plate and $\pi/2$ for a quarter-wave plate. In addition, it introduces a common phase shift for the entire wave.

These optical phase shifts transform into quantum phase shifts when applied to the single photon. The overall phase shift, common for all polarization components,

[17] The actual dark count rate can be higher. But because Bob knows the exact timing of Alice's transmission, the only dark count events that contribute to the error rate are those that occur synchronously with the clicks expected due to Alice's photons.

[18] A fuller introduction into linear operators and matrices can be found in Sections A.5 and A.6.

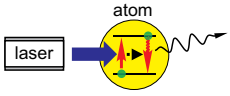
Box 1.6 How to generate a photon?



The most straightforward, but incorrect, answer to this question is as follows: attenuate a laser. Suppose one has a pulsed laser with mean power P and pulse repetition rate R. Then each laser pulse contains $n = P/R\hbar\omega$ photons, where ω is the frequency of the laser radiation. Therefore, one might argue, one could place an attenuator (dark glass) into the laser beam that would reduce its power by a factor of n, so that each pulse would contain precisely one photon.

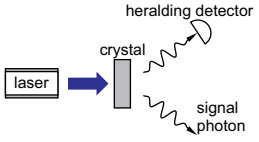
The mistake in this argument is that the photon numbers in the pulses transmitted through the attenuator will be stochastically distributed according to the Poisson distribution (see Sec. B.3). While there may indeed be one photon per pulse on average, this does not mean that each pulse will contain *exactly* one photon. Sometimes there will be zero photons, sometimes one, sometimes two or more.

In spite of this criticism, the attenuated laser provides a useful replacement for a true photon source in some applications. In particular, in practical quantum cryptography, the laser is attenuated to an extremely weak level, so the probability for each pulse to contain even a single photon is quite low. Then the probability that the pulse contains more than one photon is negligible, and hence the communication security is not compromised.



In order to guarantee production of a single photon "on demand", more sophisticated schemes are required. For example, a single two-level atom, when excited, will spontaneously decay back to the ground state, emitting precisely one photon. Practical realization of such a source, however, is challenging. First, one needs to trap a single atom and hold it in position for the duration of the experiment. Second, the photon will be emitted in a random direction. To make the atom emit in a specific direction, physicists sometimes put a Fabry-Perot cavity around it. This arrangement is commonly referred to as *cavity quantum electrodynamics*.

To obviate the need to trap the atom, experiments are conducted with solid atom-like sources, such as single-site defects in crystal lattices or quantum dots. The idea is the same: to consider an object in which only one quantum excitation of a certain energy is allowed. At the time this manuscript is being written, these experiments are evolving rapidly towards better efficiency and reproducibility of the protons produced.



In the meantime, a powerful alternative approach to single photon preparation is offered by *spontaneous parametric down-conversion*. This is a nonlinear quantum optical process that takes place when a strong laser beam propagates through a crystal with nonlinear optical properties. Each photon of the beam can spontaneously split into two photons of lower energy. Such an event is spontaneous and has a very low probability. However, it has a fundamental property: whenever it occurs, it is always a *pair* of photons that is produced. So if we detect one of the photons, we know that its counterpart has been generated and we can experiment with it.

This setup is called the *heralded single-photon source*, because the detection of one photon "heralds" the presence of the other. It is not capable of producing photons on demand; it only announces when a spontaneously emitted photon is produced without destroying it. Hence its application in quantum technology is limited. However, because we do not yet have a reliable on-demand source, heralded sources are widely used in experimental quantum optics research.

can be neglected (see Sec. 1.3). We need to agree on how to treat this common phase shift. Our convention will be that the waveplate brings about no phase shift in the ordinary polarization component, while the extraordinary component acquires a phase shift $\Delta\varpi$. In other words, a waveplate with its optical axes oriented at angle θ to the horizontal effects the following transformations:

$$|\theta\rangle \to \mathrm{e}^{\mathrm{i}\Delta\varphi}\,|\theta\rangle\,; \tag{1.4a}$$

$$\left|\frac{\pi}{2}+\theta\right\rangle \to \left|\frac{\pi}{2}+\theta\right\rangle. \tag{1.4b}$$

Exercise 1.24. Find, in the canonical basis, the matrices of the operators associated with a half- and quarter-wave plate with its optical axes oriented at angle α to the horizontal, by following these steps.

a) Write the operator $\hat{A}_{\Delta\varphi}$ associated with the transformation (1.4) in the form of Eq. (A.25).
b) Express each bra and ket in the result of part (a) in the matrix form in the canonical basis and calculate the matrix of the resulting operator.
c) Specialize the result to the half- and quarter-wave plates.

Answer:

$$\hat{A}_{\mathrm{HWP}}(\alpha) = \begin{pmatrix} -\cos 2\alpha & -\sin 2\alpha \\ -\sin 2\alpha & \cos 2\alpha \end{pmatrix}; \tag{1.5a}$$

$$\hat{A}_{\mathrm{QWP}}(\alpha) = \begin{pmatrix} \sin^2\alpha + \mathrm{i}\cos^2\alpha & (\mathrm{i}-1)\sin\alpha\cos\alpha \\ (\mathrm{i}-1)\sin\alpha\cos\alpha & \mathrm{i}\sin^2\alpha + \cos^2\alpha \end{pmatrix}. \tag{1.5b}$$

Exercise 1.25. Using the result of the above exercise, check the following:

a) when applied to a photon linearly polarized at angle θ, a $\lambda/2$ plate with its optic axis oriented at angle α produces a photon linearly polarized at angle $2\alpha-\theta$, in agreement with Fig. C.4;
b) a $\lambda/4$-plate with the optic axis oriented horizontally or vertically interconverts between the circular and $\pm 45^\circ$ polarization photons, in agreement with Ex. C.9.

Exercise 1.26. *Pauli operators*[19] are defined as

$$\hat{\sigma}_x \equiv |H\rangle\langle V| + |V\rangle\langle H|\,; \tag{1.6a}$$

$$\hat{\sigma}_y \equiv -\mathrm{i}\,|H\rangle\langle V| + \mathrm{i}\,|V\rangle\langle H|\,; \tag{1.6b}$$

$$\hat{\sigma}_z \equiv |H\rangle\langle H| - |V\rangle\langle V|\,, \tag{1.6c}$$

or in the matrix notation,

$$\hat{\sigma}_x \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \qquad \hat{\sigma}_y \equiv \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}; \qquad \hat{\sigma}_z \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.7}$$

[19] The meaning of subscripts x, y, and z will be made clear in Chapter 4, when we study quantization of angular momenta.

Propose the implementation of these operators by means of waveplates.
Hint: Find the states onto which the Pauli operators map $|H\rangle$ and $|V\rangle$, then use Ex. 1.24.

Exercise 1.27. The matrix of the *Hadamard operator* $\hat{H}$ in the canonical basis is $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$.

a) Express this operator in the Dirac notation.
b) Onto which states does $\hat{H}$ map $|H\rangle$ and $|V\rangle$?
c) How can one implement this operator using waveplates?

1.8 Projection operators and unnormalized states

We previously postulated that physical quantum states have norm 1. Let us now extend this convention. A norm of a state vector $|a\rangle$ can be less than 1; this means that state $|a\rangle$ exists not with certainty, but with a probability that equals the square of its norm

$$\mathrm{pr}_a = \| |a\rangle \|^2 = \langle a| \, a\rangle \,. \tag{1.8}$$

Such states are called *unnormalized*.

Consider a projective measurement of the state $|\psi\rangle$ in the basis $\{|v_i\rangle\}$. The canonical formulation of the Measurement Postulate says that the measurement transforms $|\psi\rangle$ into one of the $|v_i\rangle$ with probability (1.3). Using the extended convention, we can equivalently state that this measurement transforms $|\psi\rangle$ into a set of unnormalized states $|\psi_i'\rangle = \langle v_i| \, \psi\rangle \, |v_i\rangle$. Each $|\psi_i'\rangle$ is proportional to $|v_i\rangle$, but has a probability of existing that is equal to its squared norm

$$\langle \psi_i'| \, \psi_i'\rangle = |\langle v_i| \, \psi\rangle|^2 \overset{(1.3)}{=} \mathrm{pr}_i. \tag{1.9}$$

We can rewrite this according to

$$|\psi_i'\rangle = \hat{\Pi}_i |\psi\rangle = \langle v_i| \, \psi\rangle \, |v_i\rangle \,, \tag{1.10}$$

where we have introduced the *projection operator*

$$\hat{\Pi}_i = |v_i\rangle\langle v_i| \,. \tag{1.11}$$

For example, a non-destructive measurement of the state $|\psi\rangle = (2|H\rangle + |V\rangle)/\sqrt{5}$ in the canonical basis generates the following unnormalized states:

$$|\psi_H'\rangle = \hat{\Pi}_H |\psi\rangle = |H\rangle \langle H| \, \psi\rangle = 2|H\rangle/\sqrt{5};$$
$$|\psi_V'\rangle = \hat{\Pi}_V |\psi\rangle = |V\rangle \langle V| \, \psi\rangle = |V\rangle/\sqrt{5}.$$

The state $|\psi'_H\rangle$ represents a horizontally polarized photon existing with probability $\mathrm{pr}_H = 4/5$, while the state $|\psi'_V\rangle$ represents a vertically polarized photon existing with probability $\mathrm{pr}_V = 1/5$.

The interpretation of measurements in terms of projection operators is often useful, as we shall see later in this course.

Exercise 1.28. Find the matrix of the projection operator associated with basis state $|v_2\rangle$ in the basis $\{|v_i\rangle\}$ for the Hilbert space of dimension $N = 4$.

1.9 Quantum observables

1.9.1 Observable operators

The Measurement Postulate of quantum physics, as defined in Sec. 1.4, states that a quantum measurement is performed in an orthonormal basis and the measurement result is a random element of that basis. Let us now go one step further and associate with each basis element, $|v_i\rangle$, a real number, v_i. Then, instead of saying "the result of the measurement is state $|v_i\rangle$", we say "the result of the measurement is the value v_i".

For some measurements, this association is natural. For example, a state with a certain position, such as $|x_i\rangle = |x = 3m\rangle$, is naturally associated with a value of the particle's coordinate ($x_i = 3$ m). For other measurements, such as the measurement of a photon polarization, there is no natural connection between basis elements and numbers, but it can be introduced artificially. For example, if we are measuring in the canonical basis, we can associate number 1 with state $|H\rangle$ and -1 with state $|V\rangle$.

The information about the measurement basis and the values associated therewith can be conveniently expressed in the form of the operator

$$\hat{V} = \sum_i v_i \, |v_i\rangle \, \langle v_i| \, . \tag{1.12}$$

This operator is called the *observable operator*, or simply the *observable*. As we know (Sec. A.8), the elements $|v_i\rangle$ of the measurement basis (the observable's *eigenbasis*) are the *eigenstates* or *eigenvectors* of the observable and the corresponding values v_i are its *eigenvalues*. Using Eq. (1.12), one can introduce an observable operator for almost any measurement or measurable quantity: position, momentum, angular momentum, energy, etc. As we shall see in the coming sections, observable operators are of paramount significance in quantum physics.

There is one important exception to this general statement. *Time* is never treated as an operator in quantum physics. There are no eigenstates of time, nor quanta of time. Time is simply a continuous variable.

Exercise 1.29. Find the observables associated with the $\{|H\rangle, |V\rangle\}$, $\{|+\rangle, |-\rangle\}$, and $\{|R\rangle, |L\rangle\}$ bases (i.e., the pieces of measurement apparatus in Fig. 1.2) and

the eigenvalues ± 1 (respectively) in the Dirac notation. Find the matrices of these operators in the $\{|H\rangle, |V\rangle\}$ basis.
Answer: Pauli operators (1.6):

$$|H\rangle\langle H| - |V\rangle\langle V| = \hat{\sigma}_z; \tag{1.13a}$$

$$|+\rangle\langle +| - |-\rangle\langle -| = \hat{\sigma}_x; \tag{1.13b}$$

$$|R\rangle\langle R| - |L\rangle\langle L| = \hat{\sigma}_y. \tag{1.13c}$$

Now we have seen the two roles of operators in quantum mechanics: transformation of quantum states and the description of pieces of measurement apparatus. A natural question to ask is whether the physical implementations of the same operator in these two roles are in any way similar. The example above shows that it is not the case. The pieces of measurement apparatus implementing the Pauli operator are shown in Fig. 1.2. The Pauli operators as state transformations, on the other hand, have been implemented in Ex. 1.26. We can see that these setups are quite different.

Exercise 1.30. Show that:

a) operators corresponding to physical observables (1.12) are Hermitian;
b) any Hermitian operator can be associated with a physical observable, i.e., can be expressed in the form (1.12) with real eigenvalues and eigenstates that form an orthonormal basis.

Exercise 1.31. Perform the spectral decomposition of the Pauli matrices (1.7) using the methods of linear algebra. Check the consistency of your result with the definition given in Ex. 1.29.

We see that every measurement can be associated with a Hermitian operator, and every Hermitian operator can be associated with a measurement. Furthermore, the observable operator contains, in a compact form, full information about the measurement basis and the associated eigenvalues. If somebody gives us a Hermitian matrix of an observable operator, we can retrieve this information using the spectral decomposition[20].

1.9.2 Mean value and uncertainty of an observable

Suppose we measure an observable $\hat{V} = \sum_i v_i |v_i\rangle \langle v_i|$ in the state $|\psi\rangle$. The result of this measurement is probabilistic: we will observe each value v_i with probability $\mathrm{pr}_i = |\langle v_i|\psi\rangle|^2$. We can treat the measured value of the observable as a random variable (Appendix B) and find out its salient statistical properties: the expectation value and the variance.

Exercise 1.32. Observable $\hat{V}$ is measured in the state $|\psi\rangle$.

[20] An important exception is the case where the matrix has degenerate eigenvalues. In this case, the solution for the eigenbasis is not unique. See Ex. A.68 for an example.

a) Show that the *expectation value* of this measurement is

$$\langle V \rangle = \langle \psi | \hat{V} | \psi \rangle . \tag{1.14}$$

The expression on the right-hand side of the above equation is also called the *quantum mean value* of the observable $\hat{V}$ in the state $|\psi\rangle$.

b) Show that the *variance* of the value of $\hat{V}$ is

$$\langle \Delta V^2 \rangle = \left\langle \psi \middle| \left(\hat{V} - \langle \psi | \hat{V} | \psi \rangle \right)^2 \middle| \psi \right\rangle , \tag{1.15}$$

and that this variance can be calculated according to

$$\langle \Delta V^2 \rangle = \langle \psi | \hat{V}^2 | \psi \rangle - \langle \psi | \hat{V} | \psi \rangle^2 . \tag{1.16}$$

As in probability theory, the *uncertainty* of a quantum variable is the square root of its variance.

The strange notion of the operator observable introduced in the previous section now turns out to be quite useful. Not only does it carry complete information about the measurement, but it also provides an easy way to calculate the statistical properties of this measurement when applied to a given state. Let us do a simple example.

Exercise 1.33.§ Calculate the mean, variance, and uncertainty of the observable $\hat{\sigma}_z$ in the state $|+\rangle$.

Answer: $\langle \sigma_z \rangle = 0$; $\langle \Delta \sigma_z^2 \rangle = \langle \sigma_z^2 \rangle = 1$; $\sqrt{\langle \Delta \sigma_z^2 \rangle} = 1$.

To interpret the above result, recall that the observable $\hat{\sigma}_z$ can be measured using the setup in Fig. 1.2(a). The observable takes on the value of $+1$ if the photon is transmitted (projected onto the horizontal polarization state) and -1 if the photon is reflected (projected onto the vertical polarization state). The diagonally polarized photon has an equal chance of being transmitted or reflected, so the mean value over multiple measurements would be zero. Regarding the variance, in every instance of the measurement, we have a value of either $+1$ or -1, so the mean square deviation from zero must be 1.

This is a good example of a transition between classical and quantum measurements. Quantum measurements are probabilistic: in the present case, each photon will be randomly transmitted or reflected. In classical physics, on the other hand, everything is deterministic: if we send a 45°-polarized classical wave onto a PBS, it will split exactly equally, without any uncertainty. The correspondence principle demands that quantum behavior become classical in the macroscopic limit. We trace this transition from quantum to classical behavior in the following exercise.

Exercise 1.34. A set of N +45°-polarized photons are sent onto a PBS. Calculate the mean and uncertainty of the difference N_- between the numbers of transmitted and reflected photons.

Hint: use Ex. B.5

Answer: the expectation value is zero, the uncertainty is $\sqrt{\langle \Delta N_-^2 \rangle} = \sqrt{N}$.

This may appear strange at first: as our experiment becomes more macroscopic, the uncertainty increases rather than decreases. How is this consistent with classical physics? The answer is that what matters is not the absolute uncertainty, but the relative one, i.e., $\sqrt{\langle \Delta N_-^2 \rangle}/N = 1/\sqrt{N}$. The higher N, the higher the relative precision of photometry in the two channels required to discover quantum fluctuations.

For example, if $N = 10^4$, the statistical deviation is $\sqrt{N} = 100$, so the relative uncertainty is $1/100$. But if $N = 10^6$, this uncertainty becomes ten times smaller, $1/1000$. Now remember that the photon energy is very small ($\sim 4 \times 10^{-19}$ joule for the visible range), so any experiment involving a macroscopically significant amount of light — even on a scale of nanojoules — will contain an enormous number of photons. The relative difference between the transmitted and reflected energies is minuscule, and would require photometers of extraordinary precision to register.

1.9.3 The uncertainty principle

Exercise 1.35. Show that an observable $\hat{V}$ in a certain quantum state $|\psi\rangle$ has zero uncertainty if and only if $|\psi\rangle$ is an eigenstate of the observable (i.e., $\hat{V}|\psi\rangle = v|\psi\rangle$).

Exercise 1.36. Consider two Hermitian operators $\hat{A}$ and $\hat{B}$. Show that they are simultaneously diagonalizable (become diagonal in the same orthonormal basis) if and only if[21] $[\hat{A}, \hat{B}] = 0$.
Hint: The proof is simpler if you assume that one of the operators has no degenerate eigenvalues.

The last exercise shows that any two commuting observables can be measured simultaneously. That is, one can construct a piece of apparatus that performs a measurement in an orthonormal basis that can be associated at the same time with both these observables.

Commuting observables are "compatible": there exists an eigenbasis of $\hat{A}$ such that, when a system is prepared in one of its elements $|v_i\rangle$, it will remain in this state when observable $\hat{B}$ is measured and the measurement result will be certain, namely, $|v_i\rangle$[22]. If, on the other hand, $\hat{A}$ and $\hat{B}$ don't commute, a system prepared in an eigenstate of the observable $\hat{A}$ can give a random result if $\hat{B}$ is measured[23]. The

[21] See Sec. A.9 to learn about commutators.

[22] This does not mean, though, that the measurement of *any* eigenstate of the observable $\hat{A}$ will produce a certain result when $\hat{B}$ is measured. If $\hat{A}$ has degenerate eigenvalues, its eigenbasis is not unique (see Sec. A.8), so not every eigenvector of $\hat{A}$ is guaranteed also to be an eigenvector of $\hat{B}$. For example, if $\hat{A} = \hat{\mathbf{1}}$ and $\hat{B} = \hat{\sigma}_z$, the state $|+\rangle$ is an eigenstate of $\hat{A}$ but not of $\hat{B}$, so the observable $\hat{B}$ will exhibit uncertainty when measured in this state.

[23] Even if $\hat{A}$ and $\hat{B}$ don't commute, this does not mean that measuring observable $\hat{B}$ in an eigenstate of $\hat{A}$ will *always* give a random result. For example, suppose the eigenvectors of $\hat{A}$ in a three-dimensional Hilbert space are $|v_1\rangle$, $|v_2\rangle$, and $|v_3\rangle$, and the eigenvectors of $\hat{B}$ are $|v_1\rangle$, $|v_2'\rangle$, and $|v_3'\rangle$ (all eigenvalues are nondegenerate). These sets are different, so $\hat{A}$ and $\hat{B}$ do not commute. However,

degree of this randomness is quantified by the Heisenberg uncertainty principle, which we study next.

Exercise 1.37. Show that, for any Hermitian operators $\hat{A}$ and $\hat{B}$,

$$\left\langle \{\hat{A},\hat{B}\} \right\rangle = 2\,\mathrm{Re}\left\langle \hat{A}\hat{B} \right\rangle \tag{1.17}$$

$$\left\langle [\hat{A},\hat{B}] \right\rangle = 2\mathrm{i}\,\mathrm{Im}\left\langle \hat{A}\hat{B} \right\rangle; \tag{1.18}$$

$$\left|\left\langle [\hat{A},\hat{B}] \right\rangle\right|^2 \leq 4\left|\left\langle \hat{A}\hat{B} \right\rangle\right|^2, \tag{1.19}$$

where the quantum mean value is calculated in an arbitrary state $|\psi\rangle$.

Exercise 1.38. Show that, for any two Hermitian operators $\hat{A}$, $\hat{B}$, and any state $|\psi\rangle$,

$$\left\langle \hat{A}^2 \right\rangle \left\langle \hat{B}^2 \right\rangle \geq \left|\left\langle \hat{A}\hat{B} \right\rangle\right|^2. \tag{1.20}$$

Hint: Let $|a\rangle = \hat{A}\,|\psi\rangle$ and $|b\rangle = \hat{B}\,|\psi\rangle$ and apply the Cauchy-Schwarz inequality.

Exercise 1.39. Prove the *Heisenberg uncertainty principle*: For Hermitian $\hat{A}$, $\hat{B}$, and any state $|\psi\rangle$

$$\left\langle \Delta\hat{A}^2 \right\rangle \left\langle \Delta\hat{B}^2 \right\rangle \geq \frac{1}{4}\left|\left\langle [\hat{A},\hat{B}] \right\rangle\right|^2. \tag{1.21}$$

assuming for simplicity that

$$\langle A \rangle = \langle B \rangle = 0. \tag{1.22}$$

Exercise 1.40. Redo the proof without assuming Eq. (1.22). Would the uncertainty principle (1.21) remain valid if its right-hand side were $\frac{1}{4}\left|\left\langle \{\hat{A},\hat{B}\} \right\rangle\right|^2$ or $\left|\left\langle \hat{A}\hat{B} \right\rangle\right|^2$?

Exercise 1.41. Show that, if $[\hat{A},\hat{B}] = \varepsilon \cdot \hat{\mathbf{1}}$, then the product of the variances of observables $\hat{A}$ and $\hat{B}$ is independent of $|\psi\rangle$:

$$\left\langle \Delta\hat{A}^2 \right\rangle \left\langle \Delta\hat{B}^2 \right\rangle \geq \frac{|\varepsilon|^2}{4}. \tag{1.23}$$

Exercise 1.42. For $\hat{A} = \hat{\sigma}_x$ and $\hat{B} = \hat{\sigma}_y$:

a) find $\left\langle \psi \right|\hat{A}\left|\psi\right\rangle$, $\left\langle \psi \right|\Delta\hat{A}^2\left|\psi\right\rangle$, $\left\langle \psi \right|\hat{B}\left|\psi\right\rangle$, $\left\langle \psi \right|\Delta\hat{B}^2\left|\psi\right\rangle$, and $\left\langle \psi \right|[\hat{A},\hat{B}]\left|\psi\right\rangle$ for $|\psi\rangle = |H\rangle$;
b) check that the uncertainty principle holds for $\hat{A}$, $\hat{B}$ and $|\psi\rangle = |H\rangle$;
c) give an example of a state $|\psi\rangle$ for which the uncertainty product of observables $\hat{A}$ and $\hat{B}$ is zero.

The Heisenberg uncertainty principle is one of the most important tenets of quantum physics and one of the primary signatures that distinguishes it from classical. It was also one of the most controversial ideas at the time when quantum mechanics was being developed. Similarly to the Measurement Postulate, the uncertainty

---

they have one common eigenstate $|v_1\rangle$, and the system prepared in this state will yield a certain result when either of the two observables are measured.

principle appeared to be in direct contradiction with the deterministic picture of the world accepted by classical physics. According to the latter, any uncertainty one may have in a measurement is a consequence of imperfect measurement apparatus, and can be indefinitely reduced by improving that apparatus. In the framework of quantum mechanics, this is not the case: if one builds a piece of apparatus that is precise for measuring one observable in a particular state of the system, this apparatus is bound to perform poorly when the other observable is measured, no matter how good it is.

Particularly interesting is the case of Ex. 1.41. If the commutator of the two observables is proportional to the identity operator, their uncertainty product has a lower bound for *all* states. An example of such a pair is position and momentum, which we will be studying in Chapter 3. Their commutator equals $i\hbar$, and hence the uncertainty product for any state cannot fall below $\sqrt{\hbar^2/4} = \hbar/2$.

1.10 Quantum evolution

Our goal for this section is to find out how quantum states evolve: given the initial state $|\psi(0)\rangle$ of a physical system, we need to determine its state $|\psi(t)\rangle$ at an arbitrary moment in time. In classical physics, the complete set of equations of motion can be obtained from the *Hamiltonian* (full energy) of the system. That is, the entire information about the time-dependent behavior of the system, for any initial state, is contained in that Hamiltonian. As we shall see, the same is true for quantum physics.

It is not possible to derive the rules for the quantum evolution from the postulates we have studied so far. Therefore we will follow the same tactic as when working out the Measurement Postulate. We will first develop a speculative physical argument for the evolution of a specific physical system, the photon. Then we will generalize it to other systems.

Let us have another look at Eq. (1.2). The evolution of the photon's state consists in the overall phase factor $e^{-i\omega t}$:

$$|\psi(t)\rangle = |\psi(0)\rangle e^{-i\omega t}. \tag{1.24}$$

We have neglected it so far because, as we argued, it has no effect on the physical properties of the state. But let us consider it more closely now.

Recalling that the photon energy is $E = \hbar\omega$, we can rewrite Eq. (1.24) as

$$|\psi_E(t)\rangle = |\psi_E(0)\rangle e^{-\frac{i}{\hbar}Et}, \tag{1.25}$$

where the subscript E reminds us that we are dealing with a state of a certain energy (in this case, a photon of a certain frequency).

The next step is to invoke the de Broglie hypothesis, according to which not only photons, but all freely moving particles can be associated with waves whose space-time behavior is described by the factor $e^{i\vec{k}\cdot\vec{r}}e^{-\frac{i}{\hbar}Et}$, where $k = p/\hbar$. We will

discuss this hypothesis more extensively in Chapter 3; for now, let us observe that the time dependence of the de Broglie wave is the same as in Eq. (1.25). This leads us to conclude that Eq. (1.25) is valid not only for photons, but for all freely moving quantum particles. We will postulate that such behavior is even more universal, i.e., that it is valid for *all* non-relativistic quantum objects in the universe, as long as they are in a state with a certain energy — that is, in an eigenstate of the energy (Hamiltonian) operator.

Because this operator corresponds to a physical observable, it is Hermitian, and hence permits the spectral decomposition

$$\hat{H} = \sum_j E_j \left|E_j\right\rangle\left\langle E_j\right|, \tag{1.26}$$

with its eigenstates $\{\left|E_j\right\rangle\}$ forming a basis into which any arbitrary state can be decomposed:

$$\left|\psi(0)\right\rangle = \sum_j \psi_j \left|E_j\right\rangle. \tag{1.27}$$

Each component of this decomposition evolves in time according to Eq. (1.25). Because the evolution is linear, we can write

$$\left|\psi(t)\right\rangle = \sum_j \psi_j \left|E_j\right\rangle \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}E_j t}. \tag{1.28}$$

We postulate that this equation applies universally to the evolution of all quantum states.

Exercise 1.43. Let the initial state of some system be a superposition of two energy eigenstates $\left|\psi(0)\right\rangle = (\left|E_1\right\rangle + \left|E_2\right\rangle)/\sqrt{2}$. Find the lowest $t > 0$ for which the state $\left|\psi(t)\right\rangle$ becomes physically equivalent to $(\left|E_1\right\rangle - \left|E_2\right\rangle)/\sqrt{2}$.

We see that, while the quantum evolution corresponds to an unphysical phase factor for energy eigenstates (such as in the case of polarization states of a photon of a certain frequency), other states do change their physical constitution with time.

Because energy eigenstates do not physically evolve, they are called *stationary*. As an example of stationary states, we can also think of an atom in the framework of the Bohr model. According to this model, if the electron is in an "orbital" corresponding to a certain energy value, it can stay in it for a long time period.

Equation (1.28) can be used to calculate the evolution of a quantum state directly. However, it is sometimes more practical to present the evolution in a more compact form, such as the *evolution operator* that maps any initial state onto its evolved version:

$$\left|\psi(t)\right\rangle = \hat{U}(t)\left|\psi(0)\right\rangle. \tag{1.29}$$

Let us obtain the evolution operator explicitly.

Exercise 1.44. Using Eqs. (1.27) and (1.28),

a) obtain the matrix of the evolution operator in the eigenbasis of the Hamiltonian;

b) show that[24]

$$\hat{U}(t) = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\hat{H}t}. \tag{1.30}$$

Check that this operator is unitary.

The unitarity of the evolution operator is not surprising. This operator must map a physical state onto another physical state, which means that it must preserve the norm.

Exercise 1.45[§] Check that the waveplate transformation operators (1.5) are unitary.

As we know (Ex. A.82), all unitary operators are invertible and the inverse of a unitary operator is also a unitary operator. This has quite a deep consequence. If we know the evolution operator and the state resulting from this evolution, we can reproduce the initial state by applying the inverse evolution operator to the final state. No information is ever lost during the evolution of an isolated quantum system. In the language of statistical physics, this means that the entropy of a physical system does not increase during its evolution.

Equation (1.30) tells us explicitly how to implement this inversion. Replacing $\hat{H}$ by $-\hat{H}$ in Eq. (1.30) is equivalent to replacing t by $-t$, i.e., it makes the evolution go back in time, eventually bringing the system back to its original state. This phenomenon, known as the *time reversibility* of quantum mechanics, has many interesting applications, for example, the spin echo (Sec. 4.7.4).

Exercise 1.46. For any state $|\psi(t)\rangle$, show that

$$\frac{d|\psi(t)\rangle}{dt} = -\frac{\mathrm{i}}{\hbar}\hat{H}|\psi(t)\rangle. \tag{1.31}$$

Equation (1.31) is the *Schrödinger equation*. This is yet another way to describe the law for the evolution of quantum systems, and historically, it was in fact the first.

Our next goal is to practice finding the time-dependent evolution of quantum states. The physical system we have used so far, the photon polarization, is not very convenient for this purpose because the energy of the photon equals $\hbar\omega$ independently of the polarization. However, in order to practice (before we learn about other physical systems with a diverse energy spectrum), let us suppose that, under certain conditions, the photon energy can become polarization-dependent, and study the way the polarization evolves.

Suppose we are given the initial state $|\psi(0)\rangle$ of the system and its Hamiltonian $\hat{H}$, and need to predict its state $|\psi(t)\rangle$ at any moment in time. To this end, we can use three methods:

I. decompose $|\psi(0)\rangle$ into the energy eigenbasis according to Eq. (1.27) and then apply the simple evolution Eq. (1.28) to each basis element in order to find $|\psi(t)\rangle$;

[24] See Sec. A.11 about functions of operators.

II. calculate the evolution operator from (1.30), using the tricks learned in Sec. A.11, and then apply this operator to the initial state according to Eq. (1.29);
III. solve the Cauchy problem consisting of the Schrödinger differential equation (1.31) and the initial state $|\psi(0)\rangle$. In this approach, the Schrödinger equation can be written in the matrix form

$$\begin{pmatrix} \dot{\psi}_H(t) \\ \dot{\psi}_V(t) \end{pmatrix} = -\frac{i}{\hbar} \begin{pmatrix} H_{HH} & H_{HV} \\ H_{VH} & H_{VV} \end{pmatrix} \begin{pmatrix} \psi_H(t) \\ \psi_V(t) \end{pmatrix}, \tag{1.32}$$

and solved as a system of two differential equations for a pair of functions $(\psi_H(t), \psi_V(t))$.

Exercise 1.47. Write the Schrödinger equation for the following Hamiltonians:

a) $\hat{H} = \hbar\omega\hat{\sigma}_z$;
b) $\hat{H} = \hbar\omega\hat{\sigma}_x$.

For each case, find the polarization state of the photon at time t if its initial state is either $|\psi(0)\rangle = |H\rangle$ or $|\psi(0)\rangle = |{+45^\circ}\rangle$, using each of the three methods listed above. Express your answer in the canonical basis.

Exercise 1.48. Find the values of t in Ex. 1.47 for which the action of the evolution operator is equivalent to that of half- and quarter-wave plates at angles 0 and 45° for parts (a) and (b), respectively.

We see that the evolution of photons studied in Ex. 1.47 is equivalent to that occurring in birefringent materials. However, the physics is not entirely the same. In birefringent materials, the eigenstates of the evolution operator incur different phases because of the different refraction indices for the ordinary and extraordinary polarizations (Appendix C). In Hamiltonian evolution, on the other hand, the phase shift is due to the different energies of the energy eigenstates.

1.11 Problems

Problem 1.1. Find the commutator $[(\hat{\sigma}_x + \hat{\sigma}_y)^2, \hat{\sigma}_z]$.

Problem 1.2. Two states are decomposed in the circular basis according to

$$|\psi\rangle = \frac{2|R\rangle + \mathrm{i}|L\rangle}{\sqrt{5}}, \quad |\phi\rangle = \frac{\mathrm{i}|R\rangle + 2|L\rangle}{\sqrt{5}}. \tag{1.33}$$

a) Show that these states form an orthonormal basis.
b) Find the decompositions of these states in the canonical basis using two methods:

- by expressing $|R\rangle$ and $|L\rangle$ in the canonical basis and substituting into Eq. (1.33);

- by finding the matrices of $|\psi\rangle$, $|\phi\rangle$, $|H\rangle$, and $|V\rangle$ in the circular basis and using the inner product.

c) Check that the states $|\psi\rangle$ and $|\phi\rangle$ form an orthonormal set using the inner product in the canonical basis.
d) Decompose the states $|H\rangle, |V\rangle, |R\rangle, |L\rangle, (|H\rangle + 2\mathrm{i}|V\rangle)/\sqrt{5}$ in the basis $\{|\psi\rangle, |\phi\rangle\}$. Write your answer in both the Dirac and the matrix notations.
e) States $|H\rangle, |V\rangle, |R\rangle, |L\rangle, (|H\rangle + 2\mathrm{i}|V\rangle)/\sqrt{5}$ are measured in the basis $\{|\psi\rangle, |\phi\rangle\}$. What are the probabilities of the outcomes?

Problem 1.3. Repeat Ex. 1.12 for a photon that is in a random statistical mixture described by the following ensemble:

a) either $|+\rangle$ with probability $1/2$ or $|-\rangle$ with probability $1/2$;
b) either $|R\rangle$ with probability $1/2$ or $|L\rangle$ with probability $1/2$.

Problem 1.4. Consider the modified BB84 protocol in which Alice sends and Bob analyzes the photon in a polarization basis that is randomly chosen, with the same probability for each choice, among the following three: $(0^\circ, 90^\circ)$, $(30^\circ, 120^\circ)$, $(60^\circ, 150^\circ)$. Find the bit error rate that Alice and Bob will see in the event of a straightforward "intercept-resend" attack, i.e., if Eve intercepts the photon, measures it in one of the above three bases (randomly chosen with equal probabilities), and resends whatever she detects. There are no losses, all equipment is perfect.

Problem 1.5. Consider an operator $\hat{A}$ that performs the following transformation.

$$|H\rangle \to \frac{2|H\rangle + \mathrm{i}|V\rangle}{\sqrt{5}}; \tag{1.34}$$

$$|+\rangle \to \frac{2+\mathrm{i}}{\sqrt{5}}|+\rangle. \tag{1.35}$$

a) How is the vertical polarization state mapped by $\hat{A}$?[25]
b) Write the matrix of $\hat{A}$ in the canonical basis.
c) Express $\hat{A}$ in the Dirac notation in terms of outer products of states $|H\rangle$ and $|V\rangle$;
d) Using the fact that, for any linear operator, $\hat{A}(\lambda|a\rangle + \mu|b\rangle) = \lambda\hat{A}|a\rangle + \mu\hat{A}|b\rangle$, determine how $\hat{A}$ acts on the circular polarization states.
e) Using the previous result, find the matrix of $\hat{A}$ in the circular polarization basis;
f) Find the matrix of $\hat{A}$ in the canonical basis from its matrix in the circular basis, using the resolution of the identity. Is your result consistent with that of part (b)?
g) Is $\hat{A}$ Hermitian? If not, what is its adjoint?

Problem 1.6. Solve Ex. 1.24 using an alternative method.

[25] In this case, the overall phase on the right-hand side of Eq. (1.35) does matter. This is because we are interested, not only in the transformation of the state $|+\rangle$ itself, but in the whole linear operation defined by this transformation. To see the effect of the overall phase, you may want to try solving part (a) using $|+\rangle \to |+\rangle$ instead of Eq. (1.35).

a) Write the waveplate operator matrix in the basis $\{|\alpha\rangle, |90^\circ + \alpha\rangle\}$.
b) Convert this matrix to the canonical basis using the resolution of the identity (Sec. A.6).

Problem 1.7. Using Eqs. 1.5, show that $\hat{A}^2_{\mathrm{QWP}}(\alpha) = \hat{A}_{\mathrm{HWP}}(\alpha)$, i.e., two quarter-wave plates with parallel optic axes, when placed together, comprise a half-wave plate.

Problem 1.8. Using matrix multiplication, show that a quarter-wave plate oriented at any angle, when applied to a circular polarization state, generates a linear polarization state.

Problem 1.9. Find the measurement basis associated with apparatus which consists of a

a) half-waveplate,
b) quarter-waveplate

with the optic axis oriented at angle α, followed by a polarizing beam splitter and two photon detectors.

Problem 1.10. Operator $\hat{A}$ has the following matrix in the canonical basis.

$$\hat{A} \simeq \begin{pmatrix} 41 & -12\mathrm{i} \\ 12\mathrm{i} & 34 \end{pmatrix}$$

a) Present this operator in the form $\hat{A} = v_1 |v_1\rangle\langle v_1| + v_2 |v_2\rangle\langle v_2|$, where $\{|v_1\rangle, |v_2\rangle\}$ is an orthonormal basis. Find v_1, v_2, as well as the matrices of $|v_1\rangle$ and $|v_2\rangle$ in the canonical basis.
b) Write the matrices of the outer products $|v_{1,2}\rangle\langle v_{1,2}|$ in the canonical basis and check explicitly that $\hat{A} = v_1 |v_1\rangle\langle v_1| + v_2 |v_2\rangle\langle v_2|$.
c) The observable $\hat{A}$ is measured in the circularly polarized state $|R\rangle$. What are the probabilities of the possible outcomes?
d) Calculate the expectation value of the measurement result

- using the definition of the expectation value from probability theory;
- using the expression for the quantum mean.

Check that the results are the same.
e) Calculate the variance of the observable $\hat{A}$ in the state $|R\rangle$.

Problem 1.11. Consider a piece of apparatus for measuring the photon polarization that has the following properties:

- whenever a linearly polarized photon at angle θ enters the apparatus, it displays "2";
- whenever a linearly polarized photon at angle $\pi/2 + \theta$ enters the apparatus, it displays "3".

a) Find the eigenvalues and the eigenstates of the operator $\hat{A}$ associated with the observable measured by this apparatus.
b) Find the matrices of $\hat{A}$ in its eigenbasis and in the $\{|H\rangle, |V\rangle\}$ basis.
c) Find the probability of each measurement outcome for a linearly polarized photon at angle φ.
d) Find the expectation value and uncertainty of this measurement.

Problem 1.12. Write the uncertainty principle for observables $\hat{\sigma}_x$ and $\hat{A} = |R\rangle\langle R| - 2|L\rangle\langle L|$ measured in the state $|H\rangle$. Check explicitly that it holds.

Problem 1.13. Measurements of the observable $\hat{A}$ in the state $|H\rangle$ yield results 0 and 1, each with probability $1/2$. Measurements of the observable $\hat{B}$ in the state $|H\rangle$ yield the result 2 with probability $3/4$ and result 4 with probability 1/4. It is also known that $[\hat{A}, \hat{B}] = ix\hat{\sigma}_z$. Find the upper bound on the absolute value of x.

Problem 1.14. Find $e^{i\frac{\pi}{4}\left(3|H\rangle\langle H| + \sqrt{3}i|H\rangle\langle V| - \sqrt{3}i|V\rangle\langle H| + |V\rangle\langle V|\right)}$.

Problem 1.15. An atom is described in some basis $\{|v_1\rangle, |v_2\rangle\}$ by the Hamiltonian

$$\hat{H} = \hbar\omega \begin{pmatrix} 1 & 3i \\ -3i & 9 \end{pmatrix}.$$

a) Find the energy eigenstates and eigenvalues.
b) The energy of the atom is measured in the state $|\psi_0\rangle = \frac{1}{\sqrt{2}}(|v_1\rangle + i|v_2\rangle)$. Find the probabilities of detecting each energy eigenvalue, as well as the mean and variance of this measurement.
c) The atom is initially in the state $|v_1\rangle$. Find its state $|\psi(t)\rangle$ at an arbitrary time t. How much time will elapse until the atom is once again in the state $|v_1\rangle$ (up to a phase factor)?

Problem 1.16. Suppose the operator (1.5a) associated with a half-waveplate at angle α corresponds to the evolution under some Hamiltonian for time t_0.

a) Find the matrix of this Hamiltonian in the canonical basis.
b) Check that the evolution for the time $t_0/2$ will give rise to the quarter-waveplate operator (1.5b).
c) For the Hamiltonian found in part (a) and $\alpha = 30°$, solve the differential Schrödinger equation (1.31) for the initial state $|H\rangle$. Is the result for $t = t_0$ consistent with what you would expect from the physics of polarization transformations?

Problem 1.17. A quantum system can be found in one of three orthogonal states $|a\rangle$, $|b\rangle$, $|c\rangle$. These three states form an orthonormal basis. $\hat{A}$ is an operator which cyclically permutes the states, i.e., $\hat{A}|a\rangle = \hbar\omega|b\rangle$, $\hat{A}|b\rangle = \hbar\omega|c\rangle$, $\hat{A}|c\rangle = \hbar\omega|a\rangle$ (where ω is real). The Hamiltonian is $\hat{H} = \hat{A} + \hat{A}^\dagger$.

a) Find the energy eigenvalues and eigenstates of the system.
b) Find the evolution of the system that is initially in state $|c\rangle$.

Problem 1.18. An atom has two energy eigenstates $|v_1\rangle, |v_2\rangle$ with eigenvalues 0 and $3\hbar\omega$, respectively, where $\omega > 0$.

a) Write the matrix of the corresponding Hamiltonian $\hat{H}_0$.
b) At time $t = 0$, a field is turned on which makes the Hamiltonian equal to $\hat{H} = \hat{H}_0 + \hat{V}$ with $\hat{V} = 2i\hbar\omega\,|v_1\rangle\langle v_2| - 2i\hbar\omega\,|v_2\rangle\langle v_1|$. Write the matrix of the new Hamiltonian and the associated evolution operator in the basis $\{|v_1\rangle, |v_2\rangle\}$.
c) At time $t = 0$, the atom is in the state $|v_1\rangle$. Find all values of the time t at which the probability of finding the atom in the state $|v_2\rangle$ is maximized.

Chapter 2
Entanglement

2.1 Tensor product spaces

2.1.1 Tensor product states and entangled states

Consider two physical systems that are separated in space and/or time, but are interacting, or have been interacting in the past. In order to study the states emerging after such an interaction, it is not enough to treat each system separately. They must be treated as a single Hilbert space, which unites the Hilbert spaces associated with the individual systems.

Suppose, for example, that Alice, on Venus, has[1] a horizontally polarized photon, $|H\rangle$, while Bob, on Mars, has a photon in the state $|V\rangle$. Then we say that the joint state of Alice's and Bob's photons is

$$|H\rangle_A \otimes |V\rangle_B \equiv |H\rangle\,|V\rangle \equiv |HV\rangle\,. \tag{2.1}$$

These joint states are called the *tensor product* states[2].

However, the joint Hilbert space contains not only tensor product states. For example, since it contains states $|HV\rangle$ and $|VH\rangle$ and it is a linear space, it must also contain the state $(|HV\rangle - |VH\rangle)/\sqrt{2}$. This is a physical state, because it has norm 1. But it can no longer be interpreted as a tensor product, i.e., a combination of Alice's photon being in one state and Bob's in another. This is a *nonlocal superposition*, or *entangled* state. It is a quantum superposition of two situations: one in which Alice has a horizontal photon while Bob has a vertical one, and vice versa. If they measure

[1] This is a manner of speaking, of course. Photons move at the speed of light, and no one can "have" them for any extended period of time. The notion of Alice or Bob "having" a photon corresponds, as a rule, to an instant in time just before the measurement.

[2] The four equivalent parts of Eq. (2.1) represent alternative notations for tensor product states that we will be using interchangeably. Note that the subscript A or B labelling the Hilbert space is positioned *outside* the ket. When these subscripts are omitted, the first component of a tensor product is always assumed to pertain to Alice and the second to Bob.

A. I. Lvovsky, *Quantum Physics*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-662-56584-1_2

their photon polarizations in the canonical basis, they will always detect orthogonal polarizations.

We see that uniting two Hilbert spaces results in a brand new class of states, which gives rise to new physics, the physics of *nonlocal quantum phenomena*. This is the main subject of the present chapter. Some of these phenomena are not only unthinkable from the point of view of classical physics, but they even seem to contradict the most basic common sense.

Before we study this novel physics, we need to sharpen our pencils and upgrade our theoretical machinery so that it can be applied to tensor product Hilbert spaces. We will carry out our derivations for *bipartite* tensor products (i.e., involving two parties), but they can be straightforwardly extended to systems with three or more parties.

The *tensor product space* $\mathbb{V}_A \otimes \mathbb{V}_B$ of Hilbert spaces $\mathbb{V}_A$ and $\mathbb{V}_B$ is a Hilbert space consisting of elements $|a\rangle \otimes |b\rangle$ (with $|a\rangle \in \mathbb{V}_A$ and $|b\rangle \in \mathbb{V}_B$) and their linear combinations. Operations in the space obey the following rules:

1. Multiplication by a number:

$$\lambda(|a\rangle \otimes |b\rangle) = (\lambda |a\rangle) \otimes |b\rangle = |a\rangle \otimes (\lambda |b\rangle). \tag{2.2}$$

2. Distributivity:

$$(|a_1\rangle + |a_2\rangle) \otimes |b\rangle = |a_1\rangle \otimes |b\rangle + |a_2\rangle \otimes |b\rangle; \tag{2.3a}$$
$$|a\rangle \otimes (|b_1\rangle + |b_2\rangle) = |a\rangle \otimes |b_1\rangle + |a\rangle \otimes |b_2\rangle. \tag{2.3b}$$

3. The inner product of two states $|a\rangle \otimes |b\rangle$ and $|a'\rangle \otimes |b'\rangle$ in $\mathbb{V}_A \otimes \mathbb{V}_B$ is given by

$$\langle ab| a'b'\rangle = \langle a| a'\rangle \langle b| b'\rangle. \tag{2.4}$$

Elements of $\mathbb{V}_A \otimes \mathbb{V}_B$ that can be presented in the form of a tensor product $|a\rangle \otimes |b\rangle$ are said to be *separable*. Others are *entangled*.

Exercise 2.1. For any two vectors $|a\rangle \in \mathbb{V}_A$, $|b\rangle \in \mathbb{V}_B$, show that

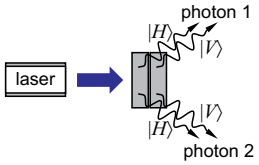
$$|\text{zero}\rangle_{\mathbb{V}_A} \otimes |b\rangle = |a\rangle \otimes |\text{zero}\rangle_{\mathbb{V}_B} = |\text{zero}\rangle_{\mathbb{V}_A \otimes \mathbb{V}_B}.$$

Exercise 2.2. Given orthonormal bases $\{|v_i\rangle\}_{i=1}^N$ and $\{|w_i\rangle\}_{i=1}^M$ in $\mathbb{V}_A$ and $\mathbb{V}_B$, respectively, construct an orthonormal basis in $\mathbb{V}_A \otimes \mathbb{V}_B$. What is the dimension of $\mathbb{V}_A \otimes \mathbb{V}_B$?
Answer: The set of tensor products $\{|v_i\rangle \otimes |w_j\rangle\}$ is an orthonormal basis. The dimension of the tensor product space is the product NM of the dimensions of its components.

For example, the Hilbert space representing the polarizations of two photons is four-dimensional. The canonical orthonormal basis in this space is $\{|HH\rangle, |HV\rangle, |VH\rangle, |VV\rangle\}$.

Box 2.1 How to make an entangled state?

Consider parametric down-conversion (Box 1.6) in a series of two nonlinear crystals, as shown in the figure*. The crystals are constructed in such a way that the first one only produces pairs of horizontally polarized photons $|H\rangle \otimes |H\rangle$, and the second one pairs of vertically polarized photons $|V\rangle \otimes |V\rangle$. The probability of generating a pair is small in both crystals. Then, whenever a pair is present, it can be in either the state $|HH\rangle$ or the state $|VV\rangle$. Moreover, because the distance between the crystals is constant, so is the optical phase φ between these two pairs. So the state of the two photons produced by the crystals is



$$|HH\rangle + e^{i\varphi} |VV\rangle .$$

By choosing the value of φ, one can generate either of the two Bell states $|\Phi^+\rangle$ or $|\Phi^-\rangle$. To convert these states into $|\Psi^+\rangle$ or $|\Psi^-\rangle$, it suffices to place a half-wave plate in one of the emission channels.

*This scheme was proposed and realized, for the first time, in P.G. Kwiat, E. Waks, A.G. White, I. Appelbaum, and P.H. Eberhard, *Ultrabright source of polarization-entangled photons*, Physical Review A **60**, R773(R) (1999).

Exercise 2.3. Find the canonical basis decomposition of the state in which Alice has a 30° polarized photon, and Bob a right circularly polarized photon. Write the matrix representation of that state. Is it separable or entangled?

Exercise 2.4. Find the inner product $\langle \Pi |\, \Omega \rangle$, where

a) $|\Pi\rangle = 5|HH\rangle + 6i|R-\rangle$ and $|\Omega\rangle = 2|+L\rangle + 3|RR\rangle$;
b) $|\Pi\rangle = i(2|H\rangle + i|V\rangle) \otimes |R\rangle$ and $|\Omega\rangle = (2i|H\rangle - 3i|V\rangle) \otimes |+\rangle$.

Exercise 2.5§ Do the sets

a) $\{|++\rangle, |+-\rangle, |-+\rangle, |--\rangle\}$,
b) $\{|RR\rangle, |RL\rangle, |LR\rangle, |LL\rangle\}$,
c) $\{|H-\rangle, |H+\rangle, |V-\rangle, |V+\rangle\}$,
d) $\{|H-\rangle, |H+\rangle, |VR\rangle, |VL\rangle\}$,
e) $\{|H-\rangle, |HH\rangle, |VR\rangle, |VL\rangle\}$

form bases in the two-photon Hilbert space? Are they orthonormal?
Answer: All five sets form bases; all but the last are orthonormal.

Exercise 2.6. Show that the *Bell states*

$$\left|\Psi^+\right\rangle = \frac{1}{\sqrt{2}}(|HV\rangle + |VH\rangle) \tag{2.5a}$$

$$\left|\Psi^-\right\rangle = \frac{1}{\sqrt{2}}(|HV\rangle - |VH\rangle) \tag{2.5b}$$

$$\left|\Phi^+\right\rangle = \frac{1}{\sqrt{2}}(|HH\rangle + |VV\rangle) \tag{2.5c}$$

$$\left|\Phi^-\right\rangle = \frac{1}{\sqrt{2}}(|HH\rangle - |VV\rangle) \tag{2.5d}$$

are entangled.

Exercise 2.7. Show that the four Bell states form an orthonormal basis.

Exercise 2.8. Rewrite the Bell states (2.5) in the diagonal basis.

Exercise 2.9. Let $|\theta\rangle$ be the linear polarization state at angle θ to horizontal. Show that, for any θ, the state $|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|HV\rangle - |VH\rangle)$ can be expressed in the form

$$\left|\Psi^-\right\rangle = \frac{1}{\sqrt{2}}\left(|\theta\rangle \otimes \left|\frac{\pi}{2}+\theta\right\rangle - \left|\frac{\pi}{2}+\theta\right\rangle \otimes |\theta\rangle\right). \tag{2.6}$$

This means that the state $|\Psi^-\rangle$ is *isotropic*, i.e., it remains the same no matter which direction we define as horizontal (as long as it is perpendicular to the direction of propagation of the photons, of course). This property of $|\Psi^-\rangle$ is unique among all the Bell states.

2.1.2 Measurements in tensor product spaces

The Measurement Postulate of quantum physics applies to tensor product states in the usual fashion. The measurement basis can consist of separable as well as entangled states. If the basis is constructed as a tensor product of bases in $\mathbb{V}_A$ and $\mathbb{V}_B$, such as in Ex. 2.2, the measurement would simply consist of Alice and Bob measuring their objects within their individual Hilbert spaces (Fig. 2.1).

Exercise 2.10. For $|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|HV\rangle - |VH\rangle)$, find the probability of detecting the state

a) $|R\rangle \otimes |-30^\circ\rangle$;
b) $\frac{1}{3}(|HV\rangle + 2|VH\rangle + 2|VV\rangle)$.

Assume that the measurement is performed in some orthonormal basis that contains the state we are interested in.

Exercise 2.11. Alice and Bob share the state

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|HV\rangle + \mathrm{e}^{-\mathrm{i}\phi}|VH\rangle).$$

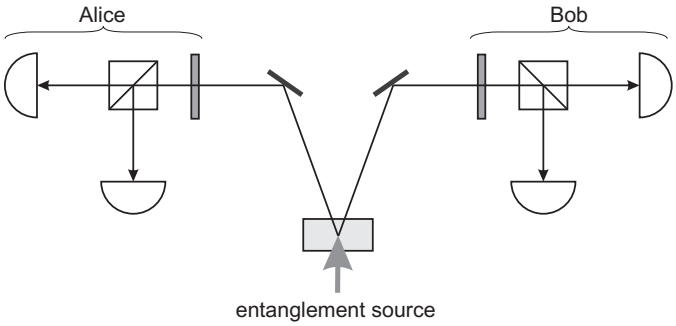


Fig. 2.1 A measurement of a polarization entangled photon pair in a tensor product basis. Alice's and Bob's apparatus each consists of waveplate(s), a polarizing beam splitter and two single-photon detectors.

a) Find the probabilities of all outcomes if Alice and Bob measure $|\Psi\rangle$ in (i) canonical and (ii) diagonal $\{|{+}{+}\rangle, |{+}{-}\rangle, |{-}{+}\rangle, |{-}{-}\rangle\}$ bases.
b) Alice and Bob share a single copy of one of the Bell states, $|\Psi^-\rangle$ or $|\Psi^+\rangle$, but they do not know which one. Can they distinguish them by measuring in the canonical basis? What about the diagonal basis?

An important conclusion we make from this exercise is that, while entangled states can only be created through interaction between the two physical systems, their measurement (e.g., distinguishing between them) does not require any interaction or even projection onto entangled states. In fact, one can perform full quantum tomography of a quantum state in a tensor product Hilbert space by measurements in bases containing only separable states. We will show this rigorously at the end of the book (Ex. 5.78).

Exercise 2.12.* Propose a procedure for performing the measurement in the basis $\{|H-\rangle, |H+\rangle, |VR\rangle, |VL\rangle\}$.
Hint: Assume that Alice and Bob are connected by a classical communication channel.

2.1.3 Tensor products of operators

Now let us extend the notion of the tensor product to operators. This extension is straightforward: in the operator $\hat{A}\otimes\hat{B}$, component $\hat{A}$ acts on Alice's Hilbert space while component $\hat{B}$ acts on Bob's. Here is a formal definition followed by a few exercises.

The *tensor product of operators* $\hat{A}$ on $\mathbb{V}_A$ and $\hat{B}$ on $\mathbb{V}_B$ is defined as a linear operator $\hat{A}\otimes\hat{B}$ on $\mathbb{V}_A\otimes\mathbb{V}_B$ such that, for any vector $|\Psi\rangle = \sum_i \lambda_i |a_i\rangle\otimes|b_i\rangle$,

$$(\hat{A}\otimes\hat{B})\,|\Psi\rangle \equiv \sum_i \lambda_i (\hat{A}\,|a_i\rangle)\otimes(\hat{B}\,|b_i\rangle). \tag{2.7}$$

Exercise 2.13. Express the matrix of the tensor product operator $\hat{C}=\hat{A}\otimes\hat{B}$ in the basis $\{|v_i\rangle\otimes|w_j\rangle\}$ through the matrices of operators $\hat{A}$ and $\hat{B}$ in the respective bases $\{|v_i\rangle\}$ and $\{|w_j\rangle\}$.
Answer: For each matrix element,[3]

$$C_{iji'j'} = \langle v_i w_j |\,\hat{C}\,|\, v_{i'} w_{j'}\rangle = A_{ii'}B_{jj'}. \tag{2.8}$$

Exercise 2.14. Find the expectation value and uncertainty of the operator $\hat{\sigma}_x\otimes\hat{\sigma}_y$ in the state $|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|HV\rangle - |VH\rangle)$.

Exercise 2.15.§ Suppose $|v\rangle$ and $|w\rangle$ are eigenstates of operators $\hat{A}$ and $\hat{B}$, respectively, with corresponding eigenvalues v and w. Show that the state $|v\rangle\otimes|w\rangle$ is an eigenstate of the operator $\hat{A}\otimes\hat{B}$ with eigenvalue vw.

Exercise 2.16. Show that, for operators $\hat{A}_1,\hat{A}_2$ in $\mathbb{V}_A$ and $\hat{B}_1,\hat{B}_2$ in $\mathbb{V}_B$,

$$\hat{A}_1\hat{A}_2\otimes\hat{B}_1\hat{B}_2 = (\hat{A}_1\otimes\hat{B}_1)(\hat{A}_2\otimes\hat{B}_2)$$

Exercise 2.17.§ Show that a tensor product operator cannot make an entangled state from a separable one.

Exercise 2.18. For two outer product operators $\hat{A}=|a_1\rangle\langle a_2|$ and $\hat{B}=|b_1\rangle\langle b_2|$ in $\mathbb{V}_A$ and $\mathbb{V}_B$, respectively, show that

$$\hat{A}\otimes\hat{B} = |a_1 b_1\rangle\langle a_2 b_2|. \tag{2.9}$$

The notion of operator tensor product is nicely illustrated by an important result known as the *no-cloning theorem*[4]. Suppose we have two objects represented by identical Hilbert spaces, $\mathbb{V}_A$ and $\mathbb{V}_B$, and the object represented by $\mathbb{V}_A$ is in some arbitrary quantum state $|a\rangle$. *Quantum cloning* is a hypothetical operation that would create a copy of $|a\rangle$ in $\mathbb{V}_B$ while preserving it in $\mathbb{V}_A$. In other words, it corresponds to an operator on $\mathbb{V}_A\otimes\mathbb{V}_B$ such that, for any $|a\rangle\in\mathbb{V}_A$ and some $|0\rangle\in\mathbb{V}_B$,

$$|a\rangle\otimes|0\rangle \to |a\rangle\otimes|a\rangle \tag{2.10}$$

Exercise 2.19. Show that quantum cloning, as defined above, is impossible.
Hint: use the fact that any physically possible evolution in quantum mechanics is described by a linear operator.

---

[3] We will normally use the intuitive double-index notation for matrices of states and operators in tensor product spaces. That is, each element $|v_i\rangle\otimes|w_j\rangle$ of the tensor product basis is identified by a pair of indices (i,j), such as in Eq. (2.8). This means, for example, that an operator matrix has four, rather than two, indices.

[4] W. Wootters, W. Zurek, *A Single Quantum Cannot be Cloned*, Nature **299**, 802 (1982); D. Dieks, *Communication by EPR devices*, Physics Letters A **92**, 271 (1982).

We define the *adjoint tensor product space* analogously to the direct tensor product space, i.e., for any tensor product state $|a\rangle \otimes |b\rangle$[5]

$$\text{Adjoint}(|a\rangle \otimes |b\rangle) \equiv \text{Adjoint}(|a\rangle) \otimes \text{Adjoint}(|b\rangle) \equiv_A \langle a| \otimes_B \langle b| \equiv \langle ab| \,. \qquad (2.11)$$

Exercise 2.20. Show that, for $\hat{A}$ in $\mathbb{V}_A$ and $\hat{B}$ in $\mathbb{V}_B$: $(A \otimes B)^\dagger = A^\dagger \otimes B^\dagger$

Exercise 2.21. Show that:

a) a tensor product of two Hermitian operators is Hermitian;
b) a tensor product of two unitary operators is unitary.

2.1.4 Local operators

Tensor product operators of the form $\hat{A} \otimes \hat{\mathbf{1}}$ or $\hat{\mathbf{1}} \otimes \hat{B}$ are called *local operators* because they affect only one of the component Hilbert spaces. An example is a waveplate which is placed in the path of Alice's photon and rotates its polarization while leaving Bob's photon untouched. Local operators are often written in a simplified notation: one writes just $\hat{A}$ instead of $\hat{A} \otimes \hat{\mathbf{1}}$ and $\hat{B}$ instead of $\hat{\mathbf{1}} \otimes \hat{B}$.

Exercise 2.22. Show that a local unitary operator cannot change a state's property of being entangled or separable.

Exercise 2.23. Suppose $|a\rangle$ is an eigenstate of operator $\hat{A}$ on Alice's Hilbert space with eigenvalue a. Show that, for any vector $|b\rangle$ in Bob's Hilbert space, $|ab\rangle$ is an eigenstate of the local operator $\hat{A} \otimes \hat{\mathbf{1}}$ with the same eigenvalue.

Exercise 2.24. $\hat{A}$ and $\hat{B}$ are observables in Alice's and Bob's spaces, respectively. A bipartite state $|\Psi\rangle$ is an eigenstate of $\hat{A} \otimes \hat{B}$ with eigenvalue x, but it is not an eigenstate of local operators $\hat{A}$ or $\hat{B}$.

a) Give an example of such a situation.
b) Show that, whenever Alice measures $\hat{A}$ and Bob $\hat{B}$ in the state $|\Psi\rangle$, the product of the values they observe is equal to x.

Hint: use Ex. A.66.

Exercise 2.25. Suppose Alice and Bob share the Bell state $|\Psi^-\rangle$. Alice performs an operation corresponding to one of the three Pauli operators locally on her qubit. Show that

a) $(\hat{\sigma}_z)_A |\Psi^-\rangle = |\Psi^+\rangle$;
b) $(\hat{\sigma}_x)_A |\Psi^-\rangle = -|\Phi^-\rangle$;

[5] The order of symbols inside a bra vector is the same as inside a ket: the first symbol pertains to Alice, the second to Bob. The subscripts A and B indicating the Hilbert spaces, if present, are usually placed to the *left* of bra vectors.

Box 2.2 Holevo bound and quantum dense coding

Suppose Alice and Bob are connected by a communication channel (for example, an optical fiber). Alice wishes to send an n-bit classical message to Bob, encoding the information in a set of quantum particles, each of which is carrying a qubit*. Can she achieve this goal using fewer than n quantum particles?

A simple argument shows that the answer is negative. Indeed, n qubits correspond to a 2^n-dimensional quantum system (Ex. 2.2). No matter how Alice encodes, Bob's measurement on this system can produce no more than 2^n possible outcomes, so the total number of different messages that can be encoded in n qubits is 2^n. The capacity of n bits of classical information is exactly the same. This restriction is a case of the so-called *Holevo bound* in quantum information science.

However, if Alice and Bob have prearranged shared entangled qubits, the Holevo bound can be overcome via the protocol known as *quantum dense coding*. Suppose Alice wishes to send 2 bits of classical information to Bob. The protocol runs as follows:

1. Alice and Bob prearrange a shared state $|\Psi^-\rangle$ of two qubits (for example, photons).
2. Depending on the value of her 2 bits, Alice performs the operation $\hat{\mathbf{1}}$, $\hat{\sigma}_x$, $\hat{\sigma}_y$, or $\hat{\sigma}_z$ on her qubit, thereby transforming the shared entangled state into one of the four Bell states as in Ex. 2.25. This can be realized using waveplates (see Ex. 1.26).
3. Alice sends her qubit to Bob.
4. Bob now possesses two qubits. He measures them in the Bell basis and recovers the value of Alice's two classical bits.

In this way, Alice is able to transfer two bits of classical information by sending only one qubit.

*A reminder: the qubit is any two-dimensional Hilbert space. Photon polarization is an example of a qubit.

c) $(\hat{\sigma}_y)_A |\Psi^-\rangle = \mathrm{i}|\Phi^+\rangle$.

The above result has an important application in a quantum communication protocol called *quantum dense coding* (Box 2.2).

Exercise 2.26. Suppose the Hamiltonian in $\mathbb{V}_A \otimes \mathbb{V}_B$ is given by a sum

$$\hat{H} = \hat{H}_A + \hat{H}_B$$

of Hamiltonians that are local operators in their component spaces. Show that:

a) if the initial state in $\mathbb{V}_A \otimes \mathbb{V}_B$ is a tensor product

$$|\psi(0)\rangle = |\psi_A(0)\rangle \otimes |\psi_B(0)\rangle$$

then the Schrödinger evolution of that state remains a tensor product

$$|\psi(t)\rangle = |\psi_A(t)\rangle \otimes |\psi_B(t)\rangle,$$

where each $|\psi_{A,B}(t)\rangle$ is the solution of the Schrödinger equation for the corresponding Hamiltonian $\hat{H}_{A,B}$;

b) if some $|\psi_A\rangle$ and $|\psi_B\rangle$ are eigenstates of their respective Hamiltonians with energies E_A and E_B, respectively, then the state $|\Psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$ in $\mathbb{V}_A \otimes \mathbb{V}_B$ is an eigenstate of the full Hamiltonian $\hat{H}$ with energy $E = E_A + E_B$;

c)* any eigenstate of the Hamiltonian corresponding to energy E can be written as a linear combination of products of the form $|\psi_A\rangle \otimes |\psi_B\rangle$, where $|\psi_{A,B}\rangle$ are the Hamiltonian eigenstates for the individual Hilbert spaces,

$$\hat{H}_{A,B}\,|\psi_{A,B}\rangle = E_{A,B}\,|\psi_{A,B}\rangle\,,$$

with $E = E_A + E_B$.

2.2 Local measurements of entangled states

As we have seen in the last exercise, extension of the Measurement Postulate to bipartite systems is straightforward if the two observers perform measurements on their respective Hilbert spaces simultaneously. However, since the two observers are independent, it may be that only one of them (e.g., Alice) performs the measurement while the other one (Bob) does not. We call this a *local measurement*.

2.2.1 Remote state preparation

Suppose Alice measures state $|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|HV\rangle - |VH\rangle)$ in the canonical basis. Because $|\Psi^-\rangle$ contains states $|HV\rangle$ and $|VH\rangle$ with amplitudes $\pm 1/\sqrt{2}$, Alice will be equally likely ($\mathrm{pr}_H = \mathrm{pr}_V = 1/2$) to observe either the horizontal or the vertical polarization. If she observes a horizontally polarized photon, we can state with certainty that Bob's photon is vertically polarized, so its state becomes $|V\rangle$, and vice versa.

This correlation by itself is not so surprising. For example, even in the classical world we can play a game in which we give Alice one shoe from a pair, and Bob the other. Each shoe is packed in an opaque box, so their color cannot be seen. Then Alice flies to Venus and Bob flies to Mars, where they open their boxes. Suppose Alice discovers that she has got the left shoe. Then she also instantly learns that Bob's shoe is right, even though he is millions of miles away.

But properties of quantum superpositions extend beyond this simple picture. In addition to polarization correlations, there is a certain *phase relation*, signified by the negative sign between terms $|HV\rangle$ and $|VH\rangle$. In this way, state $(|HV\rangle - |VH\rangle)/\sqrt{2}$ is distinctly different from, e.g. $(|HV\rangle + |VH\rangle)/\sqrt{2}$, even though both exhibit similar correlations when measured in the canonical basis. To see the implications of this phase relation, try solving the following problem.

Exercise 2.27. Suppose Alice and Bob share the state $|\Psi^-\rangle$. Alice measures her portion of the state in the basis $\{|\theta\rangle\,, \left|\tfrac{\pi}{2} + \theta\right\rangle\}$. Show that

a) if Alice detects $|\theta\rangle$, Bob's state becomes $|\pi/2+\theta\rangle$;
b) if Alice detects $|\pi/2+\theta\rangle$, Bob's state becomes $|\theta\rangle$;
c) each of these results happens with probability $1/2$.

Hint: Use the isotropic property of state $|\Psi^-\rangle$ (Ex. 2.9)

The above result is truly remarkable. By choosing the tilt angle θ of the measurement basis, Alice can *remotely prepare* an arbitrary linear polarization state (with a $\pm 90^\circ$ ambiguity) at Bob's location. This happens in spite of the fact that Alice and Bob can be millions of miles away from each other, and have no opportunity to interact. Furthermore, this happens instantly, i.e., faster than the speed of light!

This appears to be in outrageous contradiction with special relativity, and even with the principle of *causality*, which governs all the physics we know, and follows from the most basic common sense. How can it be possible to change something instantly when it is located far away, without any possibility of interacting with that location?

Perhaps the very first question a diligent physics student would ask is whether this conclusion has been verified experimentally. The answer is affirmative. To perform the experiment, one prepares the state $|\Psi^-\rangle$ and makes Alice's measurement many times, each time in the same basis. Every time Alice detects, say, $|\theta\rangle$, Bob makes measurements on this photon. From the measurement statistics, he can reconstruct the state using the technique of quantum state tomography (see Ex. 1.15) with arbitrarily high precision.

Over the last quarter of a century, physicists have studied different versions of the remote state preparation effect. Some of the setups have been arranged so that Alice's and Bob's laboratories were located kilometers away from each other, and their measurements were ensured to occur within a space-like interval, in order to exclude even the theoretical possibility for Alice to affect Bob's state through any interaction known in nature. All these experiments have unequivocally confirmed the validity of the quantum mechanical predictions.

So how can we reconcile these findings with causality? To answer this question, let us first give a formal description of local measurement.

2.2.2 Partial inner product

Suppose Alice and Bob share some entangled state and Alice performs a local measurement on her part thereof in some basis. What are the probabilities of the possible outcomes and what state will be remotely prepared at Bob's location in the case of each outcome? Before answering this question in all its generality, let us look at an example. Let the shared state be

$$|\Psi\rangle = \frac{1}{3}(|HH\rangle - 2|HV\rangle + 2|VV\rangle) \tag{2.12}$$

and suppose that Alice's measurement basis is diagonal.

Exercise 2.28. Rewrite the state (2.12), expressing the state vectors corresponding to Alice's photon in the diagonal basis.
Answer:

$$|\Psi\rangle = \frac{1}{\sqrt{18}}[|+\rangle \otimes |b_+\rangle + \sqrt{17}\,|-\rangle \otimes |b_-\rangle], \tag{2.13}$$

where

$$|b_+\rangle = |H\rangle\,,\ |b_-\rangle = \frac{1}{\sqrt{17}}(|H\rangle - 4\,|V\rangle)$$

are normalized vectors in Bob's Hilbert space.

Because the vectors $|+\rangle$ and $|-\rangle$ are orthogonal, so are $|+\rangle \otimes |b_+\rangle$ and $|-\rangle \otimes |b_-\rangle$, as in Eq. (2.4). This means that we can construct an orthonormal basis in $\mathbb{V}_A \otimes \mathbb{V}_B$ containing these states as elements. If we measure $|\Psi\rangle$ in this basis, we will observe $|+\rangle \otimes |b_+\rangle$ with the probability $\frac{1}{18}$ and $|-\rangle \otimes |b_-\rangle$ with the probability $\frac{17}{18}$. But this in turn means that, if *only Alice* performs her measurement on her photon, she will detect the state $|+\rangle$ with the probability $\frac{1}{18}$ and $|-\rangle$ with the probability $\frac{17}{18}$. This is because, if Alice observes $|+\rangle$, the state of Bob's photon will *certainly* become $|b_+\rangle$, and if Alice observes $|-\rangle$, it will become $|b_-\rangle$.

We see that, in order to answer the question posed in the beginning of this subsection, it suffices to rewrite the initial entangled state as a linear combination of tensor products, such that Alice's component in each of them is an element of her measurement basis. Let us now reproduce the same argument in a more general form.

Suppose the initial state is

$$|\Psi\rangle = \sum_{ij} \Psi_{ij}\,|v_i\rangle \otimes \left|w_j\right\rangle, \tag{2.14}$$

where $\{|v_i\rangle\}$ is the orthonormal basis in which Alice is to perform her measurement, while $\{\left|w_j\right\rangle\}$ is some orthonormal basis in Bob's Hilbert space. Let us rewrite this as

$$|\Psi\rangle = \sum_i \frac{1}{\mathcal{N}_i}\,|v_i\rangle \otimes |b_i\rangle\,, \tag{2.15}$$

where

$$|b_i\rangle = \mathcal{N}_i \sum_j \Psi_{ij}\left|w_j\right\rangle$$

is a vector in Bob's Hilbert space and

$$\mathcal{N}_i = \frac{1}{\sqrt{\sum_j |\Psi_{ij}|^2}} \tag{2.16}$$

is a normalization factor such that $\|\,|b_i\rangle\,\| = 1$ for all i (in the sum (2.15), we omit the terms with $\sum_j \Psi_{ij}\left|w_j\right\rangle = 0$, so all the $\mathcal{N}_i$ are finite).

So we have expressed the state to be measured as a sum of orthonormal components $|v_i\rangle \otimes |b_i\rangle$. These components have amplitudes $1/\mathcal{N}_i$, so the probability with

which Alice will detect the corresponding $|v_i\rangle$ is $\mathrm{pr}_{A,i} = 1/\mathscr{N}_i^2$. Whenever this happens, the state of Bob's system will become the corresponding $|b_i\rangle$.

Exercise 2.29. For a physical state $|\Psi\rangle$, show that, in Eq. (2.15), $\sum_i (1/\mathscr{N}_i^2) = 1$.

Exercise 2.30. For the state $|\Psi\rangle = \mathscr{N}(|RV\rangle + |H+\rangle)$:

a) find the factor $\mathscr{N}$ such that $|\Psi\rangle$ is normalized;
b) present this state in the form of Eq. (2.15), where $\{|v_i\rangle\}$ is the canonical basis;
c) find the probabilities of the possible results when Alice performs her local measurement on $|\Psi\rangle$ in the canonical basis, and specify the remotely prepared state of Bob's photon for each of Alice's results.

We have developed a method for predicting results of local measurements on an entangled state. This method is functional, but somewhat clumsy, so we shall now introduce a notion that will allow us to streamline the procedure.

The *partial inner product* between a local state $|a\rangle$ in the Hilbert space $\mathbb{V}_A$ and a bipartite state $|\Psi\rangle = \sum_{ij} \Psi_{ij} |v_i\rangle \big|w_j\big\rangle$ in the Hilbert space $\mathbb{V}_A \otimes \mathbb{V}_B$ (where $\{|v_i\rangle\}$ and $\{\big|w_j\big\rangle\}$ are orthonormal bases in $\mathbb{V}_A$ and $\mathbb{V}_B$, respectively) is a state in the Hilbert space $\mathbb{V}_B$ given by

$$\langle a|\,\Psi\rangle \equiv \sum_{ij} \Psi_{ij} \langle a|\,v_i\rangle \big|w_j\big\rangle; \tag{2.17a}$$

$$\langle \Psi|\,a\rangle \equiv \sum_{ij} \Psi_{ij}^* \langle v_i|\,a\rangle \big\langle w_j\big|\,. \tag{2.17b}$$

This definition is analogous for the partial inner product of $|\Psi\rangle$ with a local state in the space $\mathbb{V}_B$.

Exercise 2.31. For $|\psi\rangle = 2|H\rangle + \mathrm{i}|V\rangle$, find $\langle \psi_B|\,\Omega\rangle$ and $\langle \Pi|\,\psi_A\rangle$, where $|\Omega\rangle = 2|HH\rangle + 3|HV\rangle + 4|VH\rangle$, $|\Pi\rangle = (2|H\rangle + \mathrm{i}|V\rangle) \otimes (\mathrm{i}|H\rangle - |V\rangle)$ and subscripts A and B on a state $|\psi\rangle$ indicate that it is localized in Alice's or Bob's space, respectively.

Exercise 2.32. Show that, for any separable state $|ab\rangle \in \mathbb{V}_A \otimes \mathbb{V}_B$, and any state $|a'\rangle \in \mathbb{V}_A$,

$$\big\langle a'\big|\,ab\big\rangle = \big\langle a'\big|\,a\big\rangle\,|b\rangle \tag{2.18}$$

Exercise 2.33. Suppose $|\Psi\rangle$ is a state in the tensor product space, and $|a\rangle$ and $|b\rangle$ are states in Alice's and Bob's spaces, respectively. Show that

$$\langle a|\,(\langle b|\,\Psi\rangle) = \langle b|\,(\langle a|\,\Psi\rangle) = \langle ab|\,\Psi\rangle\,. \tag{2.19}$$

Exercise 2.34. Show that, for any two orthonormal bases $\{|v_i\rangle\} \otimes \{\big|w_j\big\rangle\}$ and $\{|v_i'\rangle\} \otimes \{\Big|w_j'\Big\rangle\}$ in $\mathbb{V}_A \otimes \mathbb{V}_B$, a local state $|a\rangle \in \mathbb{V}_A$ and a bipartite state

$$|\Psi\rangle = \sum_{ij} \Psi_{ij} |v_i\rangle \otimes \big|w_j\big\rangle = \sum_{ij} \Psi_{ij}' \big|v_i'\big\rangle \otimes \big|w_j'\big\rangle\,, \tag{2.20}$$

the partial inner product $\langle a|\,\Psi\rangle$ is independent of the choice of basis, i.e.,

$$\sum_{ij} \Psi_{ij} \langle a| v_i\rangle |w_j\rangle = \sum_{ij} \Psi'_{ij} \langle a| v'_i\rangle |w'_j\rangle . \tag{2.21}$$

Exercise 2.35. Show that, in Eq. (2.15),

a) $|b_i\rangle = \mathscr{N}_i \langle v_i| \Psi\rangle$;
b) $\| \langle v_i| \Psi\rangle \| = 1/\mathscr{N}_i$.

The last exercise offers a straightforward way to calculate the decomposition (2.15) for a given state and Alice's measurement basis, and hence also to calculate the results of local measurements. Indeed, the partial inner product gives us not only the state $|b_i\rangle$ that will be prepared remotely at Bob's location, but also the probability $\mathrm{pr}_{A,i} = 1/\mathscr{N}_i{}^2$ of each outcome on Alice's side.

These results, in essense, constitute extension of the Measurement Postulate of quantum physics to local measurements. Let us summarize them. Alice's local measurement on a bipartite state $|\Psi\rangle$ in the basis $\{|v_i\rangle\}$ will randomly collapse $|\Psi\rangle$ onto one of the states $\mathscr{N}_i |v_i\rangle \otimes \langle v_i| \Psi\rangle$ with the probability

$$\mathrm{pr}_{A,i} = \langle \Psi| v_i\rangle \langle v_i| \Psi\rangle . \tag{2.22}$$

This can be reformulated in terms of projection operators (Sec. 1.8): Alice's measurement transforms the state $|\Psi\rangle$ into a set of unnormalized states $\{\hat{\Pi}_i |\Psi\rangle\}$, where $\hat{\Pi}_i = |v_i\rangle\langle v_i|$ and the squared norm of each state in the set is the probability of the corresponding outcome.

After a local measurement, an entangled bipartite state will collapse into a separable one. If Alice destroys her system in the process of the measurement, the resulting state, $\mathscr{N}_i \langle v_i| \Psi\rangle$, will be localized with Bob.

Exercise 2.36. Solve Ex. 2.30(c) using partial inner products.

Exercise 2.37. For each Bell state, show that a local measurement by Alice in *any* orthonormal basis will yield either result with a probability of $1/2$.

Exercise 2.38.[§] Suppose Alice measures $|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|HV\rangle - |VH\rangle)$ in the circular polarization basis. Which state will Bob's photon project onto for each of Alice's results?

Exercise 2.39. Suppose Alice and Bob share the state $|\Psi^-\rangle$. Alice wishes to prepare some linear superposition $\alpha |H\rangle + \beta |V\rangle$ remotely at Bob's location, where α and β are arbitrary, but $|\alpha|^2 + |\beta|^2 = 1$ (i.e., the output state is normalized). In which basis should she measure? What is the probability of success?

2.2.3 Local measurements and causality

Now let us return to our previous discussion regarding the consistency of the remote preparation effect with causality. The fact that a measurement by Alice affects the state of Bob's photon does not in itself constitute a violation of causality. This is because the notion of a quantum state is rather abstract. The real question we need

to ask is this: will the *physical properties* of Bob's photon — that is, the behavior it will exhibit when measured, — change after Alice's measurement?

One could be tempted to give an affirmative answer. Indeed, before the measurement, Bob's state was part of a completely isotropic bipartite state; after the measurement, it is a state with a certain polarization angle — that is, one with drastically different physical properties. However, this answer misses an important point. Alice's local measurement does not always prepare the same state at Bob's location: sometimes it is $|\theta\rangle$, and sometimes $|\pi/2+\theta\rangle$. In order to know which one it is, Bob needs to receive classical communication from Alice, about the result she obtained from her measurement. Until then, Bob knows only that he has one of the two states — and thanks to this uncertainty, the measurable properties of Bob's photon are completely identical to those prior to the measurement. Before we prove this statement rigorously, let us look at an example.

Exercise 2.40. In the setting of Ex. 2.27, Bob measures the polarization of his photon in the canonical basis after Alice's measurement. What is the probability of each result given that Bob does not know the result of Alice's measurement?
Answer: $\mathrm{pr}_{\mathrm{Bob},H} = \mathrm{pr}_{\mathrm{Bob},V} = 1/2$ independently of the basis Alice uses.

Exercise 2.41. Alice and Bob perform measurements on their portions of a bipartite state $|\Psi\rangle$, in bases $\{|v_i\rangle\}$ and $\{|w_j\rangle\}$, respectively. These measurements can occur according to three alternative scenarios.

1. Alice and Bob perform their measurements simultaneously, so the original Measurement Postulate applies for a projective measurement of the state $|\Psi\rangle$ in the basis $\{|v_i\rangle \otimes |w_j\rangle\}$.
2. Alice performs her measurement first, and then Bob measures the remotely prepared state.
3. Bob performs his measurement first, and then Alice measures the remotely prepared state.

Show that the probability that Alice detects $|v_i\rangle$ while Bob detects $|w_j\rangle$ is the same for each of these scenarios: $\mathrm{pr}_{ij} = |\langle v_i w_j| \Psi\rangle|^2$.

Exercise 2.42. Test the statement of the previous exercise on the example of the state $|\Psi\rangle$ from Ex. 2.30 and the measurements performed by both parties in the canonical bases.

a) Find the probabilities pr_{HH}, pr_{HV}, pr_{VH} and pr_{VV} for Alice and Bob performing their measurements simultaneously.
b) Assume that Alice does her measurement first. Find the probabilities and the remotely prepared states of Bob's photon for each of Alice's results. Then let Bob measure each of these remotely prepared states and determine the corresponding probabilities. Use this information to evaluate pr_{HH}, pr_{HV}, pr_{VH} and pr_{VV} and check that they are the same as in part (a).
c)[§] Repeat part (b) for Bob performing his measurement first.

Exercise 2.43. For each of the scenarios of Ex. 2.41, show that the overall probability for Bob to observe the state $|w_j\rangle$ is $\| \langle w_j| \Psi\rangle \|^2$.

The above results mean that, without knowing the outcome of Alice's measurement, the physical properties of Bob's photon do not change, so Bob can extract *no information whatsoever* about Alice's actions. Although instant remote state preparation is predicted by theory and confirmed by experiment, it cannot be used for superluminal, interaction-free communication. Quantum mechanics leads us into thinking otherwise by telling us that Bob's *state* after Alice's measurement depends on the setup of that measurement. But in fact, the quantum state is a pure theoretical construct, and it is never directly observed in an experiment. We can infer information about the state only *indirectly* from the statistic obtained in multiple measurements.

So perhaps we can avoid all these paradoxes by dismissing the concept of the quantum state altogether and inventing another theory, which explains experimental results equally well but does not involve theoretical concepts that contradict common sense? We find the answer to this question in Sec. 2.3. For now, let us discuss another paradox that looks at the problem from an even more acute angle. Consider the following scenario.

1. Alice and Bob share many copies of state $|\Psi^-\rangle$.
2. On each copy, Bob *first* performs a measurement in the canonical, diagonal, or circular basis (he chooses randomly). *Then* Alice measures her photon in the basis $\{|\theta\rangle, |\frac{\pi}{2}+\theta\rangle\}$ and tells Bob her result.
3. After all measurements have been completed, Bob reconstructs the quantum state of his photon from the data he recorded using the techniques of quantum tomography (Ex. 1.15), taking into account (*postselecting*) only those events in which Alice detected $|\theta\rangle$.

If Bob's measurements occurred after Alice's, he would reconstruct the state $|\pi/2+\theta\rangle$ thanks to the effect of remote state preparation. But, as we know from Ex. 2.41, the correlated probabilities of Alice's and Bob's results do not depend on the sequence of measurements. That is, Bob will have exactly the same statistics for the outcomes of his measurements — that is, the same pr_H, pr_V, pr_+, pr_-, pr_R, pr_L — no matter whether his measurements occur before or after Alice's, and hence reconstruct the same state $|\pi/2+\theta\rangle$. So the remote state preparation effect takes place even *after* Bob has measured and destroyed his photon.

Exercise 2.44.* Show that, if quantum cloning were possible, superluminal communication would also be possible.
Hint: use remote preparation and quantum tomography.

2.2.4 Mixed states

Let us now consider a situation in which Alice loses her share of the entangled state or fails to inform us about her measurement results. The photon is absorbed on its way toward Alice's detector, or the detector fails to function, or Alice simply lets her

photon fly out of the lab window into the sky, where it may eventually get measured by distant aliens. What can we say about the quantum state of Bob's photon[6]?

One thing we do know (Ex. 2.41) is that no matter what happens to Alice's photon, the experimentally measurable properties of Bob's photon do not change. Therefore, as long as we are interested in describing Bob's photon, we can make any convenient assumption about the fate of Alice's photon. So let us assume that Alice has measured it in the canonical basis and did not tell us the result.

Specializing once again to the initial state being $|\Psi^-\rangle$, we know that, Alice can detect either $|H\rangle$ (in which case Bob's photon projects onto $|V\rangle$) or $|V\rangle$ (in which case Bob's photon projects onto $|H\rangle$). But since we do not know Alice's result, we can only describe the state of Bob's photon verbally as the *ensemble* "either $|H\rangle$ with probability $1/2$ or $|V\rangle$ with probability $1/2$".

This is the best we can do. Assuming other bases that Alice could have used, we could also describe Bob's photon as "either $|+45^\circ\rangle$ with probability $1/2$ or $|-45^\circ\rangle$ with probability $1/2$" (Ex. 2.9) or "either $|R\rangle$ with probability $1/2$ or $|L\rangle$ with probability $1/2$" (Ex. 2.38), and so on. All these descriptions are equivalent (Ex. 1.12). The polarization of Bob's photon is *completely mixed* — similar to that of natural light. *Its state is not represented by a certain vector in the Hilbert space.*

In Chapter 5 we will study properties of mixed states and ways to describe them mathematically. It is important to understand now, though, that if we lose a part of an entangled state, the remaining part *loses coherence*: it stops being in a superposition state and becomes just a statistical mixture. It is now described in terms of the classical theory of probabilities rather than quantum mechanics.

Let me note that we already discussed the loss of quantum coherence in the context of Welcher-Weg measurements in a quantum interference experiment (Sec. 1.5). In fact, this phenomenon is of the same nature as what we are studying here, as we shall see in Sec. 2.4.

Exercise 2.45. Alice and Bob share an entangled two-photon state:

a) $|\Psi\rangle = (|HH\rangle + 2|VV\rangle)/\sqrt{5}$;
b) $|\Psi\rangle = (|HH\rangle + |HV\rangle + |VV\rangle)/\sqrt{3}$.

Describe, in the form of an ensemble, the state of Bob's photon assuming that Alice measures the polarization of her photon (i) in the canonical basis and (ii) in the diagonal basis, but does not tell Bob the result.

In each part of this exercise, the ensemble describing Bob's mixed state depends on the basis in which Alice performs her measurement. But let me emphasize again: these different ensembles correspond to the same set of probabilities should Bob perform a measurement on his portion of the state. If this were not the case, Bob

[6] One may naïvely be tempted to answer that, when Alice's photon is lost from state $|\Psi^-\rangle = (|HV\rangle - |VH\rangle)/\sqrt{2}$, for example, Bob's photon becomes $(|V\rangle - |H\rangle)/\sqrt{2} = |-\rangle$. This is incorrect, of course. To see that, recall Ex. 2.9, where we found that $|\Psi^-\rangle$ can also be written as $(|+-\rangle - |-+\rangle)/2$. This means that Bob's photon has equal probabilities of being in states $|+\rangle$ and $|-\rangle$.

would be able to make inferences about Alice's actions, and this, as we found in Sec. 2.2.3, is impossible[7].

2.3 Quantum nonlocality

2.3.1 Einstein–Podolsky–Rosen paradox

In Sec. 2.2 we discussed local measurements on entangled states. We found that a local measurement by Alice leads to instant collapse of the nonlocal state into a state that is localized with Bob and depends on the measurement that Alice chooses to perform. We have shown that remote state preparation does not "violate causality, i.e., that the measurable properties of Bob's photon remain unaffected by Alice's measurement. We then argued that the quantum state is a pure theoretical construct, so it is "okay" for it to exhibit seemingly unphysical properties on paper as long as there are no consequences in the experiment. The problem, however, is not solved completely: if a theoretical model has nonsensical, counterintuitive elements that are not related to measurable physics, perhaps it's not a very good model!

This paradox was rigorously formulated for the first time in 1935, in a paper by Albert Einstein, Boris Podolsky, and Nathan Rosen (EPR)[8]. The original EPR paradox was proposed for the mechanical motion of a pair of particles, so we have to postpone its discussion to Chapter 3. Here we will instead discuss an alternative formulation similar to the one proposed by David Bohm in 1951[9].

The EPR argument relies on the notion of *physical reality*. An observable is defined to be an element of physical reality when the result of its measurement can be correctly predicted prior to measurement. For example, suppose Alice and Bob (two remote, non-interacting parties) share an entangled state $|\Psi^-\rangle = (|HV\rangle - |VH\rangle)/\sqrt{2}$ of two photons. Let Alice measure the polarization of her photon in the canonical basis. This measurement will remotely prepare state $|H\rangle$ or $|V\rangle$ at Bob's. If now Bob chooses to measure his photon in the canonical basis, his measurement result can be predicted with certainty, which means that the observable $\hat{\sigma}_z$ is an element of physical reality of Bob's photon.

If Alice instead measures in the diagonal basis, Bob's photon is remotely prepared as either $|+45^\circ\rangle$ or $|-45^\circ\rangle$. Now, if Bob chooses to measure his photon in the diagonal basis, his measurement result can be predicted with certainty — so in this case the observable $\hat{\sigma}_x$ corresponds to physical reality for Bob's photon.

Next, EPR argued that, if the two parties are far apart and/or cannot interact with each other, then no action by one party can change the physical reality at the ot-

[7] We show rigorously that Bob's ensembles obtained for the two Alice's measurement bases are identical in Ex. 5.40.

[8] A. Einstein, B. Podolsky, N. Rosen, *Can Quantum-Mechanical Description of Physical Reality be Considered Complete?*, Physical Review **47**, 777 (1935).

[9] D. Bohm, *Quantum Theory*, Prentice-Hall, Englewood Cliffs, 1951.

her. They called this the *locality principle* or *local realism*. Applying this commonsense principle to our case, we are forced to conclude that *both* $\hat{\sigma}_x$ and $\hat{\sigma}_z$ are parts of physical reality as far as Bob's photon is concerned. This is, however, impossible, because observables $\hat{\sigma}_x$ and $\hat{\sigma}_z$ have a non-overlapping set of eigenstates (see Ex. 1.35).

This contradiction lead EPR to conclude that "quantum-mechanical description of reality ... is not complete." Under *completeness*, EPR understand the requirement that "every element of physical reality must have a counterpart in physical theory". In the present case, two elements of physical reality — $\hat{\sigma}_x$ and $\hat{\sigma}_z$ — can have no more than one counterpart in the quantum theory.

EPR concluded their paper by saying:

> *While we have thus shown that the wave function does not provide a complete description of the physical reality, we left open the question of whether or not such a description exists. We believe, however, that such a theory is possible.*

In other words, maybe one day a theory will be developed which would predict experimental results as well as quantum mechanics, but at the same time exhibit no paradoxical features. Specializing to our case, the "new" theory will allow one to predict the results of Bob's measurement in any basis, independently of Alice's actions.

One may object that, according to the experiment, Bob's results, unless he measures in the same basis as Alice, are random. Doesn't this rule out any possibility of a deterministic theory? To answer this objection, let us recall the visualization we invented in Sec. 2.2.1: somebody randomly sending one shoe from a pair to Alice and the other to Bob. To Alice and Bob, the handedness of their shoe will appear random. Yet the party who packs the shoes and sends them out does know which shoe went to which observer: they possess the knowledge of a *hidden parameter* that Alice and Bob do not have access to.

The behavior of photons is more complex than the shoes, because the correlations between measurement results depend on the bases the two parties choose. But maybe it still permits a similar explanation? Maybe the two photons, at the time of their creation, are given a set of hidden parameters which somehow fully predetermine the outcome of their polarization measurements, in *any* bases, and we just don't yet know what these parameters are?

In 1935, quantum mechanics was already established as a powerful theory capable of explaining many experimental results better than any other theory. Therefore EPR did not challenge the ability of quantum mechanics to predict and explain experimental observations. They only pointed out the holes in its logic. While they hypothesized that there may exist a theory without these holes, they said nothing specific about this theory. As such, the EPR hypothesis appeared to leave no opportunity for experimental testing, and hence found itself outside the realm of physics, an inherently experimental science.

2.3.2 The Bell inequality

John Bell

The situation changed almost thirty years later, in 1964. *John Bell* proposed[10] an experiment in which *any* local realistic theory would predict a result that differs from what is predicted by quantum mechanics. Specifically, he derived an inequality that would hold in any local realistic theory, but is violated according to quantum mechanics.

Bell's discovery is ingenious because he found a way to test a theory without knowing anything about it — except that it follows common sense in the form of local realism. He tackled this nearly impossible mission by analyzing the experimental apparatus from the "front-panel" perspective, without making any assumption about the underlying physics. This very basic description of the experiment turns out to be sufficient to make significant predictions about its results.

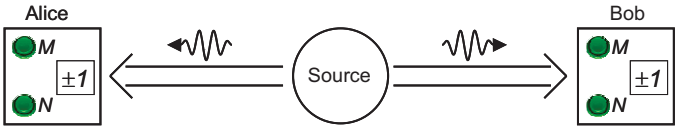


Fig. 2.2 Front panel of Bell's experiment.

Specifically, the front-panel description of the experiment is as follows. Each of the two remote observers, Alice and Bob, operates a device as shown in Fig. 2.2. Each device has two buttons marked M and N, and a display that can show either "$+1$" or "-1". Alice and Bob have no way to communicate with each other.

A "source", located about halfway between Alice and Bob, sends them a pair of particles of a certain kind. Alice and Bob receive the particles and insert them in their device. They then randomly choose and simultaneously push one of the buttons. Each device displays a value of ± 1, possibly related to the state of the particle received. We refer to this routine as an *event*.

Both observers keep records of the buttons they pressed and the numbers displayed. After acquiring the data for many events, the two parties meet and perform correlation analysis of their records. That is, they evaluate the quantity

$$\langle S \rangle = \langle M_A M_B - M_A N_B + N_A M_B + N_A N_B \rangle, \tag{2.23}$$

where $M_{A,B}$ and $N_{A,B}$ refer to the values obtained by each observer when they push the relevant button. Of course, each pair of particles contributes to only one term in Eq. (2.23). For example, if Alice pushes M and Bob N, the values they observed are used to evaluate $\langle M_A N_B \rangle$, and so on.

[10] J. S. Bell, *On the Einstein–Poldolsky–Rosen paradox*, Physics **1**, 195 (1964).

Let us write out Eq. (2.23) in its complete form:

$$\langle S\rangle = \sum_{M_A,M_B=-1}^{+1} \mathrm{pr}_{M_A,M_B} M_A M_B - \sum_{M_A,N_B=-1}^{+1} \mathrm{pr}_{M_A,N_B} M_A N_B + \sum_{N_A,M_B=-1}^{+1} \mathrm{pr}_{N_A,M_B} N_A M_B + \sum_{N_A,N_B=-1}^{+1} \mathrm{pr}_{N_A,N_B} N_A N_B, \tag{2.24}$$

where, e.g., pr_{M_A,N_B} for $M_A = 1, N_B = -1$ is the probability that Alice's display showed 1 and Bob's -1 under the condition that Alice pressed M and Bob N.

Now look at the structure of these probability distributions. Under local realism, each device determines the value displayed for each button on the basis of the local information which happens to be available — the hidden parameter of the particle that has arrived (which we denote λ_A and λ_B for Alice's and Bob's particles, respectively) and some algorithm. This algorithm may possibly involve randomness, so it is characterized by a set of probabilities $\mathrm{pr}_{M_A|\lambda_A}$, $\mathrm{pr}_{M_B|\lambda_B}$, $\mathrm{pr}_{N_A|\lambda_A}$, $\mathrm{pr}_{N_B|\lambda_B}$. For example, $\mathrm{pr}_{M_A|\lambda_A}$ determines the probability of the value M_A that will be displayed on Alice's apparatus when she presses the M button if the value of the hidden parameter of the incoming particle is λ_A.

Using the expression (B.6) for conditional probabilities, we can write the probability to obtain a certain *pair* of values on Alice's and Bob's displays as

$$\mathrm{pr}_{M_A,M_B} = \sum_{\lambda_A,\lambda_B} \mathrm{pr}_{\lambda_A,\lambda_B} \mathrm{pr}_{M_A|\lambda_A} \mathrm{pr}_{M_B|\lambda_B}, \tag{2.25}$$

for the case where both Alice and Bob press the M buttons. Here, $\mathrm{pr}_{\lambda_A,\lambda_B}$ is the probability of having a pair of particles with the hidden parameters λ_A and λ_B. Note that these parameters can be correlated because the particles come from the same source, so we may not express $\mathrm{pr}_{\lambda_A,\lambda_B}$ as a product of probabilities. For the other three possible combinations of buttons, the expressions are similar.

Exercise 2.46. Given the above result, show that Eq. (2.24) can be rewritten in the form

$$\langle S\rangle = \sum_{M_A,M_B,N_A,N_B=-1}^{+1} \mathrm{pr}_{M_A,M_B,N_A,N_B} [M_A M_B - M_A N_B + N_A M_B + N_A N_B], \tag{2.26}$$

where $\mathrm{pr}_{M_A,M_B,N_A,N_B}$ is a non-negative variable with the property

$$\sum_{M_A,M_B,N_A,N_B=-1}^{+1} \mathrm{pr}_{M_A,M_B,N_A,N_B} = 1.$$

Express $\mathrm{pr}_{M_A,M_B,N_A,N_B}$ in terms of $\mathrm{pr}_{\lambda_A,\lambda_B}$, $\mathrm{pr}_{M_A|\lambda_A}$, $\mathrm{pr}_{M_B|\lambda_B}$, $\mathrm{pr}_{N_A|\lambda_A}$, and $\mathrm{pr}_{N_B|\lambda_B}$.

The importance of Eq. (2.26) is that the set of four values $\{M_A, M_B, N_A, N_B\}$ obeys a valid probability distribution. This means that, for any local realistic experiment with Bell's front panel (Fig. 2.2), it is mathematically possible to construct an al-

ternative setup that would generate and display these four values for every event (Fig. 2.3), and these values would exhibit exactly the same statistics for each pair (M_A, M_B), (M_A, N_B), (N_A, M_B), (N_A, N_B) as does the original setup.

Note that this is not the case if the locality principle does not hold — for example, if M_A depends not only on λ_A, but also on the button that Bob pressed. This dependence would invalidate Eq. (2.25) and hence also Eq. (2.26).

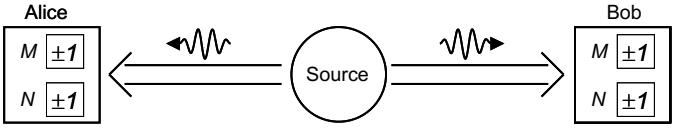


Fig. 2.3 Alternative scheme of Bell's experiment valid under local realism.

Exercise 2.47. Derive the Bell inequality

$$|\langle M_A M_B - M_A N_B + N_A M_B + N_A N_B \rangle| \leq 2 \tag{2.27}$$

for any apparatus whose front panel is represented by Fig. 2.3.
Hint: rewrite Eq. (2.26) as $\langle S \rangle = \langle M_A(M_B - N_B) + N_A(M_B + N_B) \rangle$.

The above result readily extends to any local realistic apparatus with Bell's front panel (Fig. 2.2). Indeed, if the Bell inequality did not hold for such a setup, it would also be violated for its counterpart in Fig. 2.3, and we have just shown this to be impossible.

Let me emphasize once again that our derivation did not rely on any assumption about the physics of the particles or the measurement devices, but only on very general principles of causality and local realism.

2.3.3 Violation of the Bell inequality

We now describe a specific experimental setup which has a front panel consistent with the above description, yet violates the Bell inequality. The two particles received by Alice and Bob are two photons in the Bell state $|\Psi^-\rangle$. Both Alice's and Bob's setups consist of a half-waveplate followed by a PBS with two photon detectors at its output ports (Fig. 2.1). When Alice and Bob press their buttons, the waveplates rotate through the angle $\theta/2$, where the values of θ are given in Table 2.1. The detectors are wired to the display so that the registration of the photon in the transmitted (reflected) port results in the number $+1$ (-1) being shown. This is equivalent to Alice and Bob each measuring the observable

$$\hat{\sigma}_\theta = |\theta\rangle\langle\theta| - \left|\frac{\pi}{2} + \theta\right\rangle\left\langle\frac{\pi}{2} + \theta\right| \tag{2.28}$$

In the following exercises, we shall make a quantum mechanical prediction for the statistics of the measurement outcome, so we will be able to determine the expectation value of the observable S.

Table 2.1 Angle θ in the observable (2.28) in Bell's experiment.

		Observer	
		Alice	Bob
Button pressed	M	0	$\pi/8$
	N	$\pi/4$	$3\pi/8$

Exercise 2.48. Write the observable (2.28) in the Dirac notation in the canonical basis.

Exercise 2.49. Calculate the expectation values of the following operators in the state $|\Psi^-\rangle$:

a) $\hat{M}_A \otimes \hat{M}_B$;
b) $\hat{M}_A \otimes \hat{N}_B$;
c) $\hat{N}_A \otimes \hat{M}_B$;
d) $\hat{N}_A \otimes \hat{N}_B$.

Hint: To reduce calculations, use the isotropicity of $|\Psi^-\rangle$ (Ex. 2.9).
Answer: a) $-\frac{1}{\sqrt{2}}$; b) $\frac{1}{\sqrt{2}}$; c) $-\frac{1}{\sqrt{2}}$; d) $-\frac{1}{\sqrt{2}}$.

We now find that, according to quantum mechanics, the expectation value of S is

$$\langle S\rangle = \langle \hat{M}_A\hat{M}_B - \hat{M}_A\hat{N}_B + \hat{N}_A\hat{M}_B + \hat{N}_A\hat{N}_B\rangle = -2\sqrt{2}, \tag{2.29}$$

which violates the Bell inequality (2.27).

This result concludes Bell's argument, which provides us with experimental means to test the Einstein–Podolsky–Rosen hypothesis.

Experimental tests of the Bell inequality began shortly after it had been formulated, and they continue today. All of them testify in favor of quantum mechanics. However, all the experiments done until recently contained *loopholes* — additional assumptions that had to be made in order to exclude a local realistic explanation of the observed results. At the time of writing this manuscript, in 2015, three experiments have been reported in which all the significant loopholes have been closed (Box 2.3).

There are two main types of loopholes. The *locality loophole* occurs if the observers are too close to each other (for example, on the same optical table), and they do not make their M versus N decisions quickly enough. Then they are, at least theoretically, able to influence each other. In experiments in which this loophole is eliminated, Alice's and Bob's laboratories are set up a few hundred meters or even kilometers apart. To decide on the basis, they use fast random number generators, typically based on quantum principles. Instead of rotating waveplates, they

change their measurement bases using electro-optical modulators — optical elements whose birefringence can be controlled within a few nanoseconds by means of an applied voltage. In this way, the decisions made by the two parties are separated by a space-like interval, thereby precluding any communication between them.

The *detection loophole* arises due to optical losses or inefficient detection. These imperfections result in a nonzero probability that, at Alice's or Bob's location, neither detector will register a photon. In this case, the value on the display of the corresponding party will be uncertain, which means that the front panel is no longer consistent with Fig. 2.2[11]. Many experiments deal with this issue by invoking the so-called *fair sampling assumption* that the losses occur randomly, and are not affected by the local hidden variables. Operating under this assumption, they calculate $\langle S \rangle$ taking into account only those events in which a photon has been detected by both Alice and Bob. This fair sampling assumption, while quite natural in the physical context of the setup in Fig. 2.1, is incompatible with the general ideology of the Bell theorem which does not allow any assumptions whatsoever about the physics of the experiment.

Exercise 2.50.[§] For the quantum optical setup discussed in this section, show that Alice and Bob, taken individually, will observe results $+1$ and -1 with equal probabilities, no matter what buttons they press.
Hint: Check out Ex. 2.37.

Exercise 2.51.[*] Suppose that the efficiency of each photon detector is 50%. The rest of the apparatus is ideal, so, under the fair sampling assumption, $\langle S \rangle = 2\sqrt{2}$. Propose a local realistic model for the particles and detectors that would reproduce this behavior.

Exercise 2.52. To perform a Bell experiment with imperfect detectors, the electronic circuits in Alice's and Bob's devices are programmed to randomly display $+1$ or -1 in those events when neither photon detector has clicked. Assuming that the rest of the apparatus is ideal, find the value for the left-hand side of the Bell inequality that would be obtained in this experiment as a function of the detector efficiency η. What is the minimum η for which the Bell inequality will be violated?

2.3.4 Greenberger–Horne–Zeilinger (GHZ) nonlocality

Following Bell's discovery, there have been many proposals for demonstrating the nonlocal nature of quantum mechanics. In this section, we will study an argument that is particularly spectacular because it contains no inequalities[12]. In discussing it,

[11] Of course, one can set up the electronics so that, if neither detector clicks, the display shows a value of ± 1 randomly. With this programming, the experiment will be consistent with Fig. 2.2, but then a problem emerges elsewhere (see Ex. 2.52).

[12] Theoretical proposal: D. M. Greenberger, M. A. Horne, A. Shimony, A. Zeilinger, in *Bell's Theorem, Quantum Theory, and Conceptions of the Universe* (M. Kafatos, ed.), p. 73 (Kluwer

Box 2.3 Experimental tests of the Bell inequality

The first tests were performed by *John Clauser* and his group* (1972) and, in a more complete form by *Alain Aspect* and coworkers** (1981-1982). At that time, parametric down-conversion was not sufficiently well understood, so atomic ensembles were used to prepare the required entangled states.

The locality loophole was closed by the group of *Anton Zeilinger**** (1998). Alice and Bob were separated by a distance of 400 meters and quantum random number generators were used to choose their measurement bases.

The detection loophole was closed for the first time in the group of *David Wineland*† (2001), but using qubits based on beryllium ions in a magnetic trap, rather than photons. Trapped ions can remain in a trap for a very long time and their quantum states can then be measured with high efficiency. However, the two ions on which that experiment was performed were located in the same trap, separated by a distance of only a few micrometers. Therefore the experiment was severely affected by the locality loophole.

In 2015, three experiments were published within three months of each other reporting the simultaneous closure of both loopholes. The first one, under the leadership of *R. Hanson*††, avoided the detection loophole by using entanglement swapping (Ex. 2.69) to entangle long-lived electron spin states of two nitrogen-vacancy centers in diamonds positioned 1.3 kilometers apart. The other two experiments, led by *A. Zeilinger*††† and *L. Shalm*§, used elaborate down-conversion setups and high-efficiency photon detectors to minimize the losses associated with the propagation and detection of the photons below the threshold required to violate the Bell inequality.

*S. J. Freedman and J. F. Clauser, *Experimental test of local hidden-variable theories*, Physical Review Letters **28**, 938 (1972).

A. Aspect, P. Grangier, G. Roger, *Experimental Realization of EinsteinPodolsky–Rosen–Bohm Gedankenexperiment: A New Violation of Bell's Inequalities*, Physical Review Letters **49, 91 (1982).

***G. Weihs, T. Jennewein, C. Simon, H. Weinfurter, A. Zeilinger, *Violation of Bell's inequality under strict Einstein locality conditions*, Physical Review Letters **81**, 5039 (1998).

† M. A. Rowe, D. Kielpinski, V. Meyer, C. A. Sackett, W. M. Itano, C. Monroe, D. J. Wineland, *Experimental violation of a Bell's inequality with efficient detection*, Nature **409**, 791 (2001).

†† B. Hensen *et al.*, *Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres*, Nature **526**, 682 (2015).

††† M. Guistina *et al. Significant-loophole-free test of Bell's theorem with entangled photons*, Physical Review Letters **115**, 250401 (2015).

§ L. K. Shalm *et al. A strong loophole-free test of local realism*, Physical Review Letters **115**, 250402 (2015).

we will follow the same logic as for Bell's theorem. We first consider the experiment from the front-panel perspective and draw conclusions on the assumption of local realism. We then describe the physics underlying the front panel and calculate the theoretical predictions in accordance with the laws of quantum mechanics.

In GHZ, there are three remote observers, Alice, Bob, and Charley. Each of them operates a device similar to Bell's, but the buttons are marked σ_x and σ_y. In each event, the source simultaneously sends three particles to Alice's, Bob's, and Charley's devices, where they measure them by pressing one of the buttons. After many events, the parties meet and discuss the results.

Academic, Dordrecht, 1989). Experiment: J. W. Pan, D. Bouwmeester, M. Daniell, H. Weinfurter and A. Zeilinger, *Experimental test of quantum nonlocality in three-photon GHZ entanglement*, Nature **403**, 515 (2000).

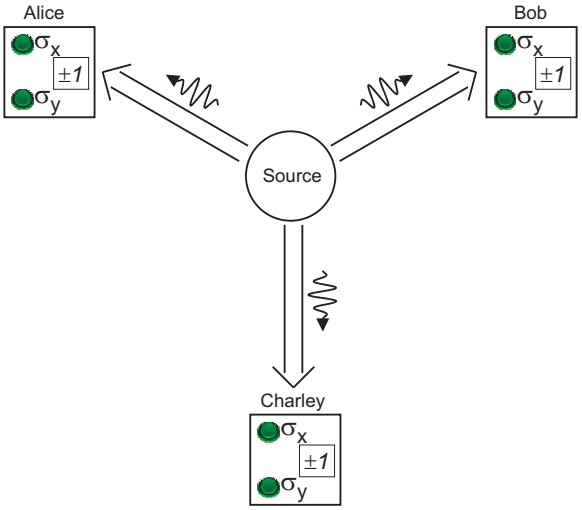


Fig. 2.4 Setup for the Greenberger–Horne–Zeilinger experiment

The apparatus is known to have the following property (which Alice, Bob, and Charley check by analyzing the statistics of their measurement results): whenever any two of the parties press the σ_y button, and the third one the σ_x button, the product of the three results is always -1.

$$\sigma_{x_A}\sigma_{y_B}\sigma_{y_C} = -1; \tag{2.30a}$$
$$\sigma_{y_A}\sigma_{x_B}\sigma_{y_C} = -1; \tag{2.30b}$$
$$\sigma_{y_A}\sigma_{y_B}\sigma_{x_C} = -1. \tag{2.30c}$$

Exercise 2.53. Assuming local realism and using hidden parameters akin to Sec. 2.3.2, show that one can define a common probability distribution $\mathrm{pr}_{\sigma_{xA},\sigma_{yA},\sigma_{xB},\sigma_{yB},\sigma_{xC},\sigma_{yC}}$ that simultaneously governs all possible observations that can be made in a GHZ experiment. Show that this probability is always nonnegative and its values add up to one.

We now follow the same line of logic as when deriving the Bell inequality. Because the possible sets of results $(\sigma_{iA}, \sigma_{jB}, \sigma_{kC})$ (where each index i, j, and k can be x or y) obey a common probability distribution, one can construct an alternative experiment in which the three devices have no buttons, but display both values of σ_x and σ_y for every event. This alternative experiment would exhibit the same statistical properties as the original one. In particular, Eqs. (2.30) would hold for every event.

Let us now multiply the left- and right-hand sides of these three equations together and conclude that the following is true for every triplet of particles:

$$\sigma_{x_A}\sigma_{x_B}\sigma_{x_C} = -1. \tag{2.31}$$

Because this result holds for the alternative experiment, it should also hold for the original one. That is, whenever all three observers push the "σ_x" button, the product of the displayed values will be -1. This conclusion is true under local realism.

Now we consider the quantum argument. The source sends three photons in the state

$$|\Psi_{GHZ}\rangle = \frac{1}{\sqrt{2}}(|HHH\rangle + |VVV\rangle)$$

to Alice, Bob, and Charley. When each of the observers presses one of the buttons, the Pauli observable corresponding to that button is measured on that observer's photon and the measurement result is displayed, corresponding to one of the eigenvalues of that observable.

Exercise 2.54. Show that $|\Psi_{GHZ}\rangle$ is an eigenstate of the operators

a) $\hat{\sigma}_{x_A} \otimes \hat{\sigma}_{y_B} \otimes \hat{\sigma}_{y_C}$, $\hat{\sigma}_{y_A} \otimes \hat{\sigma}_{x_B} \otimes \hat{\sigma}_{y_C}$, $\hat{\sigma}_{y_A} \otimes \hat{\sigma}_{y_B} \otimes \hat{\sigma}_{x_C}$ with eigenvalue -1;
b) $\hat{\sigma}_{x_A} \otimes \hat{\sigma}_{x_B} \otimes \hat{\sigma}_{x_C}$ with the eigenvalue $+1$.

Part (a) of the above exercise means that, whenever two of the three observers measure $\hat{\sigma}_y$, and the third measures $\hat{\sigma}_x$ on their portions of $|\Psi_{GHZ}\rangle$, the product of their measurement results will be -1 (see Ex. 2.24). This implies that the setting fits the front-panel description given above.

Part (b), on the other hand, proves that whenever all three observers measure σ_x, the product of their results will be $+1$. This result is in direct contradiction with the local realistic prediction (2.31). In contrast with the Bell inequality, where the violation of local realism is observed by collecting multiple data and calculating averages, the GHZ setting shows a discrepancy *every time* the observers perform a measurement. This absence of statistical uncertainty makes the GHZ argument especially appealing as a means of demonstrating quantum nonlocality.

2.4 An insight into quantum measurements

2.4.1 Von Neumann measurements

At the end of the previous chapter, we learned that *any* quantum process is described by a unitary operator. On the other hand, the Measurement Postulate states that a measurement converts a quantum superposition into a statistical mixture of elements of the measurement basis[13]. This process cannot be described by a linear operator, which, by definition, reversibly maps any element of the Hilbert space to another element of the Hilbert space. How can this inconsistency be resolved?

If this question sounds too abstract, let us restate it in more concrete terms. Suppose an observer, Alice, is measuring a diagonally polarized photon

[13] This standard approach to quantum measurements is referred to as the *Copenhagen interpretation* in honor of Niels Bohr.

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(\alpha\,|H\rangle + \beta\,|V\rangle) \tag{2.32}$$

John von Neumann

(where α and β are real and positive) in a canonical basis [Fig. 1.2(a)]. The photon propagates through or reflects off the PBS, then hits the sensitive area of one of the photodetectors (Box 1.2), triggering an avalanche, which, in turn, produces an audible click that Alice can hear. *At which point* does the superposition (2.32) collapse into a set of probabilities? Is it when the photon passes the PBS? Or when an avalanche occurs in one of the detectors? Or when the click sounds?

To answer these questions, let me introduce a model of quantum measurements proposed by John von Neumann. This model treats the quantum system to be measured and the measurement apparatus as two Hilbert spaces, which become *entangled* during the measurement. Suppose the system is initially in the state $|\psi\rangle = \sum_i \psi_i\,|v_i\rangle$, where $\{|v_i\rangle\}_{i=1}^N$ is the measurement basis. The initial state of the apparatus is $|w_1\rangle$, an element of the orthonormal basis $\{|w_i\rangle\}_{i=1}^M$ (with $M > N$) in the Hilbert space of the apparatus. During the measurement, the system interacts with the measurement apparatus and entangles itself with it through a unitary evolution, generating the state[14]

$$\hat{U}(|\psi\rangle \otimes |w_1\rangle) = \sum_{i=1}^{N} \psi_i\,|v_i\rangle \otimes |w_i\rangle\,. \tag{2.33}$$

The states $|w_1\rangle, \ldots, |w_n\rangle$ are macroscopically distinct (for example, different lights coming on or a needle being displaced to a different position). The observer has access to the apparatus and can learn the state of the system.

In the example mentioned at the beginning of this section, the entanglement of the system with the apparatus produces the state[15]

$$\begin{aligned} |\Psi_{SA}\rangle = {} & \alpha\,|H\rangle \otimes |\text{avalanche in detector 1}\rangle \\ & + \beta\,|V\rangle \otimes |\text{avalanche in detector 2}\rangle\,. \end{aligned} \tag{2.34}$$

The superposition (2.34) is relevant to the Welcher-Weg measurement situation of Sec. 1.5. Even if there is no observer present to read off the measurement result, the photon alone can no longer exhibit interference because the superposition state now involves an additional object: the apparatus.

Suppose now that the experiment is carried out by an observer, Alice, who can repeat it many times. In principle, Alice is able to verify the entangled nature of the superposition (2.34) experimentally. To that end, she would first perform multiple measurements of the photon in the canonical basis and correlate these results with

[14] It may appear that Eq. (2.33) is equivalent to quantum cloning (Sec. 2.1.3) because, for each basis element of the system, the apparatus evolves into the corresponding basis element of its Hilbert space. In fact, this is not the case. A proper cloning operation would also clone superposition states, i.e. it would make the right-hand side of Eq. (2.33) look like $\left(\sum_{i=1}^N \psi_i\,|v_i\rangle\right) \otimes \left(\sum_{i=1}^N \psi_i\,|W_i\rangle\right)$. Transformation (2.33) is different and hence does not contravene the no-cloning theorem.

[15] For the sake of argument, let us assume that the photon is not destroyed at the time of detection, and let us also neglect the different spatial paths taken by the horizontal and vertical photons.

the indications of the detectors. This would tell her the absolute values of α and β for the two terms in Eq. (2.34). Additionally, Alice would acquire the measurement statistics for both the photon and the detectors in the diagonal basis [for the detectors, it is $\{(|\text{avalanche in detector 1}\rangle \pm |\text{avalanche in detector 2}\rangle)/\sqrt{2}\}$] and determine the phase relation between the superposition terms (see Ex. 2.11). Such a measurement is, of course, far beyond our technical capability, but theoretically feasible.

But what if Alice remains passive and hears the click from one of the detectors in the state (2.34)? Since she is also a quantum object, we can continue our line of argument and say that she becomes a part of the same entangled superposition:

$$|\Psi_{SAO}\rangle = \frac{1}{\sqrt{2}}(|H\rangle \otimes |\text{avalanche in detector 1}\rangle \otimes |\text{Alice heard click in detector 1}\rangle + \beta\,|V\rangle \otimes |\text{avalanche in detector 1}\rangle \otimes |\text{Alice heard click in detector 2}\rangle).$$

This point marks a fundamental change for Alice as the observer. However intelligent and well-equipped she may be, she is unable, even in principle, to project *herself* onto the diagonal basis. As a result, she has no possibility of knowing she is in a superposition state. For her, whenever the photon is horizontal, she hears the click in detector 1, and vice versa. She has no possibility of finding out experimentally that there exists another part of the superposition, because everything she can observe (the photon and the detector) is consistent with her own state. For Alice, it will appear that the quantum state of the photon has collapsed and one of the two outcomes has occurred randomly.

Another observer, Bob, who is outside Alice's lab and has not yet become a part of the superposition, is still able to check the existence of the superposition by measuring the photon, the detector, *and* Alice in their diagonal bases. Here, again, I am asserting only the theoretical possibility of such a measurement, abstracting from its practical implementation (which is prohibitively complicated)[16].

We see that the von Neumann model answers the question asked at the beginning of this section. The collapse of the superposition need not be a part of the quantum theory — it is a *subjective* phenomenon that *seems* to happen when the observer becomes a part of the superposition. In reality, or at least in the theoretical reality put forward by quantum mechanics, the superposition never collapses, but lives on, involving an ever increasing part of the universe.

From this point of view, Einstein's sentiment that God does not play dice turns out to be valid. The evolution of the wavefunction of the universe is deterministic. The quantum randomness is only an illusion, a consequence of our limited observational power.

This interpretation, while eliminating the logical inconsistency noted at the beginning of this section, is highly unpractical. If our goal is purely utilitarian — to use quantum mechanics to predict the experimental results that are relevant to us as

[16] This is known as *Wigner's friend* Gedankenexperiment. It has been proposed by Eugene Wigner, who placed himself in the position of Bob and and a hypothetical friend in the position of Alice.

observers, it makes no sense to speculate about measurements that are possible only in theory. Instead we should adopt the Copenhagen interpretation and assume that the state collapses as soon as its degree of macroscopicity becomes so large that we are no longer able to measure the phase between the two terms of the superposition. In the example above, this would happen with the emergence of the avalanche in one of the detectors[17].

2.4.2 Decoherence

Quantum measurements do not only occur in laboratories. Measurement-like phenomena, in which the environment plays the role of the apparatus, take place around us all the time. Suppose, for example, that we prepare a single atom in a state $|\psi\rangle$ with a well-defined momentum. According to the uncertainty principle, the position observable is uncertain in this state, which means we can write it as a superposition of many position eigenstates[18].

$$|\psi\rangle = \sum_i \psi_i |x_i\rangle \tag{2.35}$$

Unless the atom is in a perfect vacuum, it will interact with other particles, such as gas molecules and photons. Most such interactions have a highly local character. For example, collisions between atoms are governed by the Lennard-Jones potential, whose strength falls off inversely proportionally to at least the sixth power of the distance. Accordingly, any such interaction changes the state of the environment particles depending on the atom's position. One can therefore say that the environment performs a measurement of the atom's state in the eigenbasis of the position observable. The joint state of the atom and the environment becomes

$$\sum_i \psi_i |x_i\rangle_{\text{atom}} \otimes |x_i\rangle_{\text{environment}} . \tag{2.36}$$

A realistic observer cannot keep track of the multiple objects that have interacted with "our" atom. Therefore, from that observer's point of view, this atom will eventually find itself in the situation discussed in Sec. 2.2.4. It will no longer be in a coherent superposition of position eigenstates (which constitutes a momentum eigenstate), but in a statistical mixture thereof. This loss of coherence due to interaction of a quantum system with its environment is called *dephasing* or *decoherence*.

[17] Such an "instrumentalist" approach was particularly favored by Richard Feynman, whose views are nicely summarized by the fictional slogan "Shut up and calculate" (N. D. Mermin, *Could Feynman have said this?*, Physics Today **57**, 10 (2004)).

[18] The precise form of the momentum eigenstate will be discussed in the next chapter; for the present argument, Eq. (2.35) is sufficient.

Exercise 2.55. The atom is initially in the state (2.35). It experiences decoherence, entangling itself with the environment according to Eq. (2.36). Write down the ensemble describing the state of the atom state after the decoherence.

Because the environment-induced measurement occurs in the position basis, it has no detrimental effect on position eigenstates. Indeed, if the atom is prepared in a state with a certain position, the sum Eq. (2.35) consists of only a single term. The interaction with the environment would then not lead to entanglement and the state (2.36) would be separable.

It is quite typical for the interactions of the system with the environment to be dominated by a single physical mechanism. Accordingly, there will be a certain basis in the system's Hilbert space whose elements will not entangle themselves with the environment and hence will not decohere[19]. It is called the *decoherence-preferred basis*.

Because of local nature of physical interactions, the position basis is often the decoherence-preferred basis for motional degrees of freedom. This is why it is much easier to prepare objects in a state with a certain position rather than a certain momentum, even though mathematically there is no preference for either case. A similar line of argument can also explain why superpositions of dead and living cats are never observed in practice, even though mathematically these states are as legitimate as each of the superposition terms. The environment persistently measures the location of various body parts of the cat by interacting with them. Because the results of these measurements contain information about whether the cat is dead or alive, any coherent superposition of these states will instantly decohere. In other words, the decoherence-preferred basis of the cat, whatever it is, does not include superpositions of dead and alive states.

For internal states of quantum objects, as well as motion on a microscopic scale, such as the motion of electrons in atoms, the position basis is not decoherence-preferred. This is because those electrostatic interactions that cause decoherence are typically caused by objects that are located much further away from the atom than the size of the atom itself.

A much more likely candidate for the status of the decoherence-preferred basis for internal states is the eigenbasis of the energy operator, i.e., the Hamiltonian. This is a consequence of the adiabatic theorem (Box 2.4). Because the electronic energy levels in atoms are quite far from each other (Sec. 4.4), the fields arising during a collision are typically "smooth" enough so an atom initially in an energy eigenstate will stay in that state[20]. On the other hand, a collision will unpredictably affect the *quantum phase* of each energy eigenstate, which evolves according to $|\psi_E(t)\rangle = \mathrm{e}^{-\mathrm{i}\int E(t)/\hbar \mathrm{d}t}|\psi_E(0)\rangle$ as per Eq. (1.25). So, while energy eigenstates are often preserved in a collision, coherence between them is typically not. This behavior is characteristic of a decoherence preferred basis, and the main reason why it is

[19] W. H. Zurek, *Decoherence, einselection, and the quantum origins of the classical*, Reviews of Modern Physics **75**, 715 (2003).

[20] Such collisions are called *elastic*.

Box 2.4 Adiabatic theorem

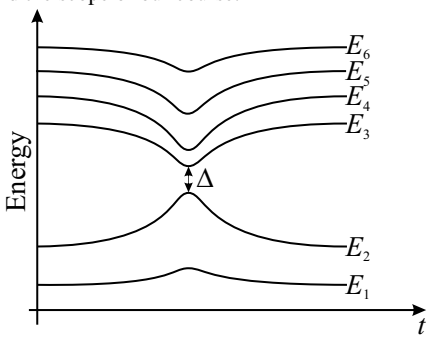
Suppose that at time $t = 0$ a certain quantum system is in an eigenstate $|\psi(0)\rangle = |E_m\rangle$ of its Hamiltonian. The Hamiltonian $\hat{H}(t)$ is time-dependent and has a discrete energy spectrum $\{E_n(t)\}$. The adiabatic theorem, due to Max Born and Vladimir Fock (1928), states that, if the Hamiltonian changes sufficiently slowly, the system will to a good approximation remain in the same energy eigenstate.

As a visual example, consider the following experiment which can be performed at home. Place a compass between the poles of a U-shaped magnet. The needle will orient itself along the field lines of the magnet. Now, if we slowly turn the magnet, the needle will follow its orientation. If, on the other hand, we flip the magnet quickly, or the magnet is weak, the needle will lose its alignment with it, and take some time to regain it. This, in essence, is the adiabatic theorem.

The adiabaticity condition can be loosely formulated as

$$\frac{\mathrm{d}}{\mathrm{d}t} E_n \ll \hbar \Delta^2, \tag{2.37}$$

where Δ is the minimum distance between E_m and other energy eigenvalues during the evolution process (see the figure below). A complete proof of the adiabatic theorem is relatively complicated and beyond the scope of our course.



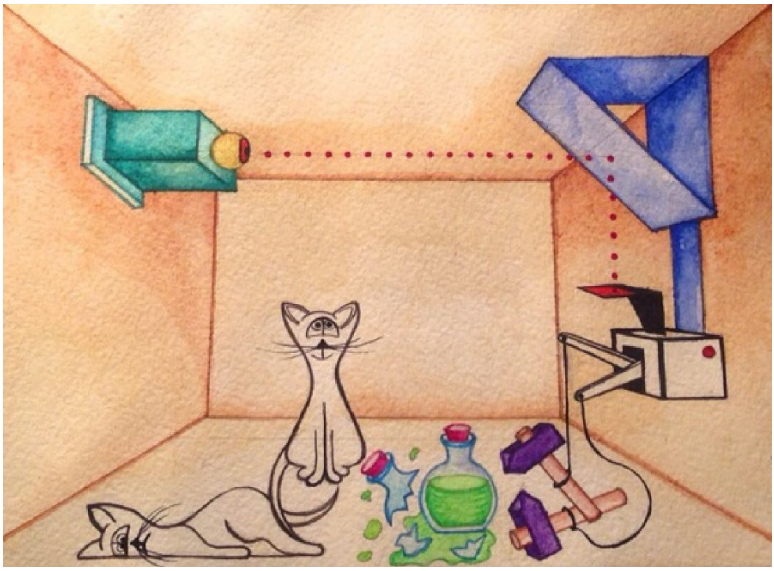
Schematic view of the evolution of atomic energy eigenvalues during a collision. The value of Δ shown is relevant to the adiabaticity condition for $m = 2$.

the energy eigenstates in which atoms and molecules are most frequently observed in experiments.

2.4.3 Interpretations of quantum mechanics

In Sec. 2.4.1, we analyzed the measurement process: the quantum object becomes entangled with the macroscopic measurement apparatus and, subsequently, with the experimentalist. Thereafter, we discussed a process of similar nature, in which the role of the experimentalist is played by the natural environment. In both cases it is clear that the expansion of the superposition to encompass further objects will con-

Box 2.5 Schrödinger's cat



Schrödinger's cat is something more complex than just a superposition of the dead and alive states of a cat. Here is a quote from the 1935 Schrödinger's article in the German magazine Naturwissenschaften ("Natural Sciences").*

> *A cat is penned up in a steel chamber, along with the following device (which must be secured against direct interference by the cat): in a Geiger counter, there is a tiny bit of a radioactive substance, so small that perhaps in the course of the hour, one of the atoms decays, but also, with equal probability, perhaps none; if it happens, the counter tube discharges, and through a relay releases a hammer that shatters a small flask of hydrocyanic acid. If one leaves this entire system to itself for an hour, one would say that the cat still lives if meanwhile no atom has decayed. The psi-function of the entire system would express this by having in it the living and dead cat (pardon the expression) mixed or smeared out in equal parts.*

In modern language, the quantum state of the cat and the atom is entangled and described by the equation

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|\text{atom not decayed}\rangle \otimes |\text{cat alive}\rangle + |\text{atom decayed}\rangle \otimes |\text{cat dead}\rangle).$$

From the subjective point of view of the cat inside the box, the quantum superposition has collapsed as soon as this entangled state has been formed (Sec. 2.4.1). An experimentalist outside with infinite technological capabilities, however, is in principle able to verify the presence of the superposition by projecting both the cat and the atom onto diagonal bases.

*E. Schrödinger, *Die gegenwartige Situation in der Quantenmechanik*, Naturwissenschaften **23**, 807–812, 823–828, 844–849 (1935).

Original artwork: courtesy R. Kazaryan.

tinue beyond the point where we stopped our analysis. The experimentalist, Alice, will act at least slightly differently depending on the detector in which she observed an event; this difference, however small, will affect atoms and photons surrounding her. Similarly, the particles that collided with the atom in the decoherence example will interact with other objects, undergo optical transitions, and so on. The more macroscopic a quantum superposition is, the more likely it is to involve more and more objects. The entanglement brought about by any measurement, deliberate or inadvertent, inevitably expands, eventually to encompass the entire universe and create a giant superposition state.

Measurement-like situations, in which the state of a microscopic object affects that of macroscopic objects, occur uncountably often in nature. Accordingly, the universe finds itself in an unimaginably complex superposition state. Taking this logic to the extreme, one would argue that all randomness in the world is of a quantum nature. For example, when we toss a coin, its motion is affected by minute vibrations of our hands and the motion of air molecules, both of which are affected by quantum fluctuations. Every hurricane is a result of a quantum fluctuation somewhere back in time, somewhere in the world. For any possible outcome of a random event or series thereof, however unlikely, there exists a "universe" in which it took place.

This is called the *many-worlds interpretation* of quantum mechanics. It was proposed by Hugh Everett in 1957. While it is a logical consequence of the quantum theory, the conclusion about the existence of multiple universes is purely speculative in the sense that it cannot be verified experimentally. As discussed above, once we become a part of the entangled superposition state, we no longer have the means to characterize that state.

Moreover, this conclusion is valid only if one believes quantum physics to be the ultimate, universal theory of the world. While all experiments so far have suggested this to be the case, these experiments are limited by our ability to isolate quantum systems from the environment. The largest objects for which quantum superpositions have been observed are organic molecules consisting of a few thousand atoms. One could imagine that, once one exceeds a certain degree of complexity, quantum superpositions cease to exist for fundamental reasons; in fact, certain arguments stemming from general relativity seem to suggest that. Finding the limits of applicability of quantum physics is therefore an important open question in modern physics. To address it, we need to construct increasingly macroscopic superposition states and check whether they retain their properties in spite of their large size.

The rather shocking nature of the many-worlds interpretation is often seen as the strongest argument against it. One must remember, however, that the Copenhagen interpretation is also full of paradoxes, some of which we have seen on the pages of this book. These paradoxes arise entirely due to the notion of quantum state collapse associated with measurements, and they disappear in the many-worlds picture, where such collapse does not exist.

Consider the Elitzur-Vaidman "Bomb" paradox (Box 1.4), for example. In the framework of the the many-worlds interpretation, we would say that the photon, initially in a localized superposition state $|+\rangle = \frac{|H\rangle+|V\rangle}{\sqrt{2}}$, experiences evolution as it

propagates through the interferometer, and at some point entangles itself with the bomb. The state of the two objects then becomes

$$\frac{1}{\sqrt{2}}\left[|H,\ \text{lower path}\rangle \otimes |\text{bomb exploded}\rangle + |V,\ \text{upper path}\rangle \otimes |\text{bomb not exploded}\rangle\right].$$

This entanglement changes the state of the photon, so not surprisingly, it will continue to evolve through the interferometer in a different manner compared to the case in which the bomb is absent. Eventually, it will become absorbed by one of the two detectors, so the state becomes[21]

$$\begin{aligned}
&\frac{1}{\sqrt{2}}\,|\text{bomb exploded}\rangle \otimes |\text{detector ``}+\text{'' not clicked}\rangle \otimes |\text{detector ``}-\text{'' not clicked}\rangle \\
&+\frac{1}{2}\,|\text{bomb not exploded}\rangle \otimes |\text{detector ``}+\text{'' clicked}\rangle \otimes |\text{detector ``}-\text{'' not clicked}\rangle \\
&+\frac{1}{2}\,|\text{bomb not exploded}\rangle \otimes |\text{detector ``}+\text{'' not clicked}\rangle \otimes |\text{detector ``}-\text{'' clicked}\rangle
\end{aligned}$$

It is therefore incorrect to say that a click of detector "−" has occurred without interaction of the photon with the bomb. The interaction has in fact occurred, and given rise to the entangled superposition above, in which the event in the detector "−" is only one of the terms. But because this superposition includes macroscopic objects, it will quickly encompass the entire universe, including the observers. Therefore the observers in the "universe" in which detector "−" clicked will have no possibility to see the other terms of the superposition. In their view, the other terms do not exist, and hence the bomb has been detected without interaction.

Exercise 2.56. Two photons in the Bell state $|\Psi^-\rangle$ are distributed to Alice and Bob. They perform non-destructive von Neumann measurements on their photons in the bases $\{|\theta\rangle, |\frac{\pi}{2}+\theta\rangle\}$ and $\{|H\rangle, |V\rangle\}$, respectively. What is the state of the two photons and two measurement apparata after this measurement? Denote the relevant basis elements in the Hilbert spaces of Alice's and Bob's apparata by $\{|w_{A1}\rangle, |w_{A2}\rangle\}$ and $\{|w_{B1}\rangle, |w_{B2}\rangle\}$, respectively.
Hint: use Eq. (2.6).

*2.4.4 The superposition tree**

Before we conclude our discussion of the many-worlds interpretation, we must address an important outstanding question. We argue that the collapse of a quantum state is a subjective phenomenon that occurs only in the view of the observer once that observer becomes a part of the entangled state. But then it must follow that the Measurement Postulate of quantum mechanics is not really a postulate: it must be a consequence of the Hilbert Space Postulate. That is, we should be able to derive

[21] I omitted the state of the photon for brevity.

Born's rule — that the measurement output probability, *as it appears to the observer,* is the square of the absolute value of the amplitude. This is our goal in this section.

Before proceeding, I would like to caution the reader that this section is quite difficult (even more so than the other chapters of this book) and not a part of "mainstream" quantum mechanics. Therefore I would recommend skipping it in the first reading.

Without attempting to be extremely rigorous, let us develop an argument for Born's rule for the state (2.32). How does the observer, Alice, determine the probability? She repeats the experiment many times and counts how many times each result is observed. The trouble is, though, that Alice herself is a part of the superposition state, so these apparent probabilities are different in each term of the superposition. For example, there exists a "universe" in which Alice repeated the experiment a thousand times and observed $|H\rangle$ every time. Alice in this universe will then conclude that the probability of detecting $|H\rangle$ is unity, in direct contradiction with Born's rule.

However, we can prove that Born's rule holds *in the overwhelming majority of universes*.

We start with a simplified case of equal probabilities for the two measurement outcomes, so that $\alpha = \beta = 1/\sqrt{2}$. Let us suppose that Alice performs measurements on multiple copies of the superposition $(|H\rangle + |V\rangle)/\sqrt{2}$ in the canonical basis. After the first measurement, she becomes part of an entangled state which contains two terms:

$$|\Psi\rangle = \tfrac{1}{\sqrt{2}}\,(|H\rangle \otimes |\text{Alice observed } H\rangle + |V\rangle \otimes |\text{Alice observed } V\rangle). \tag{2.38}$$

After the second measurement, there will be four terms:

$$\begin{aligned}|\Psi\rangle = \tfrac{1}{2}\,(&|HH\rangle \otimes |\text{Alice observed } H \text{ in 1st measurement, } H \text{ in 2nd measurement}\rangle \\ &+ |HV\rangle \otimes |\text{Alice observed } H \text{ in 1st measurement, } V \text{ in 2nd measurement}\rangle \\ &+ |VH\rangle \otimes |\text{Alice observed } V \text{ in 1st measurement, } H \text{ in 2nd measurement}\rangle \\ &+ |VV\rangle \otimes |\text{Alice observed } V \text{ in 1st measurement, } V \text{ in 2nd measurement}\rangle),\end{aligned} \tag{2.39}$$

and so on. This superposition can be visualized as a tree-like structure, with every measurement doubling the number of terms in the superposition, and doubling the branches of the tree [Fig. 2.5(a)]. After n measurements, the number of terms will be 2^n. Each term has amplitude $\sqrt{1/2^n}$ and corresponds to a unique downward path along the tree branches.

Although in each term of the superposition Alice cannot see other terms, she is aware of the full history of the measurement results that occurred within her term. Accordingly, she can calculate the frequency of occurrences for each of her results and interpret these statistics as probabilities. Specifically, if she observed $|H\rangle$ k times and $|V\rangle$ $n-k$ times, she will conclude that the probability of detecting $|H\rangle$ is k/n.

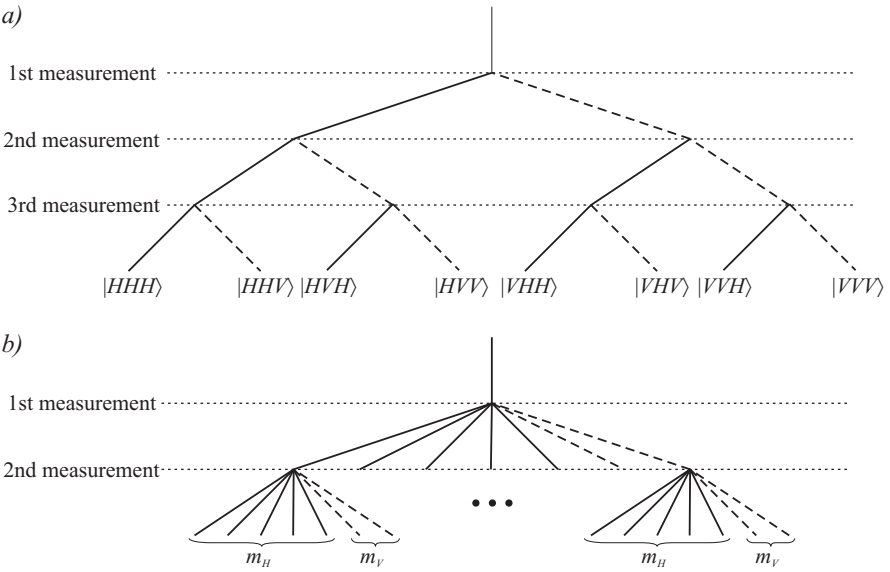


Fig. 2.5 The superposition tree. Solid lines correspond to observations of the horizontal polarization, and dashed lines to vertical. a) For an equal superposition of $|H\rangle$ and $|V\rangle$, each measurement doubles the number of branches. b) For unequal superpositions, we use multiple branching to equalize the amplitude of each branch.

Exercise 2.57. Suppose Alice performs a large number n of measurements on copies of the state $(|H\rangle + |V\rangle)/\sqrt{2}$. Calculate the fraction of paths along the superposition tree that contain k results $|H\rangle$ and $n-k$ results $|V\rangle$.
Hint: Review Ex. B.8.
Answer:

$$\frac{1}{2^n}\binom{n}{k} = \frac{1}{2^n}\frac{n!}{k!(n-k)!}. \tag{2.40}$$

Exercise 2.58.§ Calculate the above result numerically and plot it as a function of k for $n = 100$.
Answer: see Fig. 2.6(a).

Exercise 2.59.* Show that, for $n \gg 1$, Eq. (2.40) can be approximated by the Gaussian function

$$A(n)\exp\left[-2\frac{\left(k-\frac{n}{2}\right)^2}{n}\right], \tag{2.41}$$

where $A(n)$ depends on n alone.
Hint: This well known mathematical result can be obtained using the following approximations:

1. Approximate the natural logarithm of $\binom{n}{k}$ using the Stirling formula $\ln x! \approx x(\ln x - 1)$.
2. Set $k = \frac{n}{2} + \delta$ and assume $\delta \ll n$. Then approximate $\ln\left(\frac{n}{2} \pm \delta\right)$ using the *second* order Taylor decomposition.

As we can see from these exercises, an overwhelming majority of the superposition tree paths will have $k \approx \frac{n}{2}$, i.e., contain an approximately equal number of the H and V events, with the standard deviation scaling as the square root of the number of measurements [Fig. 2.6(a)]. The experience of observers in these paths is consistent with Born's rule. While, as discussed above, there exist "deviant" universes in which Born's rule does not hold, their number is unimaginably minuscule.

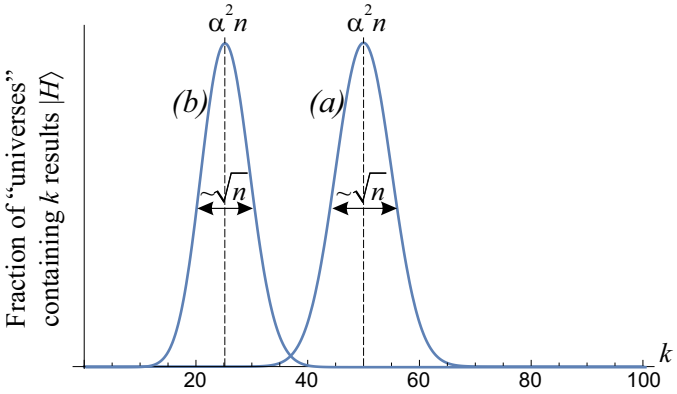


Fig. 2.6 Number of paths in the superposition tree containing k horizontal polarization detection events with $n = 100$ measurements for $\alpha^2 = \beta^2 = \frac{1}{2}$ (a) and $\alpha^2 = \frac{1}{4}$, $\beta^2 = \frac{3}{4}$ (b).

Let us now redo our derivation for a more complex setting. Suppose that the initial photon state is $|\psi\rangle = \alpha|H\rangle + \beta|V\rangle$, with amplitudes α and β not necessarily equal (we still assume them to be real, though)[22]. After the first measurement, the joint state of the photon and Alice is

$$|\Psi\rangle = \alpha|H\rangle \otimes |\text{Alice observed } H\rangle + \beta|V\rangle \otimes |\text{Alice observed } V\rangle . \qquad (2.42)$$

The superposition tree branches with subsequent measurements. However, the argument we developed above for the case $\alpha = \beta$ will not work because different branches will enter the superposition tree with different amplitudes. To address this issue, we modify the superposition tree in the following fashion.

[22] This derivation mainly follows the papers by W. H. Zurek, *Environment-Assisted Invariance, Entanglement, and Probabilities in Quantum Physics*, Physical Review Letters, **90**, 120404 (2003); *Probabilities from entanglement, Born's rule from invariance*, Physical Review A **71**, 052105 (2005).

Let us approximate

$$\alpha = \sqrt{\frac{m_H}{m_H + m_V}},\ \beta = \sqrt{\frac{m_V}{m_H + m_V}}, \tag{2.43}$$

where m_H and m_V are natural numbers. By choosing these values sufficiently high, we can approximate any real α and β with arbitrarily high precision. Alice is a complex quantum system, so her Hilbert space is highly multidimensional. Then, according to Ex. A.25, we can introduce a set of orthonormal states of Alice $\left\{\left|h_i^{(\ell)}\right\rangle\right\}_{i=1}^{m_H}$ and $\left\{\left|v_i^{(\ell)}\right\rangle\right\}_{i=1}^{m_V}$ such that

$$|\text{Alice observed } H \text{ in the } \ell\text{th measurement}\rangle = \frac{1}{\sqrt{m_H}} \sum_{i=1}^{m_H} \left|h_i^{(\ell)}\right\rangle; \tag{2.44}$$

$$|\text{Alice observed } V \text{ in the } \ell\text{th measurement}\rangle = \frac{1}{\sqrt{m_V}} \sum_{i=1}^{m_V} \left|v_i^{(\ell)}\right\rangle. \tag{2.45}$$

Substituting these decompositions along with Eq. (2.43) into Eq. (2.42), we find

$$|\Psi\rangle = \sqrt{\frac{1}{m_H + m_V}} \left(\sum_{i=1}^{m_H} |H\rangle \otimes \left|h_i^{(1)}\right\rangle + |V\rangle \otimes \sum_{i=1}^{m_V} \left|v_i^{(1)}\right\rangle \right). \tag{2.46}$$

This superposition has $m_H + m_V$ orthogonal terms, m_H of them corresponding to the observation of horizontal polarization, and m_V to vertical polarization. Subsequent measurement will result in further branching of the superposition tree, in such a way that there is a total of $(m_H + m_V)^n$ branches after n measurements [Fig. 2.5(b)]. Importantly, all branches now have equal amplitudes, so we can proceed in a similar fashion to the case of $\alpha = \beta$ studied previously.

Exercise 2.60. For a superposition state that is prepared after n measurements:

a) calculate the fraction of terms that contain k results $|H\rangle$ and $n-k$ results $|V\rangle$;
b) evaluate the above result numerically and plot it as a function of k for $n = 100$, $\alpha^2 = \frac{1}{4}$, $\beta^2 = \frac{3}{4}$;
c)* calculate the Gaussian approximation in the neighborhood of $k = \alpha^2 n$, akin to Ex. 2.59.

Answer:

a) $\binom{n}{k} m_H^k m_V^{n-k}$;
b) see Fig. 2.6(b);
c)

$$A(n) \exp\left[-\frac{\left(k - \alpha^2 n\right)^2}{2\alpha^2 \beta^2 n} \right]. \tag{2.47}$$

We see once again that Born's rule holds in the great majority of the worlds. We can conclude that the Measurement Postulate of quantum mechanics follows from the Hilbert space postulate and the unitarity of quantum evolution. Does this mean that we should now drop this postulate — given its redundancy and logical inconsistency?

Unfortunately, we cannot do so. The very example studied here shows how difficult it is — both computationally and psychologically — to use this approach for practical purposes. Effectively, we need to calculate the wavefunction of the universe every time we want to predict a measurement on a photon! If the goal is to make predictions for the phenomena experienced by finite observers such as humans, it makes much more sense to just assume that the wavefunction collapses — because this is indeed what happens in the subjective view of those observers. Then the Copenhagen interpretation is the tool of choice. Therefore, in the remainder of this book, we shall "shut up and calculate", only rarely referring to the many-worlds interpretation in order to get a broader perspective.

2.5 Quantum computation

The idea of quantum computation is to use quantum bits as basic units of information. In contrast to a classical bit, a qubit can be not only in a definite state $|0\rangle$ or $|1\rangle$, but also in a superposition of these states. Accordingly, multiple qubits can also be in superposition states, which are entangled with respect to individual qubit Hilbert spaces.

It is the entanglement that makes the quantum computer much more powerful than a classical one. Consider, for example, three qubits in a superposition

$$\begin{aligned} &a_{000}\,|000\rangle + a_{001}\,|001\rangle + a_{010}\,|010\rangle + a_{011}\,|011\rangle \\ &+ a_{100}\,|100\rangle + a_{101}\,|101\rangle + a_{110}\,|110\rangle + a_{111}\,|111\rangle\,. \end{aligned} \tag{2.48}$$

Performing a set of logical operations with these three qubits in this state, we simultaneously perform them with all $2^3 = 8$ sets of qubit values contained in state (2.48). In this way, we achieve an exponential degree of parallelism in our calculations. For example, even a tiny, 30-qubit quantum computer will work a billion ($2^{30} \approx 10^9$) times faster than its classical counterpart.

Of course, quantum computation is not as simple as it may appear from this example. Problems arise at both the theoretical and the practical level. Just to give one example of the many fundamental issues, let me mention the following. Suppose the quantum computer has completed its calculation on a superposition of qubit strings. Now we need to read out the answer. But the answers associated with each input string are also in a superposition state! If we now try to measure this state, the only thing we will obtain is one of the terms of the superposition. A systematic readout of a specific term associated with the input of interest is not possible.

Thus it turns out that the parallelism offered by quantum computers is useful for only a very limited class of problems. One such problem is the factorization of large numbers, which is known to be computationally difficult for classical computers and is thus used as the basis for public-key cryptography (Sec. 1.6). A technology for quickly deciphering public-key codes would pose a significant threat for society's information security. This is one of the reasons why quantum computation remains a subject of intense research.

Fortunately, this threat is not immediate because the quantum computer is very difficult to build. As we discussed in Sec. 2.4, any interaction of a quantum state with the environment makes the environment a part of the superposition. In the view of an observer who does not have control of that environment, this is equivalent to a superposition collapse. The likelihood of this happening is especially high for a multipartite entangled state, because interaction of *any* of the qubits with the environment will cause decoherence of the entire superposition.

This is one of the primary reasons why quantum computation technology has been developing so slowly. At the moment, we don't even know what physical system is best suited to carry quantum information. Different research groups around the world are investigating different systems — trapped atoms and ions, superconducting junctions, quantum dots, and even liquids — to determine their potential for this role. As it turns out, the photon is also a promising candidate. This is because the average energy of the optical photon (2–4 electronvolts) corresponds to a few tens of thousands of kelvins, i.e., much higher than the typical temperature of our environment. As a result, photons are not too likely to interact with this environment, and are hence robust against decoherence. In addition, it is easy to *encode* the qubit in the polarization of a photon: for example, the logical state $|0\rangle$ can correspond to the horizontal polarization and the state $|1\rangle$ to vertical.

It is also easy to implement single-qubit logical operations with this encoding. For example, we can perform the logical "not" operation using a $\lambda/2$ plate with its optic axis oriented at angle $3\pi/4$ to horizontal: state $|0\rangle$ ($|H\rangle$) will become $|1\rangle$ ($|V\rangle$) and vice versa (see Ex. 1.24). Another important operation is the Hadamard gate (see Ex. 1.27) with the matrix $\hat{H} \simeq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ which interconverts between the canonical and diagonal polarization states. The Hadamard gate is implemented using a half-waveplate oriented at angle $\pi/8$ to horizontal.

Additionally, in order to obtain the full range of computations accessible to a classical computer (Turing machine), we need *conditional* operations in which qubits could interact: the state of one qubit would affect the state of another. Theoretical research has shown that, in order to build a universal quantum computer, it is sufficient, in addition to single-qubit operations, to be able to implement only one kind of two-qubit operation: the *conditional "not" gate*, or the *c-not gate*. Implementation of this operation is a "holy grail" of quantum computation in any physical system. It is particularly complicated with photons.

The c-not gate involves two qubits: *control* and *target*. If the state of the control qubit is $|0\rangle$, the gate does not change the qubit values. But if the control qubit is

$|1\rangle$, the value of the target "flips": $|0\rangle$ becomes $|1\rangle$ and $|1\rangle$ becomes $|0\rangle$. This is summarized in Table 2.5[23].

Table 2.2 Truth table of the c-not gate.

input		output	
control	target	control	target
$\|0\rangle$	$\|0\rangle$	$\|0\rangle$	$\|0\rangle$
$\|0\rangle$	$\|1\rangle$	$\|0\rangle$	$\|1\rangle$
$\|1\rangle$	$\|0\rangle$	$\|1\rangle$	$\|1\rangle$
$\|1\rangle$	$\|1\rangle$	$\|1\rangle$	$\|0\rangle$

The c-not gate can be thought of as a "gremlin" that looks at the polarization of the control photon and, if the polarization is vertical, inserts a half-waveplate at 45° into the path of the target photon. The problem is that the gremlin must do this somehow without measuring the control photon, because such a measurement would entangle it with the qubits and collapse their quantum state (Sec. 2.4.1). As the following exercises show, this is theoretically possible.

Exercise 2.61. Write the matrices of operators corresponding to the following operations on a pair of qubits. The logical state $|0\rangle$ is encoded by the horizontal polarization and the logical state $|1\rangle$ by the vertical.

a) The c-not gate.
b) An operation that leaves states $|00\rangle$, $|01\rangle$, $|10\rangle$ unchanged, but multiplies the state $|11\rangle$ by a phase factor of -1 (*the conditional phase shift, or c-phase gate*).
c) A tensor product of the identity operator on the first qubit and the Hadamard gate upon the second (target) photon.

Are these operators unitary?

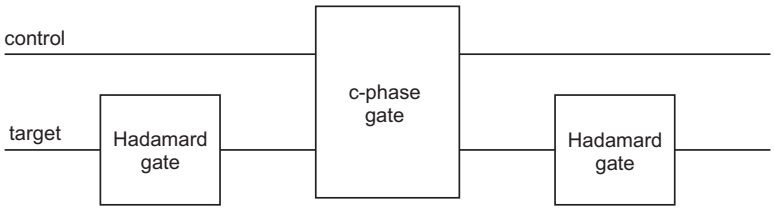


Fig. 2.7 Implementation of the c-not gate using a c-phase gate and two Hadamard gates.

[23] Note that the output value of the target qubit corresponds to the result of the classical exclusive-or (XOR) gate.

Exercise 2.62. Show that the c-not gate can be constructed by applying, in sequence, a Hadamard gate in Bob's space, a conditional phase shift, and a Hadamard gate in Bob's space again (Fig. 2.7).

Exercise 2.63. Show that the c-phase gate between two photons can be realized by the action of the Hamiltonian $\hat{H} = \hbar\omega |VV\rangle\langle VV|$ for a time π/ω.
Hint: other Hamiltonian eigenstates ($|HH\rangle$, $|HV\rangle$, and $|VH\rangle$) correspond to zero energy values.

Exercise 2.64. Show that the c-not gate constitutes the von Neumann measurement in the sense of Eq. (2.33) for $N = M = 2$ with $|w_0\rangle = |w_1\rangle$.

Exercise 2.62 shows us that, if we have a c-phase gate available, it can be used to construct the c-not gate. This does not solve the problem, but reduces it to a somewhat simpler one: rather than changing the values of the qubits, we only need to modify their phases. In application to photons, the implementation of the c-phase gate requires an optical element in which a photon would experience different phase shifts (i.e., different indices of refraction) depending on the polarization of another photon present in it. This is not something we would normally observe in optics: typically, if multiple light waves are present in the same medium, they will not interact, but propagate independently of the presence of the other waves. Situations involving the mutual influence of electromagnetic waves belong to the class of *nonlinear* optical phenomena. Nonlinear properties are observed in everyday media, such as glass or crystals, only when at least one of the fields is extremely strong, on the scale of trillions of photons. Making nonlinear optical effects significant at the optical intensity level of a few photons is a difficult problem and is currently being investigated by many research groups.

Exercise 2.65. Show that the operators from Ex. 2.61 (a,b) can create an entangled state from a separable one (*cf.* Ex. 2.17).

Exercise 2.66. Suppose you have available a c-not gate for photons. Propose a scheme that uses this gate to implement a measurement of two photons in the Bell basis.

2.6 Quantum teleportation and its applications

2.6.1 Quantum teleportation

Suppose Alice holds a single copy of a photon in some quantum state that she would like to transfer to Bob. However, she does not know the state of that photon, nor is there a direct quantum communication channel from Alice to Bob. Then it would appear that Alice's mission is impossible. Indeed, if she cannot send the photon to Bob directly, then the only option she is left with is to measure it. But, as we discussed in Sec. 1.4.2, a measurement on a single copy of a quantum state reveals very

little information about that state, and definitely not enough to recreate an exact copy elsewhere. Yet, as we shall see here, Alice can exploit the power of entanglement to transfer the state of her photon to Bob indirectly with a 100% probability and perfect fidelity.

In Sec. 2.2.1, we studied remote state preparation — a quantum communication protocol that allows the transfer of a quantum state from Alice to Bob by means of an entangled "resource" state shared between the two parties and a classical communication channel. In order to remotely prepare a desired state at Bob's station, Alice must know what state it is. By measuring her portion of the entangled resource in a basis chosen in accordance with that knowledge, Alice remotely converts Bob's portion of that resource into the desired state or the one orthogonal to it.

The quantum teleportation protocol is similar in some respects. However, in contrast to remote state preparation, Alice has no knowledge of the state she wishes to transfer to Bob. Instead, she possesses one copy of that state. It turns out that, by implementing a joint measurement on it and her share of the entangled resource, Alice can accomplish a similar goal: convert Bob's share of the resource into the desired state or one that can be transformed into it by means of a local operation[24].

So in contrast to science fiction, in which teleportation is relocation of an object, quantum teleportation is relocation of the *quantum state* of an object. While some may view it as less spectacular, I would argue that the phenomenon of quantum teleportation is not a bit less amazing. We know that, in order to *determine* an unknown quantum state, we must measure many copies of it in a variety of ways. Furthermore, we know that it is theoretically impossible to *clone* a quantum state, i.e., make a copy of it while keeping the original intact. Yet we can *recreate* a state at a remote location, while destroying the original one, and for that purpose only one copy of that state is needed.

The quantum teleportation protocol is shown schematically in Fig. 2.8. Alice has one copy of the input state $|\chi\rangle = \alpha|H\rangle + \beta|V\rangle$ in channel 1, associated with the Hilbert space $\mathbb{V}_1$; additionally, Alice and Bob share the entangled state $|\Psi^-\rangle$ in the Hilbert space $\mathbb{V}_2 \otimes \mathbb{V}_3$ encompassing channels 2 and 3. The following exercise explains, step by step, how teleportation works.

Exercise 2.67. a) Express the state $|\chi\rangle \otimes |\Psi^-\rangle$ in the canonical basis of $\mathbb{V}_1 \otimes \mathbb{V}_2 \otimes \mathbb{V}_3$.

b) Express the canonical basis states of $\mathbb{V}_1 \otimes \mathbb{V}_2$ in the Bell basis.

[24] The theoretical proposal for quantum teleportation was first published in C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, W. K. Wootters, *Teleporting an Unknown Quantum State via Dual Classical and Einstein–Podolsky–Rosen Channels*, Physical Review Letters **70**, 1895–1899 (1993). The first experiments, in different settings, were realized by several groups: D. Bouwmeester, J.-W. Pan, K. Mattle, M. Eibl, H. Weinfurter, A. Zeilinger, *Experimental Quantum Teleportation*, Nature **390**, 6660, 575–579 (1997); D. Boschi, S. Branca, F. De Martini, L. Hardy, S. Popescu, *Experimental Realization of Teleporting an Unknown Pure Quantum State via Dual classical and Einstein–Podolsky–Rosen channels*, Physical Review Letters **80**, 1121–1125 (1998); A. Furusawa, J. L. Sorensen, S. L. Braunstein, C. A. Fuchs, H. J. Kimble, E. S. Polzik, *Unconditional quantum teleportation*, Science 282, 706–709 (1998).

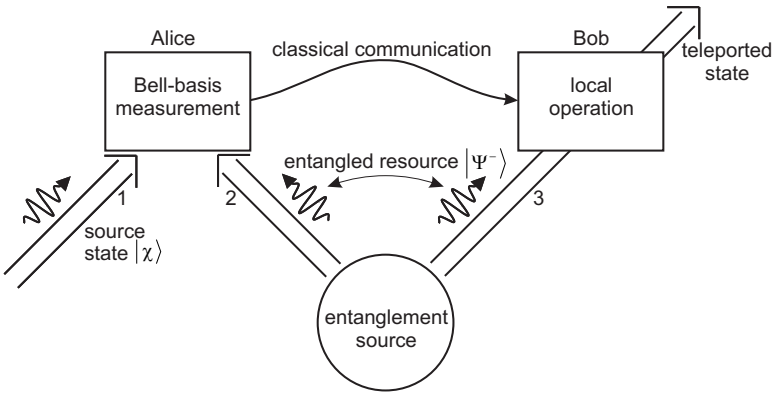


Fig. 2.8 Quantum teleportation

c) Express the state $|\chi\rangle \otimes |\Psi^-\rangle$ in the form of Eq. (2.15), i.e., as a linear combination of tensor products between Bell basis elements of $\mathbb{V}_1 \otimes \mathbb{V}_2$ and normalized states in $\mathbb{V}_3$.
d) Suppose Alice performs a local measurement on $\mathbb{V}_1 \otimes \mathbb{V}_2$ in the Bell basis. Calculate the probability of each measurement outcome and the state onto which the space $\mathbb{V}_3$ is projected.
e) Alice communicates her measurement result to Bob via the classical channel. Show that, with this information, Bob can convert the state of $\mathbb{V}_3$ into $|\chi\rangle$ via a local operation. Write this operation as an operator and propose implementation with waveplates.

Answer: see Table 2.3.

Table 2.3 Output states of quantum teleportation.

Bell state observed by Alice	Probability	State in Bob's channel	Bob's local operation
$\|\Phi^+\rangle$	1/4	$-\beta\|H\rangle + \alpha\|V\rangle$	$\hat{\sigma}_z\hat{\sigma}_x = \mathrm{i}\hat{\sigma}_y$
$\|\Phi^-\rangle$		$\beta\|H\rangle + \alpha\|V\rangle$	$\hat{\sigma}_x$
$\|\Psi^+\rangle$		$-\alpha\|H\rangle + \beta\|V\rangle$	$\hat{\sigma}_z$
$\|\Psi^-\rangle$		$-(\alpha\|H\rangle + \beta\|V\rangle)$	none

We see that, by receiving classical communication from Alice stating which Bell state she has detected and performing one of the Pauli operations on his photon, Bob will obtain a copy of the source state $|\chi\rangle$. The original source state is destroyed in Alice's measurement.

Importantly, the probabilities for each of Alice's measurement outputs are equal to $1/4$, independently of the parameters α and β of the source state. This means that Alice does not learn anything about the source state through her measurement. Neither does Bob (unless he chooses to measure his state at some point). This ignorance of both parties is a prerequisite for the perfect transfer of the source state. If we

were able to learn at least some information about a quantum state, while keeping its exact copy, we could use it for faster-than-light communication in a similar way to what was found in Ex. 2.44.

A prerequisite for implementing the teleportation protocol is a scheme for measuring two photons in the Bell basis. While such a measurement is theoretically feasible, in practice it is about as hard to implement as the c-not gate for photons (see Ex. 2.66). If only linear-optical tools are available, one can distinguish just two out of the Bell states. The latter approach is much easier to implement in practice, and has been used in most experiments on the teleportation of photon polarization qubits.

Exercise 2.68. Suppose a pair of photons in one of the Bell states enters the apparatus shown in Fig. 2.9. Show that:

a) if the input is in the state $|\Phi^+\rangle$, the detectors in the two gray boxes will simultaneously observe identical diagonally polarized photons (i.e., either both detectors 1 and 4 or both detectors 2 and 3 will click);
b) if the input is in the state $|\Phi^-\rangle$, the detectors in the two gray boxes will simultaneously observe orthogonal diagonally polarized photons (i.e., either both detectors 1 and 3 or both detectors 2 and 4 will click);
c) if the input is in the state $|\Psi^+\rangle$ or $|\Psi^-\rangle$, photon detection events will occur in only one of the two gray boxes.

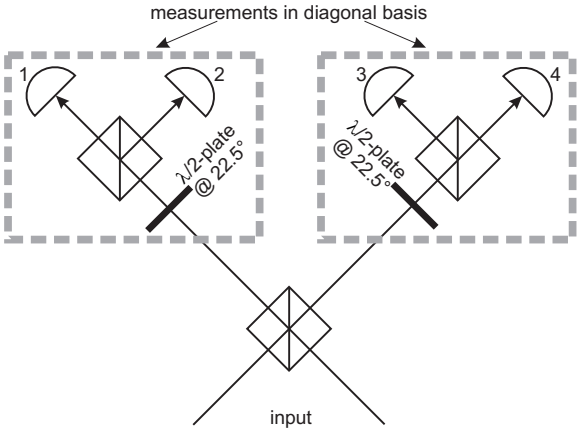


Fig. 2.9 A scheme for partial Bell-basis measurement. Squares denote polarizing beam splitters.

A protocol closely related to quantum teleportation is *entanglement swapping*[25]. It starts with four photons prepared in a pairwise entangled state $\left|\Psi_{12}^-\right\rangle \otimes \left|\Psi_{34}^-\right\rangle$. A

[25] Theoretical proposal: M. Żukowski, A. Zeilinger, M. A. Horne, and A. K. Ekert, *"Event-ready-detectors" Bell experiment via entanglement swapping*, Physical Review Letters **71**, 4287 (1993). Experiment: J.-W. Pan, D. Bouwmeester, H. Weinfurter, and A. Zeilinger, *Experimental Entang-*

Box 2.6 Can we teleport a human being?

Quantum physicists are sometimes asked how long it will be before we can teleport a person. Now you can answer this question. In order to teleport a quantum object, one requires two copies of it in a fully entangled state, i.e., a state that encompasses all possible quantum states of that object, in addition to the original one. So to teleport the *Star Trek* captain Picard from *USS Enterprise* to planet *Betazed*, we need first to make two exact copies of him, one on the ship and one on Betazed, and prepare them — that is, each pair of molecules in their respective bodies — in a fully entangled state!

measurement is then performed on photons in channels 2 and 3 in the Bell basis (Fig. 2.10). Through this action, the photon in channel 2 gets teleported into channel 4 (or, equivalently, the photon in channel 3 is teleported into channel 1). As a result, photons in channels 1 and 4 become entangled, even though they have never interacted with each other.

The following exercise provides a more rigorous analysis.

Exercise 2.69. A measurement is performed on channels 2 and 3 of state $|\Psi_{12}^-\rangle \otimes |\Psi_{34}^-\rangle$ in the Bell basis (Fig. 2.10). Determine the resulting state of channels 1 and 4 after each possible measurement outcome.

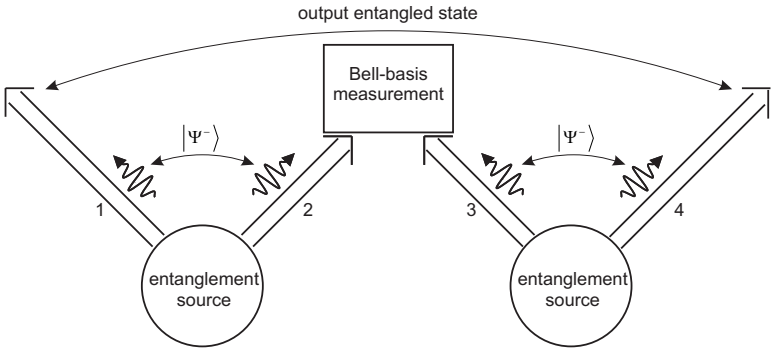


Fig. 2.10 Entanglement swapping.

lement Swapping: Entangling Photons That Never Interacted, Physical Review Letters **80**, 3891 (1998).

2.6.2 Quantum repeater

Both quantum teleportation and entanglement swapping find applications in quantum communications. In Chap.1 we learned that the primary issue preventing widespread practical application of quantum cryptography is losses in the transmission line. The exponential character of Beer's law, which governs the losses, leads to the transmissivity declining by many orders of magnitude on a scale of a few hundred kilometers, disabling quantum communications at any reasonable rate.

Of course, similar losses also occur in classical fiber optic communication lines. However, in the classical case the problem can be solved by means of the *repeater* — a device that receives the signal and retransmits it at a higher power. In quantum lines, classical repeaters are unusable because their action involves measurement of the incoming state. From the point of view of communicating parties, a classical repeater is indistinguishable from an eavesdropper.

In this section, we discuss the concept of the *quantum repeater*. While its principles are drastically different from those of its classical counterpart, the purpose is the same — to enhance the communication rate over a lossy line.

The first cornerstone of the quantum repeater is teleportation. If Alice and Bob share an entangled resource, Alice need not send the photon to Bob over a direct channel, but can teleport it to Bob. Because the Bell measurement can be performed at Alice's station, the source photon will travel a very short distance, hence experiencing negligible loss.

The problem of losses still arises, however, when we attempt to create the entangled resource required for teleportation and distribute it between Alice and Bob. The quantum repeater addresses this issue and permits quick and efficient distribution of entanglement over long distances.

This is shown schematically in Fig. 2.11. The repeater consists of multiple links, each covering the distance of a few dozen kilometers. Each link consists of two entanglement sources, an device for measuring pairs of photons in the Bell basis, and two quantum optical memory cells. The latter are devices that can store a quantum state of light for a relatively long time and subsequently retrieve it on demand.

Each entangled state source generates a pair of photons [Fig. 2.11(a)]. One of these photons propagates towards the Bell-state analyzer while the polarization state of the other is stored in memory. A Bell measurement on the central pair makes the stored states entangled thanks to entanglement swapping.

The sources are positioned in the neighborhood of the memory cells, so the loss for the photons that are to be stored is minimized. The photons that undergo the Bell measurement, on the other hand, have a significant chance of being lost, albeit much less than if they had to traverse the entire distance from Alice to Bob. Therefore multiple attempts may be required before entanglement swapping is successful. The length of the links is chosen on a scale of a few dozen kilometers — such that the expectation value for the number of attempts needed is reasonably short.

The significance of the quantum memory — the second cornerstone of the quantum repeater technology — is that once the entanglement within a link is created,

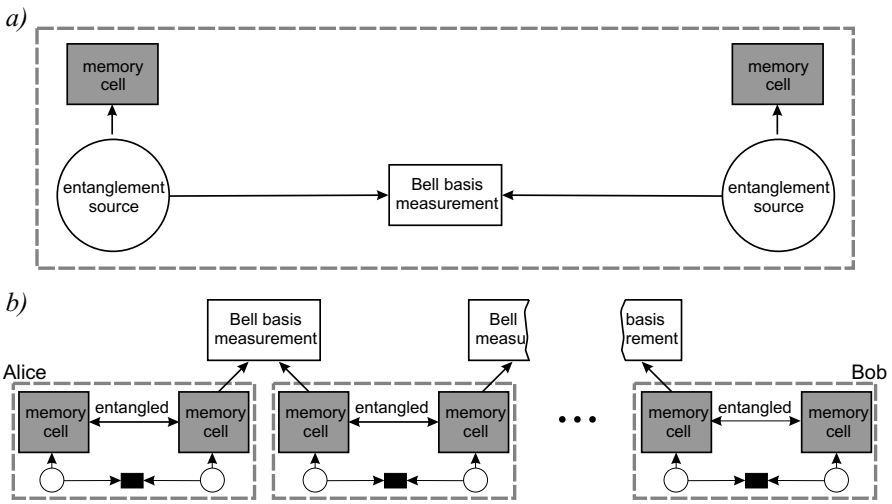


Fig. 2.11 Quantum repeater. a) An individual link creates entanglement between two memory cells. b) After all individual links have been prepared, entanglement swapping distributes entanglement through the links.

it can stay there for an extended period of time, until the entanglement is created in *all other* links.

Once this is accomplished, the procedure shown in Fig. 2.11(b) is performed. Photons are released from adjacent pairs of memory cells and subjected to Bell measurements. In this way, entanglement is swapped sequentially through the entire length of the communication line, resulting in Alice and Bob possessing a pair of entangled memory cells.

The advantage of the quantum repeater over direct transmission can be understood intuitively as follows. In order for direct transmission to be successful, the photon must not be lost anywhere in the fiber, and the probability of that happening is exponentially low. In the quantum repeater protocol, on the other hand, the loss in one of the links does not destroy the entanglement constructed in other links, so the probability of success falls off with distance at a much lower rate.

Exercise 2.70. A quantum repeater consists of two links. Each of the entanglement sources generates a state $|\Psi^-\rangle$. The Bell measurements in the first and second links detect states $|\Phi^+\rangle$ and $|\Phi^-\rangle$, respectively. A subsequent Bell measurement on the two adjacent memory cells of the two links detect $|\Psi^+\rangle$. What is the resulting joint state of the two memory cells adjacent to Alice and Bob?

Exercise 2.71. An $L = 500$ km quantum communication line between Alice and Bob consists of $k = 10$ quantum repeater links. The fiber loss coefficient $\beta = 0.05$ km^{-1}. The distance between each entangled source and the Bell basis analyzer

within each link is the same, viz., $L/2k = 25$ km. All entanglement sources generate photon pairs simultaneously, at a rate of $f = 10^6$ per second.

a) For a single link, find the probability of obtaining entanglement in its memory cells after a single attempt and after n attempts.
b) Find the probability of obtaining entanglement in all k links after n simultaneous attempts in each link.
c) Find the time t required to obtain entanglement in all links (and therefore entanglement between Alice's and Bob's cells) with a probability of at least $1/2$.
d) Instead of using a quantum repeater, Alice sends photons directly to Bob via a fiber line of length L km using a photon source with an emission rate of $f = 10^6$ photons per second. Find the time t' such that the probability that at least one of the photons sent by Alice during that period reaches Bob is $1/2$.

Assume the performance of the quantum optical memory cells and Bell basis measurements to be ideal.

We see that the quantum repeater offers a multiple order of magnitude advantage in comparison with direct transmission. However, practical realization of this device presents a challenge, associated primarily with the construction of high-performance quantum optical memory cells. This memory must be able to hold a quantum state for a long time and it must be possible to retrieve the stored state in a faithful and loss-free fashion. At the time of writing, quantum optical memory with performance characteristics suitable for use in quantum repeaters has not been achieved, but the field is developing quickly and further breakthroughs are frequently announced[26].

2.7 Problems

Problem 2.1. Modify the quantum dense coding protocol for the state that Alice and Bob pre-share being $|\Psi^+\rangle$, $|\Phi^+\rangle$, or $|\Phi^-\rangle$.

Problem 2.2. For the observable $\hat{\sigma}_z \otimes \hat{\sigma}_\theta$, where

$$\hat{\sigma}_\theta = |\theta\rangle\langle\theta| - \left|\frac{\pi}{2} + \theta\right\rangle\left\langle\frac{\pi}{2} + \theta\right|,$$

perform the following calculations.

a) Find the matrix in the canonical basis$\{|HH\rangle, |HV\rangle, |VH\rangle, |VV\rangle\}$.
b) Find the matrix in the Bell basis.
c) Determine the eigenstates and eigenvalues.
Hint: you need not solve any equations.
d) Calculate the expectation value and uncertainty in the Bell state $|\Psi^-\rangle$.

[26] A. I. Lvovsky, B. C. Sanders, and W. Tittel, *Optical Quantum Memory*, Nature Photonics **3**, 706–714 (2009); N. Sangouard, C. Simon, H. De Riedmatten, and N. Gisin, *Quantum repeaters based on atomic ensembles and linear optics*, Reviews of Modern Physics **83**, 3380 (2011).

Problem 2.3. Two qubits interact according to the Hamiltonian

$$\hat{H} = \hbar\omega\hat{\sigma}_x \otimes \hat{\sigma}_x.$$

The initial state of the qubits is $|\Psi(0)\rangle = |HH\rangle$. Find $|\Psi(t)\rangle$ in the canonical basis.

Problem 2.4. The tensor product Hilbert space of Alice's and Bob's photons evolves according to the Hamiltonian

$$\hat{H} = \hbar\omega(\hat{\sigma}_x \otimes \hat{\sigma}_x + \hat{\sigma}_y \otimes \hat{\sigma}_y + \hat{\sigma}_z \otimes \hat{\sigma}_z).$$

a) Find the 4×4 matrix of the Hamiltonian in the canonical basis.
b) Find the matrix of the evolution operator $\mathrm{e}^{-\mathrm{i}Ht/\hbar}$.
c) What is the final state of the system after the period $\omega t = \pi/4$ if the initial state is an arbitrary separable state $(a|H\rangle + b|V\rangle) \otimes (c|H\rangle + d|V\rangle)$?

Problem 2.5. The Greenberger–Horne–Zeilinger state $|\Psi_{GHZ}\rangle = \frac{1}{\sqrt{2}}(|HHH\rangle + |VVV\rangle)$ is distributed among Alice, Bob, and Charley. Rewrite $|\Psi_{GHZ}\rangle$:

- in the basis that is canonical in Alice's Hilbert space, diagonal in Bob's Hilbert space, and circular in Charley's Hilbert space;
- in the Bell basis in the Hilbert space of Alice and Bob and canonical in Charley's Hilbert space.

Problem 2.6. Alice and Bob share two photons in the polarization state

$$|\Psi\rangle = \frac{1}{\sqrt{11}}(|HH\rangle + \mathrm{i}|VH\rangle + 3|VV\rangle).$$

a) Alice and Bob each perform measurements on their respective photons. Find the probabilities of all possible results.
b) Only Alice performs a polarization measurement on her photon. Find the probability of each outcome and the remotely prepared state of Bob's photon after the measurement. Apply each of the two alternative techniques to solve the problem in each basis:
   - using the partial inner product;
   - decomposing the initial state according to Eq. (2.15).
c) Suppose Bob does not know Alice's result. Based on part (b), describe the state of Bob's photon after Alice's measurement as an ensemble.
d) Check that the probability values found in parts (a) and (b) are mutually consistent.

Solve this problem for all measurements performed in the (i) canonical and (ii) circular bases.

Problem 2.7. Alice and Bob perform measurements on multiple copies of some bipartite state $|\Psi\rangle$ and find the following:

- if Alice measures in the diagonal basis:
  - whenever Alice detects $|+\rangle$, Bob gets $|H\rangle$;
  - whenever Alice detects $|-\rangle$, Bob gets $|V\rangle$.
- if Alice measures in the canonical basis:
  - whenever Alice detects $|H\rangle$, Bob gets $|L\rangle$;
  - whenever Alice detects $|V\rangle$, Bob gets $|R\rangle$.

What is $|\Psi\rangle$?

Problem 2.8. The Greenberger–Horne–Zeilinger state is distributed among Alice, Bob, and Charley. Alice and Bob perform a joint measurement on $|\Psi_{GHZ}\rangle$. What is the probability for them to detect

a) $|\Psi^-\rangle$,
b) $|HR\rangle$,
c) $|\Theta\rangle = (3|HH\rangle + 4|VV\rangle)/5$.

and onto which state will Charley's particle project? For each of the above states, assume any measurement basis that contains the state in question.

Problem 2.9. Alice, Bob, and Charley share an entangled state of three photons

$$|\Psi\rangle = (3|+-+\rangle + 4|-++\rangle)/5. \tag{2.49}$$

Alice and Bob measure their photons in the canonical basis. Alice detects a horizontal polarization, and Bob vertical.

a) What is the probability of this event?
b) Onto which state will Charley's photon project?

Problem 2.10. Modify observables $\hat{M}_A$, $\hat{M}_B$, $\hat{N}_A$, $\hat{N}_A$ as necessary to violate the Bell inequality for the state produced when the source is $|\Psi^+\rangle$, $|\Phi^+\rangle$, or $|\Phi^-\rangle$.

Problem 2.11. Reproduce the Greenberger–Horne–Zeilinger argument for

$$\left|\Psi'_{GHZ}\right\rangle = \frac{1}{2}(|HHH\rangle + |HVV\rangle + |VVH\rangle + |VHV\rangle)$$

and operators

$$\begin{array}{c} \hat{\sigma}_z \otimes \hat{\sigma}_y \otimes \hat{\sigma}_y \\ \hat{\sigma}_y \otimes \hat{\sigma}_z \otimes \hat{\sigma}_y \\ \hat{\sigma}_y \otimes \hat{\sigma}_y \otimes \hat{\sigma}_z \\ \hline \hat{\sigma}_z \otimes \hat{\sigma}_z \otimes \hat{\sigma}_z \end{array}$$

Problem 2.12. A von Neumann measurement of the photon polarization state $|\psi\rangle = \alpha|H\rangle + \beta|V\rangle$ is performed in the diagonal basis.

a) Write the joint state of the system and the apparatus after the measurement in the measurement basis.
b) Give an ensemble description of the state of the system alone after the measurement.

Problem 2.13. A photon is initially in the state $|\psi\rangle = (3|H\rangle + 4|V\rangle)/5$. Describe, in the form of an ensemble, the photon's state after it has decohered in either of the following decoherence-preferred bases:

a) canonical;
b) circular.

Problem 2.14. An atom has two energy eigenstates $|v_1\rangle, |v_2\rangle$ with eigenvalues 0 and $3\hbar\omega$, respectively, where $\omega > 0$. The atom is initially in the state $|v_1\rangle$. At time $t = 0$, a field is turned on which makes the Hamiltonian equal to $\hat{H} = \hat{H}_0 + \hat{V}$ with $\hat{V} = 2i\hbar\omega|v_1\rangle\langle v_2| - 2i\hbar\omega|v_2\rangle\langle v_1|$. The atom experiences decoherence with the eigenbasis of the new Hamiltonian being the decoherence-preferred basis. Write the ensemble defining the atom's state after it has decohered.

Problem 2.15. A Bell inequality test, as described in Sec. 2.3, is performed with a defective entangled source which produces a statistical mixture of the state $|\Psi^-\rangle$ with probability η and $|\Psi^+\rangle$ with probability $1-\eta$. What is the range of η values for which the Bell inequality is violated?

Problem 2.16. Show that teleportation will work with other Bell states as the entangled resource. For each Bell state, determine the local operations Bob would need to perform on $\mathbb{V}_3$ after receiving classical communication from Alice.

Problem 2.17. The quantum teleportation protocol is implemented with state $|\Psi\rangle = (|HV\rangle - 2|VH\rangle)/\sqrt{5}$ as the entangled resource, instead of $|\Psi^-\rangle$. Alice's input state is $|\chi\rangle = \alpha|H\rangle + \beta|V\rangle$. Determine:

a) the state in which Bob's photon will be prepared in the event of each of the four outcomes of Alice's Bell measurement;
b) the probability of each outcome.

Problem 2.18*. In the quantum repeater described in Ex. 2.71, one of the following imperfections is present:

a) the Bell basis measurement apparatus is only able to detect states $|\Psi^\pm\rangle$, but not $|\Phi^\pm\rangle$;
b) for each photon stored in quantum memory, the retrieval efficiency is $\eta_M = 0.75$.

For each case, find the new time t required to obtain entanglement between Alice's and Bob's memory cells with a probability of at least $1/2$.

Chapter 3
One-dimensional motion

We are now ready to put the "mechanics" into quantum mechanics. In this chapter, we shall study basic quantum physics of the simplest motional system: a point-like particle with a single degree of freedom. While it may sound a bit like a "spherical horse in vacuum", this model turns out to be quite relevant to many practical physical settings, describing their properties surprisingly well. Moreover, the quantum theory of one-dimensional motion will provide us with theoretical tools to study more complex, three-dimensional, motion in the next chapter. This theory can be directly applied to the motion of electrons in atoms to calculate, for example, atomic emission and absorption spectra. These spectra can then be compared with those measured experimentally, thereby providing the basis for a confirmation or refutation of the quantum theory. The remarkably good agreement found in this comparison was the main factor in the triumph of quantum theory in the early 20th century.

3.1 Continuous observables

In classical mechanics, one-dimensional motion is described by two canonical variables, position and momentum. Accordingly, in our quantum treatment, we introduce two observable operators: position $\hat{x}$ and momentum $\hat{p}$[1].

Although the *geometric* space containing the particle is one-dimensional, the associated *Hilbert* space is of infinite dimension: there are infinitely many position eigenstates $|x\rangle$, and all these eigenstates are orthogonal[2]. Furthermore, position eigenstates form a *continuum*: for every real value of x there exists an associated eigenstate $|x\rangle$. The same is true for the momentum observable.

[1] If you are not familiar with the Dirac delta function and the Fourier transformation, please review Sections D.1 and D.2 in appendix before proceeding.

[2] Why does the continuum of position eigenstates span a Hilbert space of infinite dimension, while the continuum of linearly polarized states only a two-dimensional Hilbert space? If you do not remember the answer, refer to Sec. 1.3.

A. I. Lvovsky, *Quantum Physics*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-662-56584-1_3

We know (see Ex. 1.30) that the set of eigenstates of any physical observable forms an orthonormal basis. Position and momentum are no exception. However, the continuous nature of these observables implies that most mathematical rules (state and operator decomposition, normalization, basis conversion, etc.) derived for finite-dimensional Hilbert spaces have to be modified: summation must be replaced by integration. This is our task in this section. In order to reproduce these rules in the form that closely resembles those for the discrete case, we need to define a special normalization convention for continuous observable eigenstates. Instead of normalizing these states to one, as we would do in the discrete case, we write:

$$\langle x|\, x'\rangle = \delta(x-x'); \tag{3.1a}$$

$$\langle p|\, p'\rangle = \delta(p-p'). \tag{3.1b}$$

This may appear strange at first. According to Eq. (3.1a), the inner product of the position eigenstate $|x\rangle$ with itself is $\langle x|\, x\rangle = \delta(0)$, so this state has infinite norm. How is this consistent with the Hilbert space postulate of quantum mechanics, which says that all physical states must have norm 1? We answer this question by saying that *continuous-observable eigenstates are unphysical*: it is impossible to set a particle at an absolutely precise location or make it move at an absolutely precise velocity. Therefore, the normalization rule for physical states does not apply to $|x\rangle$ or $|p\rangle$; these states are just a mathematical abstraction[3]. All physically realistic states, which have some uncertainty both in the position and momentum, do have norm one in accordance with the Postulate.

Any quantum state $|\psi\rangle$ can be decomposed in the basis associated with a continuous-variable observable according to

$$|\psi\rangle = \int_{-\infty}^{+\infty} \psi(x)\,|x\rangle\,\mathrm{d}x. \tag{3.2}$$

This equation replaces Eq. (A.1) for the decomposition of a state into a discrete basis: the sum is replaced by an integral. The function $\psi(x)$ is called the *wavefunction* of the state $|\psi\rangle$ in the x-basis (-representation) and is the continuous-observable analog of the column representation of a vector in a Hilbert space of finite dimension. Taking the adjoint of both sides of Eq. (3.2), viz.,

$$\langle\psi| = \int_{-\infty}^{+\infty} \psi^*(x)\,\langle x|\,\mathrm{d}x, \tag{3.3}$$

we also find that the wavefunction of $\langle\psi|$ is $\psi^*(x)$.

[3] To treat this matter more rigorously, one introduces a special construction called *rigged Hilbert space*. See R. de la Madrid, *The role of the rigged Hilbert space in quantum mechanics*, European Journal of Physics **26**, 287 (2005) for details.

Box 3.1 What happens for a finite-dimensional normalization rule?

What if we wish to avoid using generalized functions and try to apply the finite-dimensional normalization rules to the continuous-variable Hilbert space? Unfortunately, it is then impossible to develop a consistent set of relations among states, wavefunctions, and observables. For example, let us set

$$\langle x|\, x'\rangle = \begin{cases} 1 \text{ if } x = x' \\ 0 \text{ if } x \neq x' \end{cases} . \tag{3.7}$$

Then by substituting Eq. (3.2) into Eq. (3.4), we would have

$$\psi(x) = \langle x|\, \psi\rangle = \int_{-\infty}^{+\infty} \psi(x')\, \langle x|\, x'\rangle \, \mathrm{d}x'.$$

The last expression in the line above contains an integral of a function that has a non-vanishing finite value at just one point $x' = x$, and hence vanishes. So under the assumption (3.7), the wave functions of all physical states would be zero.

Exercise 3.1. Show that we can construct the following continuous analogues of the major discrete-case relations:

a) instead of Eq. (A.6):

$$\psi(x) = \langle x|\, \psi\rangle\,; \tag{3.4}$$

b) instead of Eq. (A.26):

$$\int_{-\infty}^{+\infty} |x\rangle\langle x|\, \mathrm{d}x = \hat{\mathbf{1}}; \tag{3.5}$$

c) instead of Eq. (A.4):

$$\langle \psi_1|\, \psi_2\rangle = \int_{-\infty}^{+\infty} \psi_1^*(x)\psi_2(x)\mathrm{d}x. \tag{3.6}$$

Exercise 3.2. Show that, for physical states,

$$\int_{-\infty}^{+\infty} |\psi(x)|^2 \mathrm{d}x = 1. \tag{3.8}$$

Exercise 3.3. Calculate the normalization factor A for states with the following wavefunctions:

a) a "top-hat function"

$$\psi(x) = \begin{cases} 0 \text{ if } x < a \text{ or } x > b \\ A \text{ if } a \leq x \leq b \end{cases} ; \tag{3.9}$$

b) a Gaussian wavefunction

$$\psi(x) = Ae^{-\frac{x^2}{2d^2}}. \tag{3.10}$$

Exercise 3.4. Find the wavefunction of the state of definite position $|x_0\rangle$ in the position basis.

As in the discrete case, operators associated with continuous observables are given by

$$\hat{x} = \int_{-\infty}^{+\infty} x\,|x\rangle\langle x|\,\mathrm{d}x. \tag{3.11}$$

The operator functions are naturally defined by

$$f(\hat{x}) = \int_{-\infty}^{+\infty} f(x)\,|x\rangle\langle x|\,\mathrm{d}x. \tag{3.12}$$

For an arbitrary operator $\hat{A}$, the two-dimensional function

$$A(x,x') = \langle x|\hat{A}\,|x'\rangle, \tag{3.13}$$

is referred to as the operator's *matrix element*.

As we shall see below and similarly to the discrete-variable case, as a function of x and x', the matrix element $\langle x|\hat{A}\,|x'\rangle$ contains complete information about the operator. More generally, we can perform operations with states and operators represented by one- and two-dimensional functions, respectively, just as we operate with matrices in the discrete case, only replacing summation with integration.

Exercise 3.5. Show that $\hat{x}\,|x\rangle = x\,|x\rangle$.

Exercise 3.6. Prove that:

a) any operator $\hat{A}$ can be written in the form

$$\hat{A} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} A(x,x')\,|x\rangle\langle x'|\,\mathrm{d}x\,\mathrm{d}x', \tag{3.14}$$

where $A(x,x')$ is given by Eq. (3.13);

b) for an operator function of $\hat{x}$,

$$\langle\psi|\,f(\hat{x})\,|\psi\rangle = \int_{-\infty}^{+\infty} |\psi(x)|^2 f(x)\mathrm{d}x; \tag{3.15}$$

c) for any operator $\hat{A}$ and any two states $|\psi\rangle$, $|\phi\rangle$,

$$\langle \phi | \hat{A} | \psi \rangle = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \phi^*(x) A(x,x') \psi(x') \mathrm{d}x \mathrm{d}x'; \tag{3.16}$$

d) the wavefunction of the state $\hat{A}|\psi\rangle$ is

$$\langle x | \hat{A} | \psi \rangle = \int_{-\infty}^{+\infty} A(x,x') \psi(x') \mathrm{d}x'; \tag{3.17}$$

e) the wavefunction of the state $\langle \psi | \hat{A}$ is

$$\langle \psi | \hat{A} | x \rangle = \int_{-\infty}^{+\infty} \psi^*(x') A(x',x) \mathrm{d}x'; \tag{3.18}$$

f) the matrix elements of an operator $\hat{A}$ and its adjoint $\hat{A}^\dagger$ are related by

$$(A^\dagger)(x,x') = A^*(x',x); \tag{3.19}$$

g) the product of operators $\hat{A}$ and $\hat{B}$ can be written in terms of their "matrices" as

$$\langle x | \hat{A}\hat{B} | x' \rangle = \int_{-\infty}^{+\infty} A(x,x'') B(x'',x') \mathrm{d}x''. \tag{3.20}$$

Let us now reformulate the Measurement Postulate of quantum mechanics for the continuous-observable case. Suppose the observable $\hat{x}$ is measured in the quantum state $|\psi\rangle$ with wavefunction $\langle x | \psi \rangle = \psi(x)$. What is the probability distribution for the possible measurement results? In Sec. B.4, we introduced the continuous-variable probability density $\mathrm{pr}(x)$ such that the probability of observing x in a certain interval $[x',x'']$ is

$$\mathrm{pr}_{[x',x'']} = \int_{x'}^{x''} \mathrm{pr}(x) \mathrm{d}x. \tag{3.21}$$

Let us express $\mathrm{pr}(x)$ in terms of $\psi(x)$.

According to the Measurement Postulate for the discrete case, the probability of projecting onto a specific element $|v_i\rangle$ of the measurement basis is $|\langle v_i | \psi \rangle|^2$. We cannot make exactly the same statement for the continuous case, because the probability of observing the particle precisely at position x is infinitesimal. It is, however, reasonable to say that the legitimate probability measure associated with x — its probability density — must be proportional to $|\langle x | \psi \rangle|^2 = |\psi(x)|^2$. So we have $\mathrm{pr}(x) \propto |\psi(x)|^2$.

To find the proportionality coefficient, we recall to begin with that $\int_{-\infty}^{+\infty} \mathrm{pr}(x) \mathrm{d}x = 1$ according to the properties of the probability density [*cf.* Eq. (B.12)]. On the ot-

her hand, we also have, for a normalized state, $\int_{-\infty}^{+\infty} |\psi(x)|^2 \mathrm{d}x = \langle \psi | \psi \rangle = 1$ as per Eq. (3.6). Comparing these two conditions, we find

$$\mathrm{pr}(x) = |\psi(x)|^2. \tag{3.22}$$

Which state will $|\psi\rangle$ project onto after the measurement? As already discussed, the obvious answer $|x\rangle$ is unphysical. Yet it is useful as an approximation for many theoretical applications, as long as we do not forget to take the normalization issue into account. The more physically realistic answer will depend on the specifics of the measurement apparatus; generally, one would obtain some superposition or statistical mixture of multiple position eigenstates within a certain narrow neighborhood of x.

Exercise 3.7. Using the expressions (B.13) and (B.14) for the mean and variance of a continuous random variable, show that, for the continuous quantum observable $\hat{x}$ measured in the state $|\psi\rangle$:

a) the expectation value is given by

$$\langle x \rangle = \langle \psi | \hat{x} | \psi \rangle ; \tag{3.23}$$

b)[§] the variance is given by

$$\left\langle \Delta x^2 \right\rangle = \left\langle \psi \right| \hat{x} - \langle x \rangle)^2 \left| \psi \right\rangle = \left\langle \psi \right| \hat{x}^2 \left| \psi \right\rangle - \langle \psi | \hat{x} | \psi \rangle^2 . \tag{3.24}$$

The findings of this section are summarized in Table 3.1.

Table 3.1 Comparative summary of the rules for working with discrete- and continuous-variable bases.

	discrete basis $\{\|v_i\rangle\}$	continuous basis $\{\|x\rangle\}$
orthonormality	$\langle v_i \| v_j \rangle = \delta_{ij}$	$\langle x \| x' \rangle = \delta(x-x')$
decomposition of a state	$\|\psi\rangle = \sum_i \psi_i \|v_i\rangle$ $\psi(x) = \langle x \| \psi \rangle$	$\|\psi\rangle = \int_{-\infty}^{+\infty} \psi(x) \|x\rangle$ $\psi_i = \langle v_i \| \psi \rangle$
Measurement Postulate	$\mathrm{pr}_i = \| \langle v_i \| \psi \rangle \|^2$ (probability)	$\mathrm{pr}(x) = \| \langle x \| \psi \rangle \|^2$ (probability density)
decomposition of an operator	$A_{ij} = \langle v_i \| \hat{A} \| v_j \rangle$ $\hat{A} = \sum_{i,j} A_{ij} \|v_i\rangle\langle v_j\|$	$A(x,x') = \langle x \| \hat{A} \| x' \rangle$ $\hat{A} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} A(x,x') \|x\rangle\langle x'\| \mathrm{d}x\mathrm{d}x'$
decomposition of $\hat{\mathbf{1}}$	$\hat{\mathbf{1}} = \sum_i \|v_i\rangle\langle v_i\|$	$\hat{\mathbf{1}} = \int_{-\infty}^{+\infty} \|x\rangle\langle x\| \mathrm{d}x$
product of operators	$(AB)_{ij} = \sum_k A_{ik} B_{kj}$	$(AB)(x,x') = \int_{-\infty}^{+\infty} A(x,x'') B(x'',x') \mathrm{d}x''$

3.2 De Broglie wave

In the previous section, we developed the mathematical machinery for handling Hilbert spaces spanned by eigenstates of a continuous observable, such as position or momentum. But in fact, position and momentum are operators within the same physical Hilbert space associated with a particle's motion. We bring these two observables together by postulating the relation between their eigenstates:

$$\langle x|\ p\rangle = \frac{1}{\sqrt{2\pi\hbar}} \mathrm{e}^{\mathrm{i}\frac{px}{\hbar}}. \tag{3.25}$$

The relation (3.25) states that the wavefunction of the state with a definite momentum is an infinite wave which is known as the *de Broglie wave*. This wave is a manifestation of the wave–particle duality, i.e., the property of all quantum matter to exhibit both particle and wave features (*cf.* Sec. 1.5).

The de Broglie wave cannot be derived from the quantum mechanics postulates we have studied so far. Rather, it is a generalization of a multitude of experimental observations and theoretical insights. The historical path towards the de Broglie wave is briefly outlined in Box 3.2.

You may be surprised that Eq. (3.25) contains no time dependence, even though the very notion of the wave implies that such dependence must be present. We will indeed recover the wave motion by applying the Schrödinger equation in Sec. 3.4. For now, however, let us abstract from that motion and study the relations between the bases formed by the position and momentum eigenstates, which we define to be time-independent.

Exercise 3.8. Show that the wavelength of the de Broglie wave given by Eq. (3.25) is related to the momentum according to

$$\lambda_{dB} = \frac{2\pi\hbar}{p}, \tag{3.26}$$

i.e., in the same way as the momentum of the photon and the optical wavelength (Box 1.1).

Exercise 3.9. Estimate the de Broglie wavelength for

a) a car;
b) air molecules at room temperature;
c) electrons in an electron microscope with a kinetic energy of 100 keV;
d) rubidium atoms in a Bose–Einstein condensate at a temperature of 100 nanokelvin.

Exercise 3.10. Show that, according to Eq. (3.25), the position and momentum eigenstates can be expressed in terms of one another as follows:

Box 3.2 History of de Broglie's discovery

Niels Bohr

In 1913, Niels Bohr used Planck's concept to develop his model of the atom, according to which, the electron's orbital is stable if its angular momentum equals an integer multiple of $\hbar$. However, Bohr's model was purely empirical. Although it seemed to explain experimental results, the physics behind it remained a mystery.

Louis de Broglie proposed the concept of his wave in his 1924 PhD thesis. By then, Planck and Einstein had already found the relations between the photon's wavelength, frequency, energy, and momentum, and Compton had confirmed them experimentally (Box 1.1). De Broglie hypothesized that the relation $E = \hbar\omega$ did not have to be limited to light particles. Rather, *any* particle with a certain energy can be associated with a wave whose frequency is given by Planck's formula. Then de Broglie used Einstein's special relativity theory to show that the wavelength of that wave had to be given by Eq. (3.26), i.e., the same expression as for the photon.

Louis de Broglie

De Broglie used his assumption to reformulate Niels Bohr's model of the atom (Box 4.2). He hypothesized that the electron orbital is stable if its circumference contains an integer number n of de Broglie wavelengths:

$$2\pi r = n\lambda_{dB}, \tag{3.28}$$

where r is the radius of the orbit. In this way, the matter wave associated with the orbiting electron experiences constructive interference with itself, thereby forming a standing wave. This hypothesis led him to a theoretical prediction of atomic spectra that was identical to Bohr's (Ex. 4.42) and consistent with the experimental data.

This agreement constituted strong evidence in favor of de Broglie's hypothesis. Even more direct evidence was obtained at the Bell Labs in 1927. Clinton Davisson and Lester Germer observed the scattering of a flux of electrons on the crystalline lattice of nickel and found the measured angular distribution of the scattered electrons to be consistent with the laws of diffraction known from optics*. The wave-like nature of the electrons was the only possible explanation for this behavior.

* C. J. Davisson, *The Diffraction of Electrons by a Crystal of Nickel*. Bell System Technical Journal **7**, 90 (1928).

$$|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} e^{i\frac{px}{\hbar}} |x\rangle \, dx; \tag{3.27a}$$

$$|x\rangle = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} e^{-i\frac{px}{\hbar}} |p\rangle \, dp. \tag{3.27b}$$

The de Broglie wave has infinite extent in space. This is consistent with the uncertainty principle: the wavefunction of a state of definite momentum has an infinite position uncertainty. However, this becomes absurd if we interpret the square of the absolute value of the de Broglie wavefunction, a constant $|\langle x|\, p\rangle|^2 = 1/2\pi\hbar$, as a probability density. Integrating this over the entire space, we obtain infinity.

This, again, is a consequence of the momentum eigenstate being unphysical, which means that its probability density is meaningless. In practice, realistic states are linear combinations of momentum eigenstates so their position uncertainty can be limited. We will soon see this in more detail, when we discuss Gaussian wavepackets.

The de Broglie argument explains the exponential in Eq. (3.25), but not the normalization factor. The following exercise shows how to understand this.

Exercise 3.11. Expressing two arbitrary momentum eigenstates $|p\rangle$ and $|p\rangle'$ as de Broglie waves according to Eq. (3.27a) and using $\langle x|\, x'\rangle = \delta(x-x')$, calculate $\langle p|\, p'\rangle$ and check that your result is consistent with the orthonormality condition $\langle p|\, p'\rangle = \delta(p-p')$.

The *wavenumber* of the de Broglie wave is

$$k = \frac{2\pi}{\lambda_{dB}} = \frac{p}{\hbar}. \tag{3.29}$$

Sometimes it is convenient to handle momentum eigenstates $|p\rangle$ in the physically equivalent form of wavenumber eigenstates $|k = p/\hbar\rangle$, because then we need not worry about the Planck constant in the exponent.

There is a subtlety though. As for any continuous observable, the wavenumber eigenstates are normalized according to

$$\langle k|\, k'\rangle = \delta(k-k'). \tag{3.30}$$

But, as we know from Eq. (D.6), $\delta(k-k') = \delta[(p-p')/\hbar] = \hbar\delta(p-p') = \hbar\langle p|\, p'\rangle$. We are compelled to conclude that

$$|k\rangle = \sqrt{\hbar}\,|p = \hbar k\rangle\,. \tag{3.31}$$

This is yet another seemingly absurd result: two vectors that represent the same state — a state with a certain momentum — have a different norm. Once again, it is a consequence of the unphysical character of normalization for continuous observable eigenstates.

Exercise 3.12.§ Show that the de Broglie wavefunction for the wavenumber eigenstate takes the form

$$\langle x|\, k\rangle = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{\mathrm{i}kx}. \tag{3.32}$$

Show that the position and wavenumber eigenstates are expressed in terms of one another according to

$$|k\rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{e}^{\mathrm{i}kx} |x\rangle \, \mathrm{d}x; \tag{3.33}$$

$$|x\rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{e}^{-\mathrm{i}kx} |k\rangle \, \mathrm{d}k. \tag{3.34}$$

Check consistency with the normalization condition (3.30).

3.3 Position and momentum bases

3.3.1 Conversion between position and momentum bases

Here we discuss the problem of converting the representations of various states and operators between the position and momentum bases. As in the discrete case, the primary tool for such conversion is the resolution of the identity, i.e., we exploit the fact that the operator (3.5)

$$\hat{\mathbf{1}} = \int_{-\infty}^{+\infty} |x\rangle\langle x| \, \mathrm{d}x = \int_{-\infty}^{+\infty} |p\rangle\langle p| \, \mathrm{d}p \tag{3.35}$$

can be inserted into any inner product expression.

Exercise 3.13. Find explicit formulae for converting between the position $\psi(x)$ and momentum $\tilde{\psi}(p)$ representations of a given quantum state $|\psi\rangle$.
Answer:

$$\psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \tilde{\psi}(p) \mathrm{e}^{\mathrm{i}\frac{px}{\hbar}} \mathrm{d}p; \tag{3.36a}$$

$$\tilde{\psi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \psi(x) \mathrm{e}^{-\mathrm{i}\frac{px}{\hbar}} \mathrm{d}x \tag{3.36b}$$

Exercise 3.14.§ Show that conversion between the wavefunctions in the position and *wavenumber* representations and back simply corresponds to the direct and inverse Fourier transformation, respectively:

$$\psi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \tilde{\psi}(k) \mathrm{e}^{\mathrm{i}kx} dk; \tag{3.37}$$

$$\tilde{\psi}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \psi(x) \mathrm{e}^{-\mathrm{i}kx} \mathrm{d}x, \tag{3.38}$$

with

$$\tilde{\psi}(p) = \frac{\tilde{\psi}(k)}{\sqrt{\hbar}} \tag{3.39}$$

for $p = k\hbar$.

In this course, we will always use a tilde [e.g., $\tilde{\psi}(p)$ or $\tilde{\psi}(k)$] to denote wavefunctions in the momentum or wavenumber representations.

Exercise 3.15. As we know (Sec. A.4), the inner product between any two states $|\psi\rangle$ and $|\varphi\rangle$ is independent of the basis in which it is calculated. Check this explicitly for the position and momentum basis, i.e., show that

$$\int_{-\infty}^{+\infty} \psi^*(x)\varphi(x)\mathrm{d}x = \int_{-\infty}^{+\infty} \tilde{\psi}^*(p)\tilde{\varphi}(p)\mathrm{d}p$$

using *only* the relations (3.36) and the properties of the Fourier transform.

Exercise 3.16. Show that, for a state with a real wavefunction $\psi(x)$, $\mathrm{pr}(p) = \mathrm{pr}(-p)$ and the expectation value of the momentum observable is zero.

Exercise 3.17. The matrix element $A(x,x') = \langle x|\hat{A}\,|x'\rangle$ of the operator $\hat{A}$ is known for all x and x'. Find $\tilde{A}(p,p') = \langle p|\hat{A}\,|p'\rangle$.

Exercise 3.18. Consider a function $V(\hat{x})$ of the position operator. Write the matrix element of this operator:

a) in the position basis;
b) in the momentum basis.

Answer:

a)

$$V(x,x') = V(x)\delta(x-x'); \tag{3.40}$$

b)

$$V(p,p') = \frac{1}{2\pi\hbar}\int_{-\infty}^{+\infty} \mathrm{e}^{\frac{\mathrm{i}}{\hbar}x(p'-p)}V(x)\mathrm{d}x. \tag{3.41}$$

If you have studied introductory quantum mechanics, you may have encountered the expression

$$\hat{p} = -\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}x}, \tag{3.42}$$

which says that the momentum corresponds to an operator whose action on the wavefunction is differentiation. In the context of the more rigorous theory we are developing here, this statement does not make much sense. Operators act on state vectors, and the wavefunction is not a vector; rather, it is an inner product of two vectors — that is, a number. How can an operator act on a number? Let us figure this out.

Exercise 3.19. Show that the matrix element $\langle x|\,\hat{p}\,|x'\rangle$ of the momentum in the position representation is given by

$$\langle x|\,\hat{p}\,\big|x'\big\rangle = -\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}x}\delta(x-x') = \mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}x'}\delta(x-x') \tag{3.43}$$

Exercise 3.20. Show that, for an arbitrary state $|\psi\rangle$,

$$\langle x|\,\hat{p}\,|\psi\rangle = -\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}x}\langle x|\,\psi\rangle = -\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}x}\psi(x). \tag{3.44}$$

This result explains the meaning of Eq. (3.42). If the state $|\psi\rangle$ has wavefunction $\psi(x)$ in the position basis, then the state $\hat{p}\,|\psi\rangle$ has wavefunction $-\mathrm{i}\hbar\mathrm{d}\psi(x)/\mathrm{d}x$. This is the sense in which this equation is used in calculations, even though it is questionable from a rigorous mathematical point of view.

Exercise 3.21.[§] Obtain the analogues of the above results for the position operator in the momentum representation.

a) Show that the matrix element is

$$\langle p|\hat{x}\,\big|p'\big\rangle = \mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}p}\delta(p-p') \tag{3.45}$$

b) Show that, for an arbitrary state $|\psi\rangle$,

$$\langle p|\,\hat{x}\,|\psi\rangle = \mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}p}\tilde{\psi}(p). \tag{3.46}$$

Exercise 3.22. Show that $\langle x|\,\hat{p}^2\,|\psi\rangle = -\hbar^2 d^2\psi(x)/\mathrm{d}x^2$.

3.3.2 Position–momentum uncertainty

Now that we have some practice at switching between the position and momentum bases, we are ready to introduce the uncertainty relation between these observables. As we know from Sec. 1.9.3, the uncertainty relation corresponding to any two observables is determined by their commutator.

Exercise 3.23. Show that, for any state $|\psi\rangle$:

a)

$$\langle x|\,\hat{x}\hat{p}|\,\psi\rangle = -\mathrm{i}\hbar x\frac{\mathrm{d}}{\mathrm{d}x}\psi(x); \tag{3.47}$$

b)

$$\langle x|\,\hat{p}\hat{x}|\,\psi\rangle = -\mathrm{i}\hbar x\frac{\mathrm{d}}{\mathrm{d}x}\psi(x) - \mathrm{i}\hbar\psi(x); \tag{3.48}$$

c)

$$[\hat{x}, \hat{p}] = \mathrm{i}\hbar. \tag{3.49}$$

Exercise 3.24. Show that the Heisenberg uncertainty principle for the position and momentum observables and any state $|\psi\rangle$ has the form

$$\langle\psi|\,\Delta x^2\,|\psi\rangle\,\langle\psi|\,\Delta p^2\,|\psi\rangle \geq \frac{\hbar^2}{4}. \tag{3.50}$$

We have thus obtained the uncertainty principle in its original form: a particle state with simultaneously precise position and momentum is not possible[4].

Exercise 3.25. Perform the following calculations for a Gaussian wavefunction

$$\psi(x) = \frac{1}{(\pi d^2)^{1/4}} \mathrm{e}^{\mathrm{i}\frac{p_0 x}{\hbar}} \mathrm{e}^{-\frac{(x-a)^2}{2d^2}}: \tag{3.51}$$

a) Check normalization.
b) Find the corresponding wavefunction in the momentum basis.
Hint: use the standard rules for the Fourier transformation.
Answer:
$$\tilde{\psi}(p) = \left(\frac{d^2}{\pi\hbar^2}\right)^{1/4} \mathrm{e}^{-\frac{\mathrm{i}(p-p_0)a}{\hbar}} \mathrm{e}^{-\frac{(p-p_0)^2 d^2}{2\hbar^2}}. \tag{3.52}$$

c) Determine the expectation values and uncertainties of the position and momentum, as well as the product of these uncertainties.
Answer:
$$\langle x\rangle = a;\ \langle p\rangle = p_0;\ \langle\Delta x^2\rangle = d^2/2;\ \langle\Delta p^2\rangle = \hbar^2/2d^2. \tag{3.53}$$

We see that, for Gaussian states, the position–momentum variance product equals $\hbar^2/4$, the minimum allowed by the uncertainty principle. Here we can relate the position–momentum uncertainty to the properties of the Fourier transformation (Sec. D.2): if the wavefunction in the position basis becomes "narrower", its Fourier image, i.e., the wavefunction in the momentum basis, becomes "wider". The general quantum uncertainty principle, is, of course, much broader in scope: it holds for *any* pair of non-commuting observables, no matter whether they are related by the Fourier transform.

Exercise 3.26.[*§] Show that Gaussian wavepackets of the form (3.51) are the only states for which the inequality (3.50) expressing the uncertainty principle saturates[5].

[4] The original Heisenberg formulation was actually slightly different. See Box 3.3 for further discussion.

[5] A solution can be found in, e.g., Ulf Leonhardt, *Measuring the quantum state of light* (Cambridge University Press, 1997).

Box 3.3 Can position and momentum be measured simultaneously?

In his original work*, Werner Heisenberg formulated the uncertainty principle as follows:

> *The more precisely the position is determined, the less precisely the momentum is known, and vice versa.*

Here is a counterexample that shows a flaw in this formulation.** Suppose we prepare a particle of mass M in the position eigenstate $|x=0\rangle$ at time $t=0$. Because its position is certain, the momentum is completely uncertain. We let this state evolve freely for some time t_0 and then perform a measurement of $\hat{x}$, obtaining some value x_0. Now the position of the particle immediately before the measurement is precisely known. But so is the momentum! Indeed, because the position at $t=0$ was known to be precisely $x=0$, and at $t=t_0$ it is precisely $x=x_0$, we conclude that the velocity before the measurement was precisely $v=x_0/t_0$, whence the momentum must be $p_0=mx_0/t_0$.

Naturally, this example does not contradict the uncertainty principle as defined by Eq. (3.50). That equation says that the measurements of $\hat{x}$ and $\hat{p}$ will exhibit a certain degree of randomness, but it does not say that they cannot be *correlated* with each other. This is exactly the case for our particle: because its initial state has a completely uncertain momentum, the values of x_0 and p_0 that the measurement at t_0 could yield are completely unpredictable. However, they are correlated, being proportional to each other.

Our example shows that it is possible to know the position and momentum of a particle simultaneously *post factum*, after the measurement. However, it is impossible to prepare a particle whose position and momentum are both known *a priori*, before the measurement.

Let me also clarify that there is no contradiction to our discussion in Sec. 1.9.3, where we said that non-communing observables cannot be measured simultaneously. That discussion referred to the possibility of constructing an apparatus that would yield the same information about the position and momentum for *every* state. In our example, on the other hand, the simultaneous information about these observables is obtained for one *specific* state that we have deliberately concocted in order to devise a paradox.

*W. Heisenberg, *Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik*, Zeitschrift für Physik **43**, 172 (1927).

**This example is due to A. V. Belinsky and V. B. Lapshin.

3.3.3 The original Einstein–Podolsky–Rosen paradox

Now let us reproduce another research masterpiece, the 1935 Einstein–Podolsky–Rosen paradox. In Sec. 2.3.1 we studied a version of that paradox adapted to the quantum system we were considering at that time: the polarization of the photon. Now we possess enough tools to handle the EPR argument in its original form.

Suppose each of the two observers, Alice and Bob, holds a one-dimensional point-like particle. The two particles are prepared in an entangled state $|\Psi_{AB}\rangle$ whose wavefunction is

$$\Psi(x_A,x_B)=\delta(x_A-x_B) \tag{3.54}$$

(neglecting normalization). In other words, Alice's and Bob's particles (in their respective reference frames) always have the same space coordinate, but the specific value of that coordinate is completely random.

Exercise 3.27. For the original EPR state (3.54), answer the following questions:

a) Find the wavefunction of the two particles in the momentum representation.
b) Suppose Alice performs a measurement of her particle's position and obtains some result x_0. Onto which state will Bob's particle project?
c) Suppose Alice instead performs a measurement of her particle's momentum and obtains some result p_0. Onto which state will Bob's particle project?

Answer:

a) $\tilde{\Psi}(p_A, p_B) = \delta(p_A + p_B)$;
b) $|x_0\rangle$;
c) $|-p_0\rangle$.

We see that if Alice chooses to measure the position of her particle, she will remotely prepare Bob's particle in a state in which the position is certain and the momentum is uncertain. On the other hand, if Alice measures the momentum, Bob obtains a state with a certain momentum and uncertain position. In this way, Alice can remotely, without any interaction, choose to prepare one of two mutually incompatible realities at Bob's location.

One may object that the argument requires the use of singular wavefunctions, which, as emphasized previously, are unphysical. This objection is valid. However, the EPR paradox can be readily reformulated for a physically possible Gaussian state in which the correlation of the positions and the anticorrelation of the momenta is *almost*, but not exactly, precise. In this way, the state becomes physically plausible while local realism remains violated. We shall see this in Sec. 3.10.3.

Let me emphasize that the original EPR paradox does not demonstrate the nonlocality of nature to the same extent as Bell's experiment does. The Bell inequality is valid for any local realistic experiment whose front end is described by Fig. 2.2, so one need not believe in quantum mechanics in order to be convinced of nonlocality by observing the Bell inequality violation in an experiment. The original EPR Gedankenexperiment, on the other hand, will not appear paradoxical to someone who does not believe in quantum mechanics, and in particular, to someone who does not believe in the uncertainty principle. Indeed, if particles are allowed simultaneously to have certain position and momentum, the observed correlations can be readily explained by saying that Alice's and Bob's particles are prepared every time with the same (but random) positions and opposite (but random) momenta. In Bell's language, this means that the original EPR experiment, unlike Bell's experiment, can be explained by a local hidden variable model.

3.4 The free space potential

So far, we have discussed static, time-independent properties of the de Broglie wave. Now let us see how that wave evolves with time. As we postulated in Sec. 1.25, the quantum evolution is determined by the Hamiltonian, which is the sum of the

Box 3.4 Just add hats?

We obtained Eq. (3.55) by placing hats on top of the variables in the corresponding classical expression. While this operation has only a small effect on the appearance of the equation, it changes its physical essence quite drastically: variables turn into operators. By what right are we making this change?

As an example, consider the relation between the momentum and kinetic energy. The momentum observable is

$$\hat{p} = \int_{-\infty}^{+\infty} p\,|p\rangle\langle p|\,\mathrm{d}p,$$

which means, according to the definition given in Sec. 1.9.1, that the set of all kets $|p\rangle$ forms an orthonormal basis of the Hilbert space, and each of these kets denotes the state of the particle with a certain momentum value, p.

Now each such state also has a certain kinetic energy, $K = p^2/2M$. Therefore the kinetic energy observable would be written, according to the same definition, as

$$\hat{K} = \int_{-\infty}^{+\infty} \frac{p^2}{2M}\,|p\rangle\langle p|\,\mathrm{d}p.$$

But according to Definition A.25 for operator functions, this expression can be written simply as

$$\hat{K} = \frac{\hat{p}^2}{2M}.$$

kinetic and potential energies expressed as functions of the particle's position and momentum:

$$\hat{H} = V(\hat{x}) + \frac{\hat{p}^2}{2M}. \tag{3.55}$$

This Hamiltonian is identical to the classical one except that the canonical observables are written as operators (see Box 3.4 for a discussion of why we can do this). Here M is the particle mass, $\hat{p}^2/2M$ is the kinetic energy operator, and $V(\hat{x})$ is the potential energy, which is a function of the position observable.

The motion of the particle and the evolution of its state depend on the specific form of the potential $V(\hat{x})$. Let us start with the simplest case, $V(x) \equiv 0$ (*free space* evolution). Under this condition, any eigenstate $|p\rangle$ of the momentum operator with eigenvalue p is also an eigenstate of the Hamiltonian (3.55) with the eigenvalue (energy) $E = p^2/2M$.

Exercise 3.28. Show that the wavefunction describing the evolution of the state $|p\rangle$ under the Hamiltonian (3.55) with $V(x) \equiv 0$ is given by

$$\psi_p(x,t) = \frac{1}{\sqrt{2\pi\hbar}} \mathrm{e}^{\mathrm{i}\frac{p}{\hbar}x - \mathrm{i}\frac{p^2}{2M\hbar}t}. \tag{3.56}$$

According to the above, the time-dependent behavior of the wavefunction of the momentum eigenstate is similar to that of a traveling wave with wavenumber $k = p/\hbar$ and angular frequency

$$\omega = \frac{p^2}{2M\hbar} = \frac{\hbar k^2}{2M}. \tag{3.57}$$

The evolution of this wave constitutes a translation with *phase velocity* (Box 3.5) $v_{\text{ph}} = \lambda_{dB}/T = \omega/k = p/2M$, where $T = 2\pi/\omega$ is the period associated with the wave motion.

Surprisingly, this phase velocity is different from the value p/M expected classically. The explanation is that, in the (unphysical) momentum eigenstate, the position is completely uncertain and the probability of finding the particle is uniform over the entire one-dimensional universe. This probability does not change with time. Accordingly, the phase velocity of the de Broglie wave does not correspond directly to the motion of matter.

In order to understand how the Schrödinger evolution translates into motion, we have to study a state whose wavefunction is to some extent localized in space (we use the term *wavepacket* for such wavefunctions). The motion of these waves is governed by the *group velocity*:

$$v_{\text{gr}} = \frac{\mathrm{d}\omega}{\mathrm{d}k} \overset{(3.57)}{=} \frac{\hbar k}{M} = \frac{p}{M}, \tag{3.58}$$

in exact agreement with classical expectations[6].

For example, let us look at a Gaussian state with nonzero mean momentum. As we learned in Ex. 3.25, we can decompose it into a set of de Broglie waves. Each of these waves evolves according to Eq. (3.28). How will this evolution affect the wavepacket as a whole?

Exercise 3.29* Consider a wavefunction which, at time $t = 0$, has a Gaussian form (3.51).

a) Find the corresponding wavefunction $\tilde{\psi}(k,0)$ in the wavenumber representation. Find its evolution $\tilde{\psi}(k,0)$ under the free space Hamiltonian.

---

[6] In fact, the phase velocity of the de Broglie wave is a matter of convention rather than physics. Suppose we shift the potential energy reference point by $-V_0$ so that the particle is now under a constant potential $V(x) = V_0$. The same physical state as (3.56) would now have energy $E + V_0$, so its time-dependent wavefunction would be of the form

$$\psi_p(x,t) = \frac{1}{\sqrt{2\pi\hbar}} \mathrm{e}^{\mathrm{i}\frac{p}{\hbar}x - \mathrm{i}\frac{E+V_0}{\hbar}t}.$$

The spatial behavior of this wavefunction is the same as that in Eq. (3.56), because it is determined by the momentum, and the latter is related to the kinetic energy, which did not change. But the time evolution will depend on V_0 because the frequency of the wave now equals $(E + V_0)/\hbar$ rather than $E/\hbar$. The phase velocity will thus depend on V_0, too.

The group velocity, on the other hand, is proportional to the *derivative* of the energy, and is thus independent of the choice of the zero potential energy reference point.

Box 3.5 Phase and group velocities

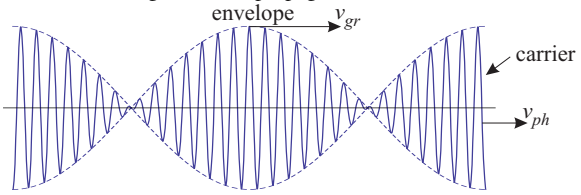
The phase and group velocities are fundamental notions from wave mechanics. Let us review them briefly here. Consider a wave propagating along the z axis:

$$W(z,t) = W_0 \mathrm{Re}\left[\mathrm{e}^{\mathrm{i}kz - \mathrm{i}\omega t}\right].$$

The exact nature of the wave does not matter: it can be an optical, acoustic, or a quantum de Broglie wave. The above equation can be rewritten as

$$W(z,t) = W_0 \mathrm{Re}\left[\mathrm{e}^{\mathrm{i}k(z - \mathrm{i}v_{\mathrm{ph}}t)}\right],$$

where $v_{\mathrm{ph}} = \omega/k$ is the *phase velocity*. It is clear from the above equation that this is the velocity with which points of constant phase (*wave fronts*) travel. It is determined by the function $k(\omega)$, known as the *dispersion relation*. This function depends on the physics of the wave and/or the medium through which it propagates.



Now suppose the wave is modulated as shown in the figure. At time $t = 0$, it has the form

$$W(z,0) = W_0 \mathrm{Re}\left[\mathrm{e}^{\mathrm{i}kz}\right] \cos \Delta k z = \frac{1}{2} W_0 \mathrm{Re}\left[\mathrm{e}^{\mathrm{i}(k+\Delta k)z} + \mathrm{e}^{\mathrm{i}(k-\Delta k)z}\right],$$

where $\Delta k \ll k$ describes the *modulation envelope*. Let us find the velocity of that envelope. Setting a nonzero time in the above equation, we find

$$\begin{aligned} W(z,t) &= \frac{1}{2} W_0 \mathrm{Re}\left[\mathrm{e}^{\mathrm{i}(k+\Delta k)z - \mathrm{i}(\omega+\Delta\omega)t} + \mathrm{e}^{\mathrm{i}(k-\Delta k)z - i(\omega-\Delta\omega)t}\right] \\ &= W_0 \mathrm{Re}\left[\mathrm{e}^{\mathrm{i}kz - \mathrm{i}\omega t}\right] \cos(\Delta k z - \Delta\omega t) \\ &= W_0 \mathrm{Re}\left[\mathrm{e}^{\mathrm{i}k(z - v_{\mathrm{ph}}t)}\right] \cos[\Delta k (z - v_{\mathrm{gr}} t)], \end{aligned}$$

where $\Delta\omega$ is the frequency shift corresponding to the shift Δk of the wavenumber and $v_{\mathrm{gr}} = \Delta\omega/\Delta k$ is the *group velocity*, the velocity at which the envelope propagates.

The group velocity determines, for example, the speed of signals carried by the wave. In systems in which the wavenumber is proportional to the frequency (for example, electromagnetic waves in vacuum), the phase and group velocities are equal. If the relation between these two quantities is more complicated, these velocities can differ dramatically, giving rise to many curious phenomena.

b) Use the inverse Fourier transform to find the wavefunction $\psi(x,t)$ in the position basis.
Hint: for the direct and inverse Fourier transformations, use properties (D.13) and (D.14).
c) Find the mean value $\langle x\rangle$ and variance $\langle \Delta x^2\rangle$ of the position as functions of time.
Answer: $\langle x\rangle = a + (p_0/M)t$, $\langle \Delta x^2\rangle = \frac{d^2}{2}\left(1+\frac{\hbar^2 t^2}{M^2 d^4}\right)$

As expected, the wavepacket moves with the effective group velocity $v_{\rm gr} = p_0/M$. But in addition, it becomes wider and wider as time goes by. This phenomenon, known as *spreading of the wavepacket*, is a consequence of *group velocity dispersion*, i.e., the group velocity (3.58) not being the same for different values of k. As a result, the simple description of the motion of the wavefunction in terms of the phase and group velocities, as in Box 3.5, is only approximately valid.

It is instructive to compare this behavior with that exhibited by laser pulses. Such pulses can propagate over long distances in a vacuum without any spreading because the group velocity of light in vacuum is a constant; it does not depend on the frequency or wavenumber. But when the propagation takes place in a refractive medium with strong dispersion, such that the refractive index and hence the group velocity vary as a function of the frequency, the pulses will spread.

We learn from the above results that the spreading can be neglected as long as $\frac{\hbar}{md^2}t \ll 1$, in which case the shape of the Gaussian wavepacket remains the same; it travels as a single unit, reproducing the classical motion of a pointlike particle. The above condition is almost always true for microscopic objects.

Even for microscopic objects, the spreading effect is quite difficult to observe experimentally. This is, in particular, because of the particle's interaction with other objects. As we discussed in Sec. 2.4.2, such interaction brings about decoherence, which collapses the state onto a position eigenstate or a mixture thereof, thereby "resetting" the spreading. The spreading will also be suppressed if the particle is localized in some kind of a potential well, which we shall study shortly.

Exercise 3.30. Estimate the time required by:

a) the wavepacket associated with a single electron with position uncertainty of the order of 1 Å to spread over a length of 1 mm;
b) the wavepacket associated with a metal ball of mass 1 g with the position uncertainty of the order of 1 Å to spread over a length of 1 mm;
c) the wavepacket associated with a 40-kg interferometer mirror in the LIGO gravitational wave project, whose position is known with a $d = 10^{-18}$ meter accuracy, to spread until its position variance doubles.

Exercise 3.31. Show that, if the mean momentum greatly exceeds the momentum uncertainty of the initial wavepacket, the distance traveled by the center of the wavepacket during time t is much greater than the length over which it spreads.

3.5 Time-independent Schrödinger equation

For the remainder of this chapter, we will study the quantum behavior of a pointlike particle in the field of some conservative force. We know that it is governed by the Schrödinger equation. Rather than finding its general solution, we will first master a more modest task: *finding the set of energy eigenvalues and eigenstates* for a certain potential. If we address this problem successfully, we will be able to determine the time-dependent dynamics as well. To that end, we would decompose the initial state into the energy eigenstates and then apply the evolution equation (1.25) to each of these states.

In addition to being useful for calculating the evolution, energy eigenstates are physically significant because they often form a decoherence-preferred basis (see Sec. 2.4.2). This means that these states and their statistical mixtures occur much more frequently than their coherent superpositions.

Furthermore, energy eigenstates can be probed with light. The transition between these states in atoms or molecules is associated with absorption or emission of a photon whose energy $\hbar\omega$ equals the difference between corresponding energies. By performing spectroscopy — measuring the wavelengths at which the absorption or emission occurs — one can determine the respective energies, and thereby check a quantum-mechanical calculation experimentally.

So our goal is to find states $|\psi\rangle$ such that

$$\hat{H}|\psi\rangle = E|\psi\rangle. \tag{3.59}$$

This equation is referred to as the *time-independent Schrödinger equation*. As a rule, we will be working in the position basis, looking for the wavefunction $\psi(x)$ of state $|\psi\rangle$. To that end, we take the inner product of both sides of Eq. (3.59) with the bra vector $\langle x|$.

Exercise 3.32. Show that, in the x-basis, the time-independent Schrödinger equation (3.59) takes the form

$$\left[V(x) - \frac{\hbar^2}{2M}\frac{d^2}{\mathrm{d}x^2}\right]\psi(x) = E\psi(x). \tag{3.60}$$

This is a second-order ordinary differential equation, which can be solved either analytically or numerically. Before we go on to find solutions for specific potentials, let us discover a few general properties of these solutions.

Exercise 3.33. Find the general solutions to Eq. (3.60) for $V(x) = V_0$. Consider the following cases:

a) $E > V_0$;
b) $E < V_0$.

Answer:

a) $Ae^{ikx} + Be^{-ikx}$, where $k = \sqrt{2M(E-V_0)}/\hbar$;

b) $Ae^{\kappa x} + Be^{-\kappa x}$, where $\kappa = \sqrt{2M(V_0 - E)}/\hbar$

with A and B being arbitrary coefficients.

We see that the solutions are fundamentally different for the energies above and below the potential level. In the former case, we obtain oscillatory behavior akin to the de Broglie wave. In the latter case, the solutions grow or fall off exponentially as a function of the position. Such a solution will blow up at $x \to \pm\infty$ — a behavior which implies infinite probabilities and thus cannot occur in a physical state (or even in an approximation thereof).

The next exercise generalizes this observation to arbitrary potentials.

Exercise 3.34. Show that the Hamiltonian (3.55) cannot have eigenvalues that are less than the minimum of the function $V(x)$ on the real axis.

In other words, there can be no energy eigenvalues such that $E < V(x)$ for all x. However, situations where the energy is lower than the potential for a *part* of the x axis are possible, as is the case, for example, with quantum tunnelling (which we will study shortly).

Exercise 3.35. Show that, if $\psi(x)$ is a solution of the time-independent Schrödinger equation, then both $\psi(x)$ and $\mathrm{d}\psi(x)/\mathrm{d}x$ must be continuous at points where the potential $V(x)$ is finite.

This result will turn out to be extremely useful for many problems in which the potential is given by a piecewise function, i.e., a set of different elementary functions each defined on its own interval of positions. It is relatively easy to find the solution for each interval, but then these solutions must be "stitched" together to form a physically meaningful wavefunction. Exercise 3.35 provides us with the guidelines for this "stitching".

Exercise 3.36. Consider the set S_E consisting of all Hamiltonian eigenstates with energy eigenvalue E. Show that there exists a spanning set of S_E which consists only of states with real wavefunctions

For example, the de Broglie wave

$$\psi_p(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{i\frac{px}{\hbar}}, \tag{3.61}$$

associated with the momentum eigenstate $|p\rangle$, is a solution of the time-independent Schrödinger equation with energy eigenvalue $E = p^2/2M$. The same is true for the wavefunction

$$\psi_{-p}(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{-i\frac{px}{\hbar}}, \tag{3.62}$$

which is the de Broglie wave for the momentum eigenstate $|-p\rangle$. The set S_E consists of states $|\pm p\rangle$ and their linear combinations. In particular, real wavefunctions

$$\psi_{p,+}(x) = \frac{\psi_p(x) + \psi_{-p}(x)}{2} = \frac{1}{\sqrt{2\pi\hbar}} \cos\frac{px}{\hbar} \tag{3.63}$$

and

$$\psi_{p,-}(x) = \frac{\psi_p(x) - \psi_{-p}(x)}{2i} = \frac{1}{\sqrt{2\pi\hbar}} \sin\frac{px}{\hbar} \tag{3.64}$$

also represent energy eigenstates with the same eigenvalue. The de Broglie wavefunctions (3.61) and (3.62) — and hence any other wavefunction corresponding to the same energy — can be written as linear combinations of these real wavefunctions.

In this way, Ex. 3.36 simplifies our search for solutions of the time-independent Schrödinger equation. We can restrict the search to real wavefunctions without fear of "missing" anything: any other solution can be written as a linear combination of real ones.

Exercise 3.37. Consider the set S_E consisting of all Hamiltonian eigenstates with energy eigenvalue E. Show that, if $V(x)$ is an even function of the position, there exists a spanning set of S_E consisting only of states with wavefunctions that are either even or odd.

3.6 Bound states

Bound states are characterized by a wavefunction that tends to zero at both $x \to \infty$ and $x \to -\infty$, so that the particle exhibits some degree of localization. This property is typical for energy eigenstates in well-like potentials, i.e., fields in which the particle is attracted towards a certain location or set of locations. Physical examples include a pea inside a cup, a ball on a spring (harmonic oscillator), or an electron within an atom. For this type of potential, we usually take advantage of Ex. 3.36 and look for solutions of the time-independent Schrödinger equation in the real domain.

Exercise 3.38. Consider a potential $V(x)$ that asymptotically approaches the values $V_{1,2}$ at $|x| \to \pm\infty$, respectively. Show that an energy eigenstate is bound if and only if its energy does not exceed $\min(V_1, V_2)$.

The boundary conditions imposed on the wavefunction at $x \to \pm\infty$ augment the differential time-independent Schrödinger equation, giving rise to a Cauchy problem. This problem has solutions only for specific, discrete values of the energy. In other words, bound states exist for a *discrete* or *quantized* spectrum of energy eigenvalues, which are called *energy levels*.

Exercise 3.39. Find the energy eigenvalues and eigenwavefunctions for the *finite square well* potential:

$$V(x) = \begin{cases} V_0 \text{ for } |x| > a/2 \\ 0 \text{ for } |x| \le a/2 \end{cases} \tag{3.65}$$

with $V_0 > 0$ (Fig. 3.1).

a) Write the general solution for each region where the potential is constant. Eliminate unphysical terms that grow at infinity.

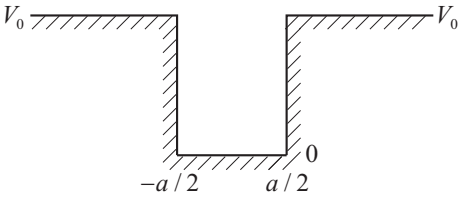


Fig. 3.1 Potential for Ex. 3.39

Hint: use the result of Ex. 3.37.
Answer: A generic odd wavefunction is of the form

$$\psi(x) = \begin{cases} -B\mathrm{e}^{\kappa x}, & x < -a/2 \\ A\sin kx, & -a/2 \le x \le a/2 \\ B\mathrm{e}^{-\kappa x}, & x > a/2 \end{cases}, \tag{3.66}$$

and a generic even wavefunction

$$\psi(x) = \begin{cases} B\mathrm{e}^{\kappa x}, & x < -a/2 \\ A\cos kx, & -a/2 \le x \le a/2 \\ B\mathrm{e}^{-\kappa x}, & x > a/2 \end{cases}, \tag{3.67}$$

where

$$k = \frac{\sqrt{2ME}}{\hbar}, \tag{3.68a}$$

$$\kappa = \frac{\sqrt{2M(V_0 - E)}}{\hbar}. \tag{3.68b}$$

b) Apply the statement of Ex. 3.35 to "stitch" these results together. Show that the energy values that can simultaneously allow continuity of the wavefunction and its derivative at $x = \pm a/2$ must obey the transcendental equations

$$\tan\theta = \sqrt{\frac{\theta_0^2}{\theta^2} - 1}$$

for even wavefunctions and

$$-\cot\theta = \sqrt{\frac{\theta_0^2}{\theta^2} - 1}$$

for odd wavefunctions, where

$$\theta = \frac{ka}{2} \text{ and } \theta_0 = \frac{\sqrt{2MV_0}}{\hbar}\frac{a}{2}.$$

c) Solve these equations numerically and plot the energies of the three lowest bound states as a function of the depth V_0 of the potential well.
Answer: see Fig. 3.2(a).

d) What is the minimum depth that is required in order for the well to contain a given number N of bound eigenstates?
Answer: $[\pi\hbar(N-1)]^2/2Ma^2$.

e) Plot the wavefunctions corresponding to all possible energy eigenvalues for $V_0 = \frac{9}{2}\hbar^2/Ma^2$, $V_0 = \frac{49}{2}\hbar^2/Ma^2$, and also the three lowest energy solutions for $V_0 = \infty$.
Answer: see Fig. 3.2(b).

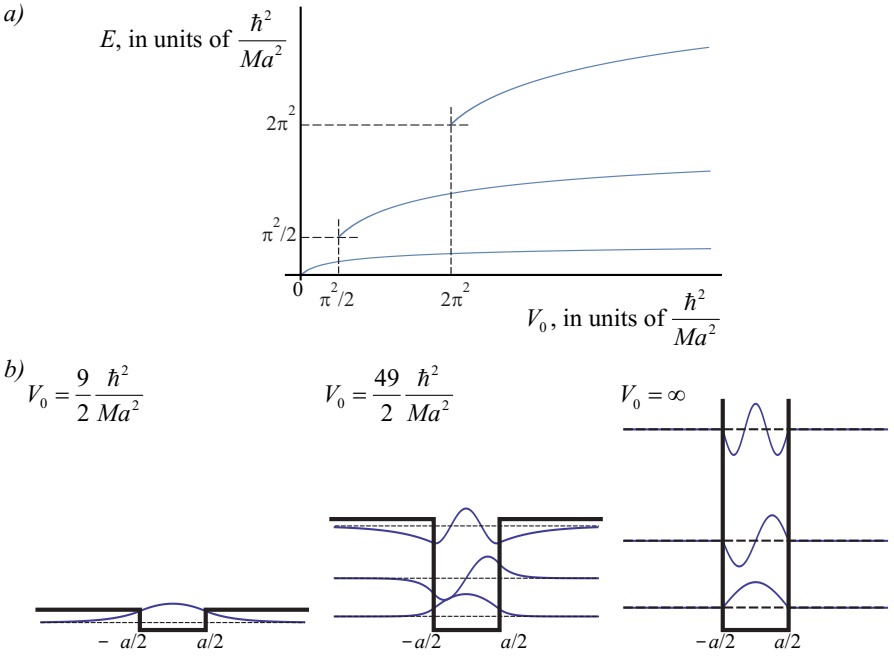


Fig. 3.2 Solution for Ex. 3.39. a) Lowest energy eigenvalues as a function of the well depth. At least one bound state exists for all values of V_0; the existence of further bound states is conditional on V_0 exceeding certain threshold values. b) Wavefunctions for the lowest energy eigenstates with different well depths. The well on the left supports only one bound state, the well in the middle three bound states, and the well on the right infinitely many.

This problem requires more work than most of our other exercises, yet I would advise you to try it, or at least carefully study the solution, because it is a good illustration of the general behavior of bound state wavefunctions. Let us discuss the salient features of this behavior.

As we can see from Fig. 3.2(b), the wavefunction extends outside the box, so there is a finite probability of finding the particle in the region where the potential is higher than the particle's energy. This is, of course, an expressly nonclassical phenomenon: if the particle were a classical ball bouncing in a gap between two walls, we would never find it outside that gap. The greater the difference between the energy E of the state and the well depth V_0, the faster the decay of the wavefunction outside the well, and the lower the probability of finding the particle in that region. In the limit $V_0 \to \infty$, this probability tends to zero. In this case, the problem admits an analytic solution, as we shall see in the exercise below.

In contrast to the exponential decay outside the well, the wavefunction exhibits oscillatory behavior inside, as per Ex. 3.33. For each subsequent energy eigenstate, the number of times the wavefunction crosses the x axis goes up by one. The increasing number of crossings is associated with faster oscillation, a higher wavenumber, and hence a higher energy value. Accordingly, for each nonzero number of crossings there exists a certain minimum potential, below which the bound state no longer exists [Fig. 3.2(a)]. The deeper and wider the potential well, the more bound states it can support. However, no matter how shallow it is, it does support the bound state with no zero crossings.

Exercise 3.40. Find the energy eigenvalues and the wavefunctions of bound stationary states for Ex. 3.39 in the case $V_0 \to \infty$ (known as the *box potential*).
Answer: A discrete energy spectrum with

$$E_n = \frac{\hbar^2 \pi^2 n^2}{2Ma^2} \tag{3.69}$$

and eigenwavefunctions

$$\psi_n(x) = \begin{cases} \left\{ \begin{array}{l} \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi x}{a}\right), \text{ even } n \\ \sqrt{\frac{2}{a}} \cos\left(\frac{n\pi x}{a}\right), \text{ odd } n. \end{array} \right\}, & -a/2 \le x \le a/2 \\ 0, & |x| > a/2 \end{cases} . \tag{3.70}$$

These wavefunctions are displayed in Fig. 3.2(b), right panel.

They exhibit a number of interesting features:

- $\psi(x) = 0$ outside the box;
- $\mathrm{d}\psi(x)/\mathrm{d}x$ exhibits discontinuities at $x = \pm a/2$;
- $\psi(x)$ is continuous for all position values.

The vanishing wavefunction outside the box can be seen as the extreme case of the exponential decay outside the well observed in the previous exercise; in the present case, the well is infinitely deep and the decay constant is infinite. The infinite value of the potential outside the box also implies that we are not affected by the conditions of Ex. 3.35, so neither the wavefunction nor its derivative need be continuous at $x = \pm a/2$. However, we observe that only $\mathrm{d}\psi(x)/\mathrm{d}x$ exhibits discontinuities, while the wavefunction does not. This can be understood as follows. According to Ex. 3.33, the derivative of the wavefunction inside the box is limited

by $|\mathrm{d}\psi(x)/\mathrm{d}x| \leq k|\psi(x)|$, where $k = \sqrt{2ME}/\hbar$. Outside the box, $|\mathrm{d}\psi(x)/\mathrm{d}x| = 0$. This means that the discontinuity in the derivative of the wavefunction at the box boundary is finite, and this in turn implies that the wavefunction is continuous.

A similar argument can be made in all practical cases, so the wavefunction can always be safely assumed to be continuous — perhaps with the exception of some extremely exotic potentials. The derivative of the wavefunction, on the other hand, can exhibit discontinuities whenever the potential is infinite or singular.

Let us now look at another extreme case of a rectangular potential well, important from both the educational and the scientific point of view.

Exercise 3.41. Find the energy eigenvalues and the wavefunctions of bound stationary states of the potential $V(x) = -W_0\delta(x)$ in the position basis.
Hint: Integrate both sides of the time-independent Schrödinger equation over an infinitesimal interval around $x = 0$ and use Eq. (D.9).
Answer: A single eigenstate with $E = -W_0^2 M/2\hbar^2$ and wavefunction (Fig. 3.3):

$$\psi(x) = \sqrt{\kappa}\begin{cases} \mathrm{e}^{-\kappa x} \text{ at } x > 0 \\ \mathrm{e}^{\kappa x} \text{ at } x \leq 0 \end{cases}. \tag{3.71}$$

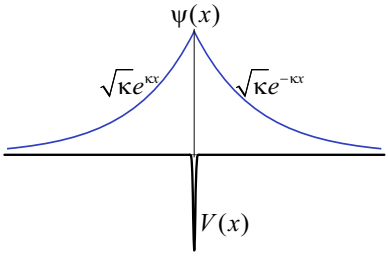


Fig. 3.3 Wavefunction of the energy eigenstate of the delta potential (Ex. 3.41).

Exercise 3.42*. Obtain the result of the previous exercise by an alternative method. Solve the time-independent Schrödinger equation for the finite well potential (3.65) analytically in the limit of an infinitely deep and narrow potential well: $a \to 0$, $V_0 = W_0/a$, with $W_0 = \text{const}$. How many bound states can this well contain?

Exercise 3.43. A particle is in the bound state of the potential $V(x) = -W_0\delta(x)$. The potential suddenly changes to $V(x) = -2W_0\delta(x)$. Find the probability that the particle will remain in a bound state.

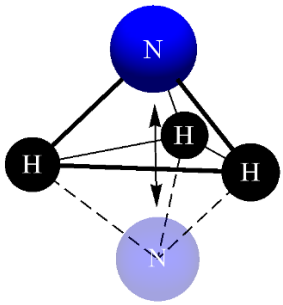
Exercise 3.44*. Investigate the bound states of the potential

$$V(x) = -W_0\delta(x-a) - W_0\delta(x+a). \tag{3.72}$$

a) Find the equation for the energy eigenvalues (consider both the even and the odd case). How many solutions does it have?

Box 3.6 Ammonia maser

The "double-delta" setting of Ex. 3.44 constitutes the theoretical basis of the first ammonia maser, the precursor of modern lasers, built in 1953 by Charles Townes and colleagues*. The source of the radiation used in the maser is the ammonia molecule NH_3, shown on the right. It has the shape of a pyramid, with a base formed by three hydrogen atoms and a nitrogen atom located "at the top". This position of the nitrogen atom corresponds to the potential energy minimum represented by one of the delta functions. The other delta function corresponds to the mirror image of that configuration, where the nitrogen atom is "below" the base plane. These configurations have the same energy, and there is a non-vanishing probability for the nitrogen atom to "hop" between them. Because of this possibility, the energy eigenstates are not the "top" and "bottom" positions of the nitrogen atom, but their symmetric and antisymmetric linear combinations, as in Ex. 3.44. It is the transition between these two states that produces the 24-GHz microwave radiation emitted by the maser.



* J. P. Gordon, H. J. Zeiger, and C. H. Townes, *Molecular Microwave Oscillator and New Hyperfine Structure in the Microwave Spectrum of NH_3*, Physical Review **95**, 282 (1954); J. P. Gordon, H. J. Zeiger, and C. H. Townes, *The Maser — New Type of Microwave Amplifier, Frequency Standard, and Spectrometer*, Physical Review **99**, 1264 (1955).

b) Show that in the limit $a \to \infty$ this equation becomes identical to the one for a single well.

c) Find the expression for the energies and wavefunctions of the Hamiltonian eigenstates for the potential (3.72) up to the first order in $\hbar^2/W_0 Ma \ll 1$.
Answer: The even and odd state energies are

$$E_{e,o} = -W_0^2 M/2\hbar^2 (1 \pm 2e^{-2\kappa_0 a}). \tag{3.73}$$

The behavior observed here is common in quantum mechanics. For example, a proton forms an attractive potential for a free electron; this potential gives rise to bound states referred to as the hydrogen atom. If there are two protons far away from each other and a single electron, the states of the electron bound to either proton correspond to the same energy eigenvalue — so this value is degenerate. But when the protons are close enough, so the electron is affected by both potentials at the same time, the energy eigenstates become nonlocal and the energy eigenvalue degeneracy is *lifted*: energy levels split as in Eq. (3.73). This splitting can be used for practical applications, as discussed in Box 3.6. Moreover, the negative energy shift of one of the new energy eigenstates can become larger than the positive potential arising due to the Coulomb repulsion of the two protons; in such a situation, a molecule will be formed.

Box 3.7 Energy: a discrete or continuous observable?

For most observables studied so far, their discrete or continuous character depends on their physical nature. For energy, it depends on the specific physical setup being studied: the energy spectrum is discrete inside potential wells and continuous for unbound states. In fact, energy spectra can contain discrete and continuous domains within the same setting. Such is the situation with a finite well (Ex. 3.39), where the states become unbound and energy spectrum continuous for $E > V_0$. To consider a more physical example, an electron can be in a bound state with respect to a nucleus, thereby forming an atom with a discrete energy spectrum, or in an unbound, continuous-spectrum state, corresponding to an ionized atom.

One may argue that energy, by nature, is a continuous variable, and the shape of the potential only determines which values of this variable associate with Hamiltonian eigenvalues. However, by definition (Sec. 1.9.1), it is exactly this association that establishes the allowed set of values of a quantum observable operator. If energy eigenstates exist for a discrete set of energy values, this makes the energy a discrete observable.

We know that discrete and continuous observables follow different normalization rules. Conveniently, the energy eigenstates are compliant with these rules. Bound states have square-integrable wavefunctions, allowing the application of the discrete-spectrum normalization rule $\langle E_i | E_j \rangle = \delta_{ij}$. Unbound wavefunctions, on the other hand, have infinite norms, as expected for continuous-spectrum states.

Another interesting feature of energy eigenstates is that, however complex their spectrum, they will form a basis in the Hilbert space of the states that are *physically allowed* under a given potential. For example, all the energy eigenwavefunctions of the infinite potential box (Ex. 3.40) vanish outside the box. The space they span is not the Hilbert space of all possible functions, but only that of the functions localized within the box, i.e., those allowed under a potential of this shape.

Exercise 3.45. Under the conditions of the previous problem (distant wells), suppose that at time $t = 0$ the particle is localized at the first well (i.e., its wavefunction is that of Ex. 3.41 centered at $x = a$). What is the probability of finding it in the second well as a function of time?

Finally, let us derive an important property of bound states that will be useful later on.

Exercise 3.46.* Show that bound energy eigenstates of a point-like particle with a single degree of freedom cannot be degenerate when the potential is bounded from below.

3.7 Unbound states

Unbound state wavefunctions take finite, nonzero values at $x \to -\infty$ or $x \to +\infty$ or both. As we found previously, this happens if the energy E satisfies the condition

$$E > V(-\infty) \text{ or } E > V(+\infty). \tag{3.74}$$

A straightforward example of an unbound state is the momentum eigenstate $|p\rangle$ in free space. The associated energy eigenvalue, $E = p^2/2M$, is greater than the potential $V(x) \equiv 0$.

Because, in contrast to the bound case, we don't have the boundary condition $\psi(x) \to 0$ at $x \to \pm\infty$, the Schrödinger equation (3.60) has a solution for *any* energy value (as long as Eq. (3.74) holds). Moreover, in some cases energy eigenstates are degenerate. Such is the case, for example, in the free-space potential where states $|\pm p\rangle$ have the same energy.

The existence of an eigenstate for any energy value satisfying Eq. (3.74) means that the energy becomes a continuous observable in this region. For this reason, unbound states are sometimes called *continuous spectrum* states. For example, in the situation of Fig. 3.1, the energy spectrum is discrete for $E < 0$ and continuous for $E \geq 0$.

As we know from Sec. 3.2, normalization for continuous-observable eigenstates is tricky and ambiguous. Typically, then, we do not worry about normalization when analyzing unbound-state wavefunctions.

3.7.1 The single-step potential

Exercise 3.47.§ Find the wavefunctions corresponding to eigenstates of the Hamiltonian with the potential

$$V(x) = \begin{cases} 0 \text{ for } x \leq 0 \\ V_0 \text{ for } x > 0 \end{cases} \tag{3.75}$$

[Fig. 3.4] corresponding to a given energy $E > V_0$, taking into account the continuity requirement for the wavefunction and its derivative at $x = 0$.
Answer: Any wavefunction of the form

$$\psi(E,x) = \begin{cases} Ae^{ik_0 x} + Be^{-ik_0 x}, & x < 0 \\ Ce^{ik_1 x} + De^{-ik_1 x}, & x \geq 0 \end{cases}, \tag{3.76}$$

where $k_0 = \sqrt{2ME}/\hbar$, $k_1 = \sqrt{2M(E-V_0)}/\hbar$ and the four amplitudes A, B, C, D satisfy

$$A + B = C + D; \tag{3.77a}$$

$$ik_0(A - B) = ik_1(C - D). \tag{3.77b}$$

We see that the general solution depends on four parameters, while the continuity conditions give rise to only two equations (3.77). An additional equation would come from normalization; however, because we agreed to neglect normalization, we simply say that any two wavefunctions that differ by a constant factor are physically identical. This leaves us with three parameters and two equations; therefore, for each energy value, there are two linearly independent solution sets. We find them by introducing an additional equation into the system.

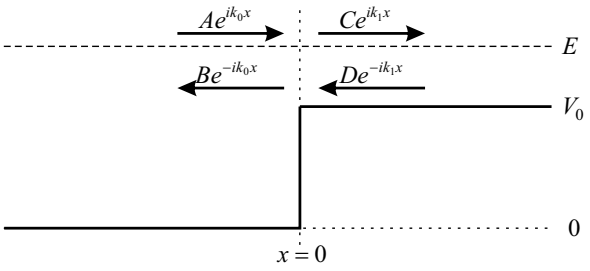


Fig. 3.4 Solution of the time-independent Schrödinger equation for the single-step potential (Ex. 3.47 and 3.51).

Exercise 3.48:§ Solve Eqs. (3.77) for B and C when the additional equation is:

a) $D = 0$,
b) $A = 0$.

Answer:

a)

$$B = A\frac{k_0 - k_1}{k_0 + k_1};\ \ C = A\frac{2k_0}{k_0 + k_1}. \tag{3.78a}$$

b)

$$B = D\frac{2k_1}{k_0 + k_1};\ \ C = D\frac{k_1 - k_0}{k_0 + k_1}. \tag{3.78b}$$

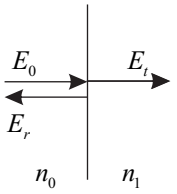
Of course, any linear combination of these solutions is also a solution.

The choice of $D = 0$ or $A = 0$ in the above exercise is dictated by the following intuition. As we found in Sec. 3.4, the evolution of a de Broglie wave of the form e^{ikx} with positive k corresponds to propagation in the positive x direction, while e^{-ikx} corresponds to propagation in the negative x direction. Accordingly, the case $D = 0$ corresponds to the de Broglie wave with amplitude A (let us call it the A-wave) coming from the left and encountering a barrier. Part of the wave passes through the barrier and becomes the C-wave; another part reflects as the B-wave. The case $A = 0$ corresponds to the particle approaching from the right (D-wave) and giving rise to B- and C-waves in transmission and reflection, respectively.

Somewhat unintuitive in this argument is perhaps our treatment of the collision of a particle with the barrier as a *stationary* state, i.e., an event of infinite duration. This is related to the infinite spatial extent of the de Broglie wave that we discussed in Sec. 3.2. Perhaps a good analogy to this effect would be a continuous laser beam which propagates from air into glass, experiencing partial reflection in accordance with the Fresnel equations (Box 3.8). Similarly to the situation with the quantum particle, the reflection is not an instantaneous event but a stationary process. Interestingly, if we compare the Fresnel equations (3.79) for the field amplitudes with Eqs. (3.78), and take into account the fact that the optical wavenumber is proporti-

Box 3.8 Fresnel equations

Consider an optical wave of amplitude E_0, propagating through a material with refractive index n_0. If it is incident on an interface with another material with refractive index n_1, it will be partially transmitted through the interface, and partially reflected. The Fresnel equations relate the amplitudes of the transmitted and reflected waves (E_t and E_r, respectively) to E_0, as functions of the angle of incidence and polarization. For normal incidence, these equations take the form



$$E_t = E_0 \frac{2n_0}{n_0 + n_1}; \tag{3.79a}$$

$$E_r = E_0 \frac{n_0 - n_1}{n_0 + n_1}. \tag{3.79b}$$

Note that for $n_0 > n_1$ we have $E_t > E_0$. However, there is no violation of energy conservation. This is because the intensity (flux of energy) of an optical wave is proportional not only to the square of its amplitude, but also to the refractive index:

$$I = 2nc\varepsilon_0 |E|^2.$$

The transmitted wave travels at a lower speed, so the flux of energy carried by that wave is reduced accordingly. The sum of the intensities of the reflected and transmitted waves, viz.,

$$I_t + I_r = 2c\varepsilon_0 (n_1 |E_t|^2 + n_0 |E_r|^2) = 2c\varepsilon_0 n_0 |E_0|^2 = I_0,$$

is equal to that of the incident wave.

onal to the reciprocal phase velocity, and hence to the index of refraction, we will find these two sets of equations to be almost identical!

A curious feature of the result (3.78a) is that the amplitude C of the transmitted de Broglie wave is higher than the amplitude A of the incident wave. Similarly to the optical case (Box 3.8), this does not contradict the law of conservation of matter because the flux of matter is proportional to both the probability density associated with the wavefunction and the phase (or group) velocity of that wavefunction. If we take this into account, we find that the conservation of matter is perfectly respected.

Exercise 3.49. Defining the *probability density current* of the de Broglie wave by $j = v_{\mathrm{ph}} |\psi(x)|^2$, find the probability density currents for the A-, B- and C-waves in Eq. (3.78a). Find the reflection and transmission coefficients for these currents, i.e., j_B/j_A and j_C/j_A. Show that their sum is 1. What is the behavior of these coefficients for $E \to V_0$ and $E \to \infty$?

Exercise 3.50. Solve Ex. 3.47 for energies below V_0. Check that the reflection coefficient is unity.

If you still feel uneasy with collisions of infinitely long duration, try the following exercise. Start with a Gaussian wavepacket moving towards the barrier, decompose it into a set of de Broglie waves, and study its evolution in a similar way to Ex. 3.29.

Exercise 3.51* Find the evolution of the state whose initial wavefunction is a Gaussian packet described by Eq. (3.51) with a positive momentum p_0 and negative center position a in the single-step potential field [Fig. 3.4]. Assume that:

- $|a| \gg d$, so the wavepacket is initially entirely to the left of the step;
- $d^2 \gg \hbar t/M$ so that spreading of the wavepacket (Ex. 3.29) can be neglected;
- the initial average energy of the particle $E = p_0^2/2M$ is greater than V_0;
- the momentum uncertainty of the wavepacket $\hbar/2d$ is small compared to the average momenta $\hbar k_0$ and $\hbar k_1$ of the incident and transmitted de Broglie waves.

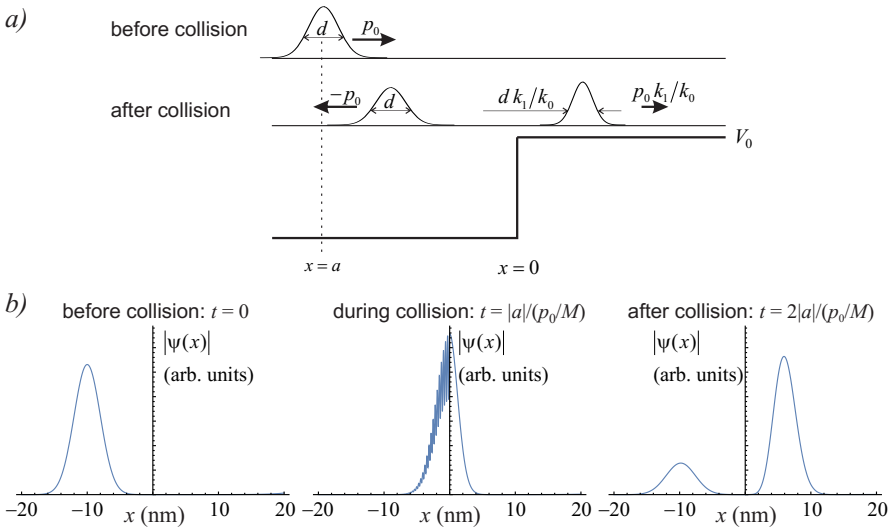


Fig. 3.5 Gaussian wavepacket interacting with a single-step potential (Ex. 3.47 and 3.51). a) Schematic diagram of the evolution. b) Numerical simulation for an electron with $a = -10$ nm, initial energy $E = 3.78$ eV (corresponding to $k_0 = 10^{10}$ m^{-1}), and potential height $V_0 = 2.42$ eV (corresponding to $k_1 = 0.6 \times 10^{10}$ m^{-1}). Interference fringes visible during the collision are between the incident and reflected waves. In the rightmost plot, the transmitted wavepacket is clearly seen to be moving more slowly than the reflected one.

The solution is presented graphically in Fig. 3.5. When it encounters the step, the initial wavepacket splits. Part of the wavepacket continues to propagate past the step with a lower group velocity, while the other part reflects off the step and begins to travel in the backward direction. It may be surprising that all this complex motion comes from simple phase rotation of the component de Broglie waves, and yet this is indeed the case!

As a final comment on the potential step problem, let us note that the particle's nonvanishing probability of bouncing off a potential step that is lower than the particle's energy or even negative [as in the case described by Eq. (3.78b)] is expressly quantum. Any classical particle will simply "fly above" the potential step, reducing or increasing its speed, but never reversing its direction of motion.

3.7.2 Quantum tunnelling

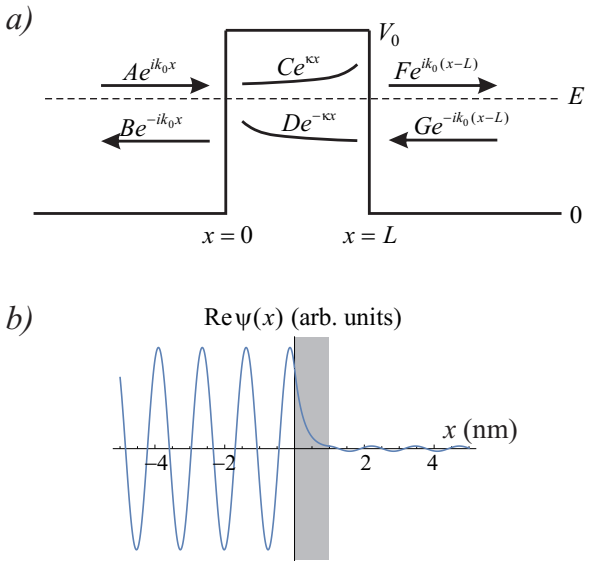


Fig. 3.6 Tunnelling through a barrier (Ex. 3.52). a) Notation for the de Broglie component waves. b) Real part of the numerical solution for an electron with initial energy $E = 0.95$ eV (corresponding to $k_0 = 0.5 \times 10^{10}$ m^{-1}) and potential height $V_0 = 1.51$ eV (corresponding to $k_0 = 0.39 \times 10^{10}$ m^{-1}). The barrier of length $L = 1$ nm is shown in gray. The three parts of the wavefunction are visible: oscillatory before the tunnel, exponentially decaying within the tunnel, and again oscillatory with a lower amplitude after the tunnel.

Exercise 3.52. Consider the potential in Fig. 3.6(a), i.e.,

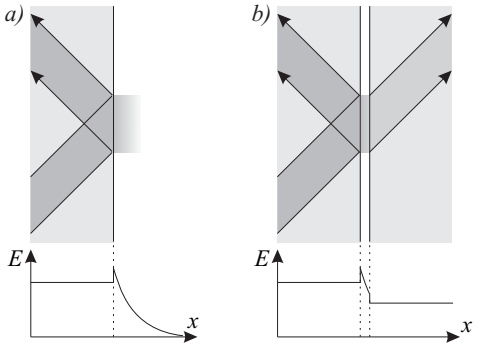
$$V(x) = \begin{cases} 0 \text{ for } x \leq 0 \text{ or } x > L \\ V_0 \text{ for } 0 < x \leq L \end{cases} \tag{3.80}$$

with the energy satisfying the condition $0 < E < V_0$.

a) What is the degeneracy of the energy levels?
b) Find the solution of the time-independent Schrödinger equation corresponding to a de Broglie wave entering from the left with energy between 0 and V_0.
c) Find the transmission and reflection coefficients for the probability current. Is their sum equal to one?
Answer:

Box 3.9 Optical analog to tunnelling

The phenomenon of quantum tunnelling also has an analog in optics. When an optical wave experiences total internal reflection, e.g., on a glass–air interface, an *evanescent wave* emerges at the opposite side of the interface (in the air). The evanescent wave normally decays exponentially on a length scale comparable to the wavelength and does not carry any energy. This situation is analogous to the one studied in Ex. 3.50. However, if another glass object is placed in close proximity to the interface, the evanescent wave will enter that object and propagate away from the interface. Similarly to the case of quantum tunnelling, the group velocity of the wave inside the air gap is infinite.



a) Evanescent wave emerging as a result of total internal reflection. b) Evanescent coupling — an optical analogue of quantum tunnelling.

$$\text{Transmission: } \frac{j_F}{j_A} = \left|\frac{F}{A}\right|^2 = \frac{4k_0^2\kappa^2}{4k_0^2\kappa^2 + (\kappa^2 + k_0^2)^2 \sinh^2(\kappa L)}; \tag{3.81a}$$

$$\text{Reflection: } \frac{j_B}{j_A} = \left|\frac{B}{A}\right|^2 = \frac{(\kappa^2 + k_0^2)^2 \sinh^2(\kappa L)}{4k_0^2\kappa^2 + (\kappa^2 + k_0^2)^2 \sinh^2(\kappa L)}, \tag{3.81b}$$

where $k_0 = \sqrt{2ME}/\hbar$ and $\kappa = \sqrt{2M(V_0 - E)}/\hbar$.

We observe that a particle encountering a finite potential barrier which is higher than the particle's kinetic energy has a finite probability of "tunnelling" through this barrier Fig. 3.6(b)]. This phenomenon has, of course, no analogy in classical physics. But even more surprising is the following.

Exercise 3.53.* Investigate the propagation of a Gaussian wavepacket through a potential shown in Fig. 3.6(a) under the same conditions and assumptions as in Ex. 3.51. Calculate and plot the center positions of the incoming and transmitted wavepackets (A- and F-waves, respectively) as a function of time.

If you do everything correctly, you should obtain a picture similar to Fig. 3.7. That is, tunnelling occurs *instantaneously*: the transmitted wavepacket emerges behind the barrier simultaneously with the initial wavepacket being absorbed. The wavepacket spends no time inside the barrier. This can be traced back to the group

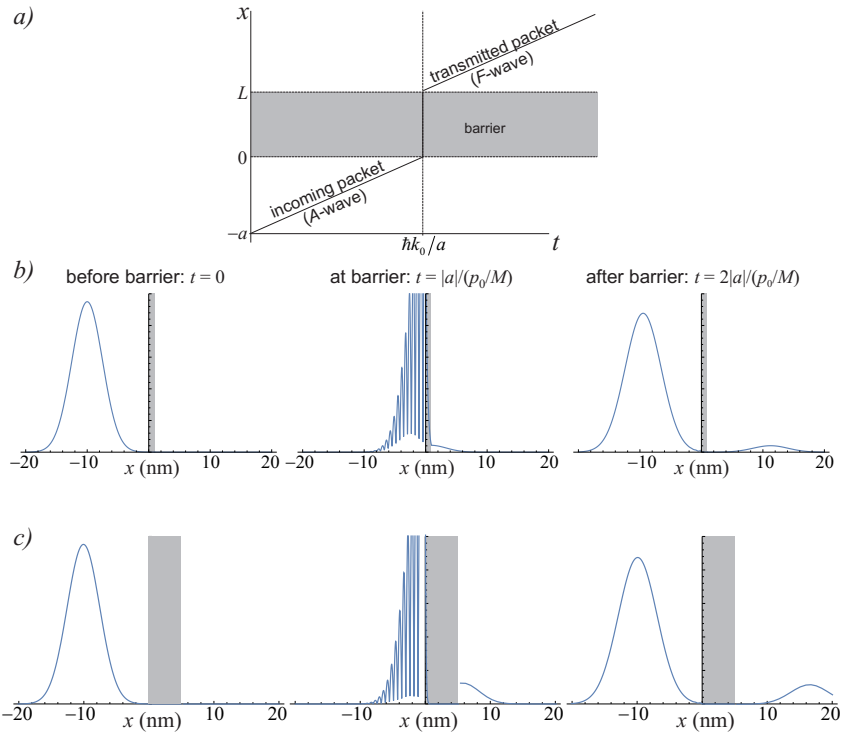


Fig. 3.7 Superluminal tunnelling of a Gaussian wavepacket (Ex. 3.53). a) Schematic diagram. b) Numerical simulation for an electron starting at $a = -10$ nm, with the same parameters as in Fig. 3.6. c) Numerical simulation with the same parameters, except $L = 5$ nm and $V_0 = 5.66$ eV; the section to the right of the barrier is magnified by a factor of 10^{23}. The center of the transmitted wavepacket emerges from the barrier at the same time as the center of the incident wavepacket enters it.

velocity's being the derivative $v_{\mathrm{gr}} = \mathrm{d}\omega/\mathrm{d}k$. Inside the barrier, the wavefunction consists of real exponentials (C- and D-waves in Fig. 3.6), and hence has a constant complex phase. This means that the effective wavevector $k = 0$ and the group velocity is infinite.

In Chapter 2 we already encountered a quantum phenomenon that appeared to enable faster-than-light communication, but after careful analysis found this to be only an illusion. In the present case, the superluminal group velocity is a correct conclusion. However, once again, it does not imply the possibility of instant signaling, for the following reason.

We found the velocity of the *center* of the wavepacket to be infinite. But let us ask ourselves: at what moment does an observer behind the barrier learn that the particle is entering the barrier? Does it have to happen when exactly a half of the wavepacket has emerged from the barrier, or is it one-fourth or perhaps one-tenth?

The correct answer is, much earlier than that. From complex analysis, we know that the Gaussian function is *analytic*: any fragment of this function allows one to reconstruct its behavior in the entire complex plane. Therefore, theoretically, any observer anywhere in space and time is aware of the presence of the particle with a Gaussian wavefunction, and can predict its evolution. With this in mind, it makes no sense to talk about instant communication.

What if we tried a different wavefunction, for example, with the top-hat shape (3.9), which takes on nonzero values only within a finite spatial region? The problem with such wavefunctions is that we cannot apply the approximations we used for the Gaussian wavepacket (see Ex. 3.51). The feature of a Gaussian wavepacket which allows us to use these approximations is that its momentum representation is also Gaussian, so it falls off exponentially away from the center point. In contrast, states with spatially limited wavepackets are not narrow in the momentum representation: for example, the Fourier transform of the top-hat function is the sinc function [Ex. D.9(f)]. This means that such a state will have significant components corresponding to arbitrarily high energies: not only above the barrier, but also extending into the relativistic domain. Hence the formalism of non-relativistic quantum mechanics, which we are studying here, cannot be applied to this problem.

To complete our study of the barrier potential, let us see what happens when the particle energy is above the barrier. For generality, we will allow V_0 to be either positive or negative, which corresponds to the cases of a barrier or a well, respectively.

Exercise 3.54. Solve Ex. 3.52 for $E > 0$ and $E > V_0$.
Answer:

$$\text{Transmission: } \frac{j_F}{j_A} = \frac{4k_0^2k_1^2}{4k_0^2k_1^2\cos^2(k_1L) + (k_1^2+k_0^2)^2\sin^2(k_1L)}; \tag{3.82a}$$

$$\text{Reflection: } \frac{j_B}{j_A} = \frac{(k_1^2-k_0^2)^2\sin^2(k_1L)}{4k_0^2k_1^2\cos^2(k_1L) + (k_1^2+k_0^2)^2\sin^2(k_1L)}, \tag{3.82b}$$

where $k_0 = \sqrt{2ME}/\hbar$ and $k_1 = \sqrt{2M(E-V_0)}/\hbar$.

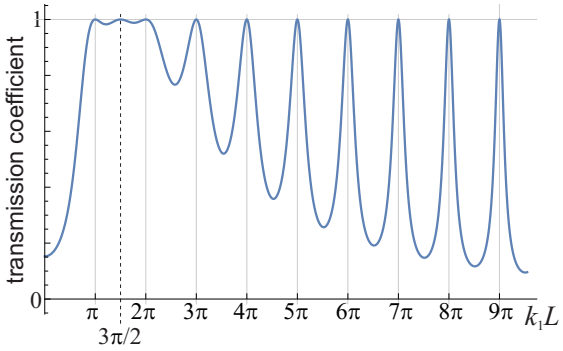


Fig. 3.8 Transmission coefficient (3.82)(a) of the potential barrier for the case of the particle's energy above the barrier, specialized to $k_0L = 3\pi/2$.

Exercise 3.55. Under which conditions is the transmission coefficient in the previous exercise equal to one?
Answer: $V_0 = 0$ (i.e. $k_0 = k_1$) or $k_lL = m\pi$, where m is a positive integer.

We see that the transmissivity of the potential barrier (or well, if $V_0 < 0$) exhibits oscillatory behavior and becomes unity when the thickness of the barrier corresponds to an integer or half-integer number of de Broglie waves inside the barrier. Again, there is an immediate analogy to optics: the optical cavity, also known as the Fabry-Pérot etalon. In such a cavity, the optical wave bounces between two highly reflective interfaces, and the multiple reflections interfere with each other. If the roundtrip distance $2L$ of the wave in the interferometer is an integer number of wavelengths (i.e., $2L = m\lambda = \frac{2\pi m}{k_1}$), this interference becomes constructive, and the transmissivity of the cavity becomes unity.

We can also see that the width of each resonance decreases with k_1. This is due to the increasing reflectivity of each "cavity mirror", which is given by the first part of Eq. (3.78a) and Eq. (3.79b) for quantum mechanics and optics, respectively. The closer the reflectivity to one, the higher the finesse of the etalon, the sharper the resonance.

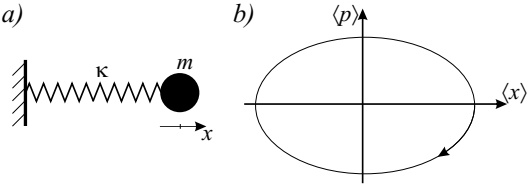
3.8 Harmonic oscillator

The harmonic oscillator is a physical system of primary importance, with applications extending far beyond pure mechanics. In fact, virtually *any* oscillatory motion is governed by a Hamiltonian that is analogous to that of the mechanical harmonic oscillator, and thus has the same quantum description. Examples include the electromagnetic field, inductor–capacitor circuits in electronics, and quasiparticles in condensed matter. Even the photon, which we discussed frequently in the pre-

Box 3.10 The classical harmonic oscillator

Figure (a) below displays the simplest harmonic oscillator — a "ball on a spring". When the ball is displaced from the equilibrium position $x = 0$, the spring will exert a force $F = -\kappa x$ according to Hooke's law, where κ is the spring constant. The potential of the spring tension is then $U(x) = \kappa x^2/2$, which corresponds to the Hamiltonian

$$H = \frac{p^2}{2M} + \frac{\kappa x^2}{2}. \tag{3.83}$$



A classical harmonic oscillator. a) Physical model; b) motion in the phase space.

When left alone, the ball will obey the equations of motion

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{p}{M}, \tag{3.84a}$$

$$\frac{\mathrm{d}p}{\mathrm{d}t} = -\kappa x, \tag{3.84b}$$

which give rise to the oscillation with frequency $\omega = \sqrt{\kappa/M}$:

$$x(t) = x(0)\cos\omega t + \frac{1}{M\omega}p(0)\sin\omega t, \tag{3.85a}$$

$$p(t) = p(0)\cos\omega t - M\omega x(0)\sin\omega t. \tag{3.85b}$$

This classical motion of the oscillator can be represented as a trajectory in the *phase space*, i.e., as a parametric plot of its momentum vs. position, as shown in part (b) of the figure above. This trajectory has an elliptical shape with half-axis ratio given by $p_{max} = M\omega x_{max}$.

vious two chapters, can be seen as an energy eigenstate of the quantum harmonic oscillator describing a mode of the light field.

Exercise 3.56. Check that the solution for the classical equations of motion (3.84) of the harmonic oscillator is given by Eqs. (3.85).

3.8.1 Annihilation and creation operators

The harmonic oscillator potential is a typical well. Therefore its energy eigenstates are bound and nondegenerate (see Ex. 3.46). The wavefunctions of these states can be found by solving the time-independent Schrödinger equation (3.60) in the posi-

tion basis. However, the harmonic oscillator permits a special, much more elegant theoretical treatment. To develop this treatment, we first rescale the position and momentum observables to make them more convenient to work with.

Exercise 3.57. Find the proportionality constants A and B such that the observables defined by $X = Ax$, $P = Bp$ have the following properties:

- In the new variables (X, P), the phase space trajectory is circular, so Eqs. 3.85 take the form

$$X(t) = X(0)\cos\omega t + P(0)\sin\omega t; \tag{3.86a}$$
$$P(t) = -X(0)\sin\omega t + P(0)\cos\omega t. \tag{3.86b}$$

- For the corresponding quantum operators,

$$[\hat{X}, \hat{P}] = \mathrm{i}. \tag{3.87}$$

Show that the rescaled observables $\hat{X}$ and $\hat{P}$ are dimensionless.
Answer:

$$\hat{X} = \hat{x}\sqrt{\frac{M\omega}{\hbar}}; \quad \hat{P} = \frac{p}{\sqrt{M\omega\hbar}} \tag{3.88}$$

Being continuous observables, the rescaled position and momentum eigenstates are normalized according to

$$\langle X \big| X' \rangle = \delta(X - X'); \quad \langle P \big| P' \rangle = \delta(P - P'). \tag{3.89}$$

As we observed in Sec. 3.2, rescaling a continuous observable, while imposing normalization conditions like those above, results in the renormalization of eigenstates of these observables, as well as the wavefunctions and operators written in terms of these eigenstates. Let us see how this applies in the present case.

Exercise 3.58. a) Show that the eigenstates of the canonical and rescaled observables are related according to

$$|X\rangle = \left(\frac{\hbar}{M\omega}\right)^{1/4} |x\rangle; \tag{3.90a}$$
$$|P\rangle = (M\hbar\omega)^{1/4} |p\rangle. \tag{3.90b}$$

Hint: use the same argument as at the end of Sec. 3.2, where we related the position and wavenumber observables.

b) Show that

$$\langle X | P \rangle = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{\mathrm{i}PX}. \tag{3.91}$$

c) If a certain quantum state has wavefunctions $\psi(x) = \langle x | \psi \rangle$ and $\tilde{\psi}(p) = \langle p | \psi \rangle$, what are the corresponding wavefunctions $\psi(X) = \langle X | \psi \rangle$ and $\tilde{\psi}(P) = \langle P | \psi \rangle$ in the rescaled variables?

d) Show that the relations for converting wavefunctions between the $\hat{X}$- and $\hat{P}$-bases are

$$\langle X|\,\psi\rangle = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\langle P|\,\psi\rangle\,\mathrm{e}^{\mathrm{i}PX}\mathrm{d}P; \tag{3.92}$$

$$\langle P|\,\psi\rangle = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\langle X|\,\psi\rangle\,\mathrm{e}^{-\mathrm{i}PX}\mathrm{d}X. \tag{3.93}$$

e) Show that

$$\langle X|\,\hat{P}\,|\psi\rangle = -\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}X}\psi(X); \quad \langle P|\hat{X}\,|\psi\rangle = \mathrm{i}\frac{\mathrm{d}}{\mathrm{d}P}\psi(P). \tag{3.94}$$

f) Show that the Heisenberg uncertainty principle for the rescaled position and momentum takes the form

$$\left\langle \Delta X^2\right\rangle \left\langle \Delta P^2\right\rangle \geq \frac{1}{4}. \tag{3.95}$$

Exercise 3.59. Write the Hamiltonian (3.83) in terms of the rescaled observables $\hat{X}$ and $\hat{P}$.

Answer:

$$\hat{H} = \frac{1}{2}\hbar\omega\left(\hat{X}^2+\hat{P}^2\right). \tag{3.96}$$

We now define and study the properties of the two operators which, as we shall see in the next subsection, effect transitions between adjacent energy eigenstates. The *annihilation* operator is defined as follows:

$$\hat{a} = \frac{1}{\sqrt{2}}\left(\hat{X}+\mathrm{i}\hat{P}\right); \tag{3.97}$$

The operator $\hat{a}^\dagger$ is called the *creation* operator.

Exercise 3.60. Show that:

a) the creation operator is

$$\hat{a}^\dagger = \frac{1}{\sqrt{2}}\left(\hat{X}-\mathrm{i}\hat{P}\right); \tag{3.98}$$

b) the creation and annihilation operators are not Hermitian;

c) their commutator is

$$[\hat{a},\hat{a}^\dagger] = 1; \tag{3.99}$$

d) the position and momentum can be expressed as

$$\hat{X} = \frac{1}{\sqrt{2}}\left(\hat{a}+\hat{a}^\dagger\right); \quad \hat{P} = \frac{1}{\mathrm{i}\sqrt{2}}\left(\hat{a}-\hat{a}^\dagger\right); \tag{3.100}$$

e) the commutation relations of the creation and annihilation operators are

$$[\hat{a}, \hat{a}^\dagger \hat{a}] = \hat{a}; \quad [\hat{a}^\dagger, \hat{a}^\dagger \hat{a}] = -\hat{a}^\dagger; \tag{3.101}$$

f) the Hamiltonian (3.96) can be written as

$$\hat{H} = \hbar\omega \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right). \tag{3.102}$$

3.8.2 Fock states

Our next goal is to find eigenvalues and eigenstates of the Hamiltonian. Because of Eq. (3.102), they are also eigenstates of the operator $\hat{a}^\dagger \hat{a}$. This is called the *number operator* and denoted by the symbol $\hat{n}$. A normalized eigenstate of this operator with eigenvalue n is denoted by $|n\rangle$:

$$\hat{a}^\dagger \hat{a} \, |n\rangle = n \, |n\rangle \tag{3.103}$$

Exercise 3.61. Show that:

a) the state $\hat{a} \, |n\rangle$ is also an eigenstate of $\hat{n}$ with eigenvalue $n-1$;
b) the state $\hat{a}^\dagger \, |n\rangle$ is also an eigenstate of $\hat{n}$ with eigenvalue $n+1$.

Hint: Use Eq. (3.101).

We know from Ex. 3.46 that energy spectra of bound states are nondegenerate, i.e., for each value of n there exists no more than a single energy eigenstate $|n\rangle$. Hence we can conclude from Ex. 3.61 that the states $\hat{a} \, |n\rangle$ and $\hat{a}^\dagger \, |n\rangle$ are proportional to the states $|n-1\rangle$ and $|n+1\rangle$, respectively. Note that I say "proportional" rather than "equal", because the states $\hat{a} \, |n\rangle$ and $\hat{a}^\dagger \, |n\rangle$ are not guaranteed to be normalized, whereas $|n-1\rangle$ and $|n+1\rangle$ are normalized by definition. In fact, the normalization condition is used to determine the proportionality coefficient.

Exercise 3.62. Using the fact that all energy eigenstates must be normalized to 1, show that, up to an arbitrary phase factor,

a)

$$\hat{a} \, |n\rangle = \sqrt{n} \, |n-1\rangle \, ; \tag{3.104a}$$

b)

$$\hat{a}^\dagger \, |n\rangle = \sqrt{n+1} \, |n+1\rangle \, . \tag{3.104b}$$

Hint: use $\langle n| \, \hat{a}^\dagger \hat{a} \, |n\rangle = n$.

The phase factor mentioned in the exercise above is our choice; we can define it however we want. We choose the simplest option and define it to equal 1, so the relations (3.104) are valid as they are.

Equation (3.104a) tells us that, if the state $|n\rangle$ with energy $\hbar\omega(n+1/2)$ exists as a physical state (i.e., if it is a normalized element of the Hilbert space), so does the state $|n-1\rangle$ with energy $\hbar\omega(n-1/2)$. Similarly, the states $|n-2\rangle$, $|n-3\rangle$, and so on must also exist. Continuing this chain for sufficiently many steps, we would end up with energy eigenstates with negative energy values. This is impossible because the Hamiltonian is a non-negative operator (Ex. A.72, A.87).

How can we resolve this contradiction? The only way is to assume that *n must be a nonnegative integer* so that the chain is broken at $n = 0$, in which case

$$\hat{a}\,|0\rangle = |\text{zero}\rangle\,. \tag{3.105}$$

Then (provided that the state $|n=0\rangle$ exists), number eigenstates exist only for non-negative n, comprising an infinite set with the corresponding energy eigenvalues $\hbar\omega(n+1/2)$.

Energy eigenstates of a harmonic oscillator are called *Fock* or *number* states. The state $|0\rangle$ is called the *vacuum state*[7].

Exercise 3.63. Express $|n\rangle$ in terms of $|0\rangle$.
Answer:

$$|n\rangle = \frac{\left(\hat{a}^\dagger\right)^n}{\sqrt{n!}}\,|0\rangle \tag{3.106}$$

Exercise 3.64. Calculate the wavefunctions of the vacuum state in the position and momentum representations.
Hint: use Eqs. (3.94), (3.97) and (3.105).
Answer:

$$\psi_0(X) = \frac{1}{\pi^{1/4}}\mathrm{e}^{-X^2/2}; \tag{3.107a}$$

$$\tilde{\psi}_0(P) = \frac{1}{\pi^{1/4}}\mathrm{e}^{-P^2/2}. \tag{3.107b}$$

As we can see, both the position and momentum observables are uncertain in the vacuum state. That is, if we prepare our "ball on a spring" in the state of minimum possible energy, and then measure its position, we will find it to be away from the equilibrium position by a random microscopic amount. Similarly, if we measure the velocity, we will find it to be microscopically small, but nonzero. This quantum phenomenon is known as the *zero-point vibration*.

The above wavefunctions are unique up to an arbitrary overall phase factor. For the vacuum state, by convention, we choose this factor so as to obtain a real and positive-definite wavefunction in the position basis. It then automatically follows that the wavefunction in the momentum basis is also real and positive. Furthermore,

[7] Let me emphasize the difference between the vectors $|0\rangle$ and $|\text{zero}\rangle$ (see Defn. A.1). The vector $|\text{zero}\rangle$ is the zero of the Hilbert space, such that, for any vector $|\psi\rangle$, we have $|\psi\rangle + |\text{zero}\rangle = |\psi\rangle$. Its norm is $\langle\text{zero}|\ \text{zero}\rangle = 0$ so this vector does not represent any physical quantum state. The vacuum $|0\rangle$, on the other hand, is a physical state: $\langle 0|\ 0\rangle = 1$ and $|\psi\rangle + |0\rangle \neq |\psi\rangle$.

as we shall see below, this convention ensures that the wavefunctions of all other Fock states are also real.

Having explicitly found the wavefunction of the vacuum state, we have proven its existence and uniqueness, and thus, automatically, the existence and uniqueness of all other Fock states — because these states are obtained from the vacuum state by applying the creation operator.

Exercise 3.65. a) By applying Eq. (3.106), calculate the wavefunctions of the Fock states $|1\rangle$ and $|2\rangle$.

Answer:

$$\psi_1(X) = \frac{\sqrt{2}}{\pi^{1/4}} X \mathrm{e}^{-X^2/2}; \tag{3.108}$$

$$\psi_2(X) = \frac{1}{\sqrt{2}\pi^{1/4}} (2X^2 - 1)\mathrm{e}^{-X^2/2}. \tag{3.109}$$

b)* Show that the wavefunction of an arbitrary Fock state $|n\rangle$ is given by

$$\psi_n(X) = \frac{H_n(X)}{\pi^{1/4}\sqrt{2^n n!}} \mathrm{e}^{-X^2/2}, \tag{3.110}$$

where $H_n(X)$ are the Hermite polynomials,

$$H_n(X) = \left(2X - \frac{\mathrm{d}}{\mathrm{d}X}\right)^n 1. \tag{3.111}$$

A special feature of the Hamiltonian of the harmonic oscillator is that its energy levels not only quantize, but are equidistant. The distance $\hbar\omega$ between the levels is called the *quantum* of energy. Physically, the equidistant energy structure means that, by pumping quanta of the same frequency into a harmonic oscillator, one can excite it to arbitrarily high energy. For example, you can rock a swing to any desired amplitude by pulling and pushing it with the same frequency; you can also amplify a laser pulse to any desired strength. An opposite example is the atom: by driving it with a resonant laser, you can bring it from a ground state to one of the higher energy eigenstates, but increasing the laser power is not likely to help you excite that atom any further[8].

Energy quanta are often interpreted as particles, especially in the context of generalizations of the harmonic oscillator mentioned at the beginning of this section. For example, the *photon* is an energy quantum in an optical pulse (see Box 3.11) and the *phonon* is the energy quantum in a vibrational mode of a solid state.

It is instructive to compare the wavefunctions of the Fock states with those of energy eigenstates of the finite potential well (shown in Fig. 3.2). In both cases, the wavefunctions exhibit oscillatory behavior inside the well and fall off exponentially

[8] These statements are valid within limits, because physical models of the harmonic oscillator or two-level system may break down for very strong excitations. This happens, for example, if a swing flies too high to invalidate the small-angle pendulum approximation, or if the electric field in a laser pulse becomes so strong that it starts to give birth to electron–positron pairs.

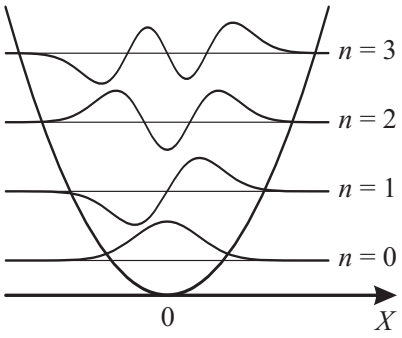


Fig. 3.9 Wavefunctions of the few lowest energy levels of a harmonic oscillator.

outside. The number of zero crossings is equal to the number of the energy level. The difference is that the energy levels are equidistant for the harmonic oscillator,

Box 3.11 What is the photon?

In the previous two chapters, we treated the photon as a *particle* and discussed the quantum states it can be found in. Now we seem to be saying that the photon is a *state* of the electromagnetic harmonic oscillator mode. How can these views be reconciled?

These two approaches are known as *first quantization* and *second quantization*, respectively. In the first quantization, we associate a Hilbert space with each particle; the elements (vectors) of this space are the various states that the particle can be in. For example, a single photon (Hilbert space) can be in different polarization states.

In the second quantization, the roles of state vectors and Hilbert spaces are exchanged. What we call a basis of the Hilbert space of the first quantization is treated as a set of separate Hilbert spaces in the second quantization. For example, the vertical and horizontal polarization modes are treated as individual Hilbert spaces. A photon in the state $|H\rangle$ in the first quantization is written as the state vector $|1\rangle_H \otimes |0\rangle_V$ in the second. A photon in the state $|+45^\circ\rangle$ becomes the entangled state $\frac{1}{\sqrt{2}}(|1\rangle_H \otimes |0\rangle_V + |0\rangle_H \otimes |1\rangle_V)$. Alternatively, we can choose the two diagonal polarization modes as Hilbert spaces; in this case, the diagonally polarized photon is a separable state while the horizontally polarized photon is an entangled one.

We thus see that the first quantization treatment is more compact and convenient when we know *a priori* that we are dealing with exactly one particle. In the case of multiple identical particles, the first quantization gives rise to complications. For example, suppose we have two photons with orthogonal polarizations. In the framework of the second quantization, we have a unique way of writing this state: $|1\rangle_H \otimes |1\rangle_V$. Using the first quantization, on the other hand, we could write this state — the same physical object — in two possible ways: $|H\rangle \otimes |V\rangle$ or $|V\rangle \otimes |H\rangle$, or any linear combination thereof. In order to eliminate this ambiguity, additional rules need to be introduced as to which state vector can be considered physical, depending on whether the particle is fermionic or bosonic.

To summarize, while both approaches are valid and can be used to treat physical phenomena, one or the other may turn out to be more practical, depending on the problem we are trying to solve.

but not equidistant for the rectangular well. Further, each eigenwavefunction of the well is defined in a piecewise fashion [see Eqs. (3.66) and (3.67)], while it is a single elementary function for the harmonic oscillator potential.

Exercise 3.66. Calculate the matrices of the position and momentum observables in the Fock basis.
Hint: Rather than integrating the wavefunctions, it is more convenient to employ Eqs. (3.100) and (3.104).

Exercise 3.67. For an arbitrary $|n\rangle$, calculate $\langle X\rangle$, $\langle \Delta X^2\rangle$, $\langle P\rangle$, $\langle \Delta P^2\rangle$ and check the uncertainty principle.
Answer:

$$\left\langle \hat{X}\right\rangle = \left\langle \hat{P}\right\rangle = 0; \tag{3.112}$$

$$\left\langle \Delta X^2\right\rangle = \left\langle \Delta P^2\right\rangle = \frac{1}{2}(2n+1). \tag{3.113}$$

We see that the product of the position and momentum uncertainties increases with the energy. The vacuum state is the only Fock state for which this product reaches the minimum (3.95).

Exercise 3.68. Consider the Schrödinger evolution $|\psi(t)\rangle$ of an arbitrary state $|\psi(0)\rangle = \sum_n \psi_n |n\rangle$ under the harmonic oscillator Hamiltonian. Derive the following behavior of operator mean values as a function of time:

a)

$$\langle \hat{a}\rangle (t) = \langle \hat{a}\rangle (0) \mathrm{e}^{-\mathrm{i}\omega t}; \tag{3.114a}$$

$$\left\langle \hat{a}^\dagger\right\rangle (t) = \left\langle \hat{a}^\dagger\right\rangle (0) \mathrm{e}^{\mathrm{i}\omega t}; \tag{3.114b}$$

b)

$$\langle X\rangle (t) = \langle X\rangle (0) \cos\omega t + \langle P\rangle (0) \sin\omega t; \tag{3.115a}$$

$$\langle P\rangle (t) = -\langle X\rangle (0) \sin\omega t + \langle P\rangle (0) \cos\omega t. \tag{3.115b}$$

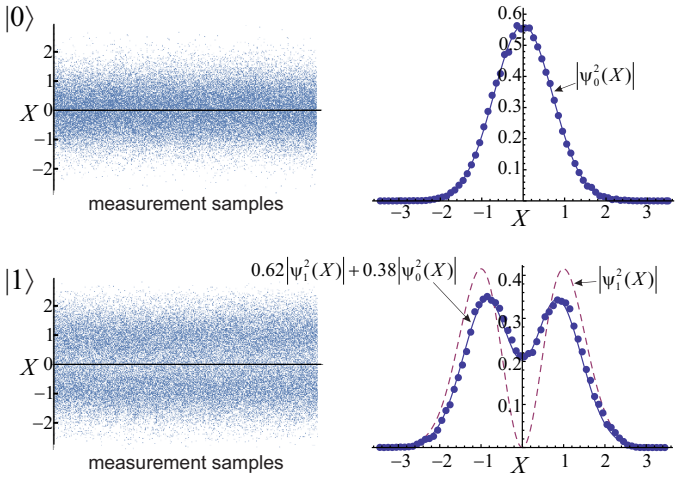
The Fock states themselves are stationary and thus exhibit no time variation of the mean position and momentum. In this sense, they are highly nonclassical, inconsistent with our familiar notion that a ball on a spring must oscillate (unless it is at rest, i.e., in the lowest energy state). On the other hand, in all other states, the mean position and momentum do change. Remarkably, in *any* quantum state, they change in the same way as the position and momentum in the classical harmonic oscillator [see Eq. (3.86)]. We will generalize this observation in Sec. 3.9.

Box 3.12 Measuring the position of the harmonic oscillator: an experiment

At the time of writing this manuscript, physicists are not yet able to prepare and measure arbitrary quantum states of mechanical oscillators. They have much better control of the optical realization of the harmonic oscillator. In particular, they are able to prepare a few of the lowest number states and their superpositions with high fidelity.

In the optical realization of the harmonic oscillator, the position and momentum observables correspond to the magnitudes of the electric field in the electromagnetic wave at certain phases. Phase-sensitive measurements of the electromagnetic field are performed using the so-called *optical homodyne detector*. I will not go into a detailed description of this technology, but it can be found in many quantum optics textbooks.

The figure below shows the experimental data on multiple measurements of the position observable in the vacuum state (top) and the single-quantum state (bottom) of an optical mode. The vacuum state is obtained by simply blocking the light; the heralded single photon is prepared using parametric down-conversion (Box 1.6). The histograms (right) of the raw experimental data (left) are expected to correspond to the probability densities $\mathrm{pr}_{0,1} = |\psi_{0,1}(X)|^2$, where the wavefunctions $\psi_{0,1}(X)$ are given by Eqs. (3.107a) and (3.108), respectively.



Data taken from A. I. Lvovsky and S. A. Babichev, *Synthesis and tomographic characterization of the displaced Fock state of light*, Physical Review A **66**, 011801 (2002).

We can see that, while the agreement between theory and experiment is almost perfect for the vacuum state, the data for the single-photon state are best fit with a mixed state of the single photon with probability 0.62 and vacuum with probability 0.38. This is because it is impossible to create a perfect single-photon state. The measured state fidelity is inevitably diminished by losses in the optical path, imperfect detection efficiency, and other issues.

3.8.3 Coherent states

The coherent state is the closest quantum approximation of the classical picture of harmonic oscillator motion. As we have seen, the mean position and momentum in any quantum state oscillate as functions of time in the same way as those of a classical ball on a spring. A special feature of the coherent state is that, while the amplitude of this oscillation can be arbitrarily high, the position and momentum uncertainties remain as low as in the vacuum state. Because of their classical-like behavior, coherent states frequently occur in nature, not only in mechanics, but also in other "incarnations" of the harmonic oscillator, such as the quantum state of the light field in a laser pulse.

The *coherent (Glauber) state* $|\alpha\rangle$ is an eigenstate of the annihilation operator with eigenvalue α:

$$\hat{a}\,|\alpha\rangle = \alpha\,|\alpha\rangle\,. \tag{3.116}$$

Because $\hat{a}$ is not a Hermitian operator, its eigenvalue α can be complex. Its absolute value $|\alpha|$ is called the *amplitude*, and the complex argument $\operatorname{Arg}\alpha$ the *coherent phase* of the coherent state.

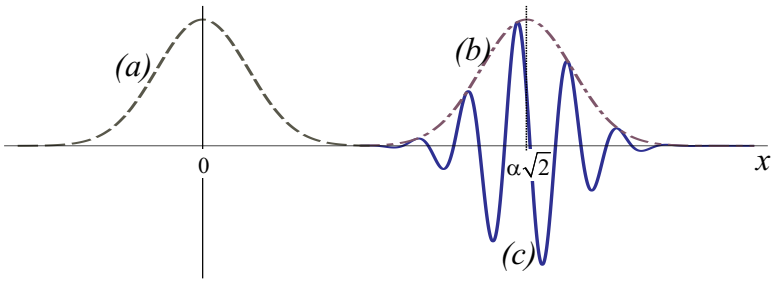


Fig. 3.10 Wavefunctions of coherent states. a) $\alpha = 0$ (vacuum state), b) $\alpha = 5$, c) $\alpha = 5 + 4\mathrm{i}$ (real part).

We begin our studies of the coherent state from its wavefunction. The wavefunction can be determined by solving Eq. (3.116) as a differential equation in the position basis akin to Ex. 3.64. However, to avoid this fairly tedious calculation, an easy answer is provided in the exercise below. We will develop an alternative way to calculate the wavefunction of the coherent state in Sec. 3.10.

Exercise 3.69. For a coherent state $|\alpha\rangle$, show that its wavefunctions in the position and momentum bases are given by

$$\psi_\alpha(X) = \frac{1}{\pi^{1/4}} \mathrm{e}^{-\mathrm{i}\frac{P_\alpha X_\alpha}{2}} \mathrm{e}^{\mathrm{i}P_\alpha X} \mathrm{e}^{-\frac{(X-X_\alpha)^2}{2}}; \tag{3.117a}$$

$$\tilde{\psi}_\alpha(P) = \frac{1}{\pi^{1/4}} \mathrm{e}^{\mathrm{i}\frac{P_\alpha X_\alpha}{2}} \mathrm{e}^{-\mathrm{i}X_\alpha P} \mathrm{e}^{-\frac{(P-P_\alpha)^2}{2}}, \tag{3.117b}$$

where

$$X_\alpha = \sqrt{2}\mathrm{Re}\,\alpha; \quad P_\alpha = \sqrt{2}\mathrm{Im}\,\alpha. \tag{3.118}$$

Check that these wavefunctions are normalized. Show that the expectation values and uncertainties of the position and momentum in the coherent state $|\alpha\rangle$ are

$$\langle X\rangle = X_\alpha, \langle P\rangle = P_\alpha \tag{3.119}$$

and

$$\langle \Delta X^2\rangle = \langle \Delta P^2\rangle = 1/2, \tag{3.120}$$

respectively.

The coherent state wavefunction is a Gaussian wavepacket. For $\alpha = 0$, the coherent state becomes the vacuum, as is evident by comparing Eqs. (3.105) and (3.116) [Fig. 3.10(a)]. For real α, the shape of the wavefunction is identical to that of the vacuum state, shifted by $\alpha\sqrt{2}$ along the x axis [Fig. 3.10(b)]. For complex α, this Gaussian wavepacket is multiplied by a linearly varying phase factor, because of the nonzero mean momentum [Fig. 3.10(c)], akin to Ex. 3.25.

We see that there exists a coherent state for any complex α and that each such state can be normalized according to $\langle\alpha|\,\alpha\rangle = 1$. This may seem to contradict our discussion early in this chapter, where we argued that eigenstates of continuous quantum observables must be normalized via the Dirac delta function, as in Eqs. (3.1). The reason why this rule does not apply to coherent states is that the annihilation operator is not a Hermitian observable. For the same reason, coherent states associated with different values of α are not orthogonal (see Ex. 3.75 below).

As per Eq. (3.120), any coherent state has the lowest possible position–momentum uncertainty (3.95), similarly to the vacuum state.

One can picture the coherent state in the phase space as a circle centered at the point ($\langle X\rangle = \sqrt{2}\mathrm{Re}\,\alpha$, $\langle P\rangle = \sqrt{2}\mathrm{Im}\,\alpha$) (Fig. 3.11). The radius of the circle, $1/\sqrt{2}$, symbolically represents the standard deviations of the position and momentum, which are independent of the coherent amplitude[9].

The overall phase factors $\mathrm{e}^{\pm\mathrm{i}P_\alpha X_\alpha/2}$ are included in Eqs. (3.117a) and (3.117b) as a matter of convention. These factors make these two equations (which are obtained from one another by means of a direct or inverse Fourier transform) look similar. Additionally, this convention is necessary for consistency with another phase convention that we introduce below for the decomposition of the coherent state into the Fock basis.

Let me emphasize the role of the phase $\mathrm{Arg}\,\alpha$ of the coherent state. This phase is the angle of the radius vector pointing to $(\langle X\rangle, \langle P\rangle)$, as shown in Fig. 3.11, and is thus directly related to measurable parameters of the state. This is in contrast to the *quantum* phase factor, which, as we have discussed several times already, does not affect a state's physical properties.

[9] In fact, this circle has more than just a symbolic value. The behavior of uncertainties in the phase space is described by the so-called *Wigner function*, which is the analog of the classical phase-space probability density.

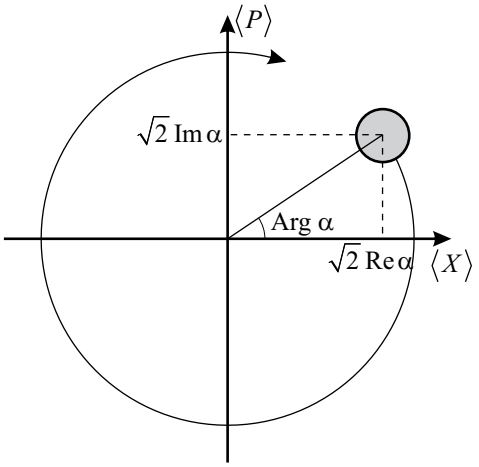


Fig. 3.11 Phase space picture and evolution of the position and momentum observables in the coherent state.

Now let us find the Schrödinger evolution of the coherent state in time. To this end, we first decompose the coherent state in the energy eigenbasis.

Exercise 3.70. Find the decomposition of the coherent state $|\alpha\rangle$ in the number basis. **Hint:** Assume some decomposition,

$$|\alpha\rangle = \sum_{n=0}^{\infty} \alpha_n |n\rangle , \tag{3.121}$$

and apply the definition (3.116) of the coherent state to it.
Answer: up to an overall phase factor,

$$|\alpha\rangle = \mathrm{e}^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle . \tag{3.122}$$

Once again, we introduce an overall phase convention according to which the overall phase factor in Eq. (3.122) is 1; that is, we set $\langle n|\ \alpha\rangle$ real for real α. We need to check whether this convention is consistent with the one chosen for the phase of the coherent state wavefunction (3.117a).

Exercise 3.71. Calculate the inner product $\langle 0|\ \alpha\rangle$ for an arbitrary α in the position and Fock bases. Check whether the results are the same.

If one performs an energy measurement on a coherent state, the probabilities of the possible results are distributed according to

$$\mathrm{pr}_n = |\langle n|\ \alpha\rangle|^2 = \mathrm{e}^{-|\alpha|^2}\frac{|\alpha|^{2n}}{n!}. \tag{3.123}$$

This turns out to be the famous Poisson distribution (see Sec. B.3). From its properties (Ex. B.15), we see that both the mean and the variance of the Fock number in the coherent state are equal to

$$\langle n\rangle = \left\langle \Delta n^2\right\rangle = |\alpha|^2. \tag{3.124}$$

This means, for example, that in a laser pulse train with n photons per pulse on average, the root mean square uncertainty of that photon number in a pulse is $\sqrt{n}$.

Actually, we don't need to know the properties of the Poisson distribution to obtain the latter result. It follows directly from the definition of the coherent state.

Exercise 3.72. Calculate the mean and variance of the Hamiltonian operator (3.102) in the coherent state using the properties of the creation and annihilation operators and check that your result is consistent with Eq. (3.124).

Exercise 3.73. Show that the action of the evolution operator $\exp(\mathrm{i}\hat{H}t/\hbar)$ on the state $|\alpha\rangle$ is given by

$$\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\hat{H}t}|\alpha\rangle = \mathrm{e}^{-\mathrm{i}\omega t/2}\left|\alpha\mathrm{e}^{-\mathrm{i}\omega t}\right\rangle. \tag{3.125}$$

Exercise 3.74. Calculate the expectation values of

a) the creation and annihilation operators;
b) the position and momentum observables

in a coherent state as a function of time using Eqs. (3.119) and (3.125). Check that your result is consistent with Eqs. (3.114) and (3.115).

The result of Ex. 3.73 is remarkable. Neglecting the unphysical quantum phase factor, a coherent state evolves into another coherent state with the same amplitude, but different coherent phase, as shown in Fig. 3.11. This means that the position and momentum uncertainties remain constant, equal to those of the vacuum state.

This result illustrates once again the difference between the quantum and coherent phases. The *quantum* phase factor $\mathrm{e}^{-\mathrm{i}\omega t/2}$, outside of the ket in Eq. (3.125), has no physical counterpart. The *coherent* phase $\mathrm{e}^{-\mathrm{i}\omega t}$, which has observable physical meaning, is inside the ket.

Finally, Ex. 3.73 makes manifest the classical analogy of large-amplitude coherent states. If the amplitude of the coherent state is macroscopic, the relative uncertainties are negligible, so the coherent state is well approximated by a classical oscillation. In contrast, for microscopic amplitudes, the uncertainties play a significant role and the classical approximation fails.

Exercise 3.75. Show that $\langle\alpha|\ \alpha'\rangle = \mathrm{e}^{-|\alpha|^2/2-|\alpha'|^2/2+\alpha'\alpha^*}$.

This result addresses once again our earlier point that coherent states associated with different values of α are not orthonormal. Because the annihilation operator is

not Hermitian, the spectral theorem (Ex. A.60), which states that the eigenstates of a Hermitian operator constitute an orthonormal basis, does not apply here. Coherent states do form a spanning set, but they are not orthogonal.

Exercise 3.76. Coherent states are eigenstates of the annihilation operator. Do their counterparts — eigenstates of the creation operator — exist, and if so, what is their decomposition in the number basis?

3.9 Heisenberg picture

By now, we have encountered a few cases in which quantum mechanics predicts behavior that is expected classically. Examples are the evolution of the mean position and momentum in free space or under the harmonic oscillator potential. Such observations are in principle not surprising because we know that the classical picture corresponds to the macroscopic limit of the quantum physics. But at the same time, the theoretical and mathematical frameworks of the two treatments are so different that, even if they do lead to similar results, it is difficult to understand the intuition behind this similarity.

If we attempt to reconcile the two treatments and find a common intuitive ground beneath them, one obstacle that we inevitably encounter is in the way that classical and quantum physics treat temporal evolution. In the classical picture, it is the observables that evolve: for example, the position of a moving particle changes with time. In the quantum world, in contrast, observables, such as the position operator $\hat{x}$, are constant; the evolution is associated with the system's state $|\psi(t)\rangle$. In this section, we shall make the connection between the two worlds more transparent by developing an alternative quantum formalism, in which states are constant and observables evolve.

3.9.1 Operator evolution

Suppose that we need to find the mean value of a certain observable $\hat{A}$ in a quantum state $|\psi\rangle$ that evolves under a Hamiltonian $\hat{H}$. Our usual approach (Sec. 1.10) prescribes calculating the evolution of the state of interest according to $|\psi(t)\rangle = \hat{U}(t)\,|\psi(0)\rangle$, where $\hat{U}(t) = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\hat{H}t}$ is the unitary evolution operator[10]. The quantum mean value is then

$$\langle A\rangle = \left\langle \psi(t)\right|\hat{A}\left|\ \psi(t)\right\rangle . \tag{3.126}$$

The Heisenberg picture

[10] Throughout this section, we assume that the Hamiltonian does not explicitly depend on time.

This approach is known as the *Schrödinger picture* of quantum evolution.

An alternative is the *Heisenberg picture*, which assumes that operators evolve according to

$$\hat{A}(t) = \hat{U}^\dagger(t)\hat{A}(0)\hat{U}(t) = \mathrm{e}^{\frac{\mathrm{i}}{\hbar}\hat{H}t}\hat{A}(0)\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\hat{H}t}, \tag{3.127}$$

while all quantum states remain constant: $|\psi(t)\rangle = |\psi(0)\rangle$. The mean value of $\hat{A}$ is then

$$\langle A\rangle = \left\langle \psi(0)\right| \hat{A}(t) \left|\psi(0)\right\rangle. \tag{3.128}$$

At time $t = 0$, the states and operators in the two pictures are assumed equal.

Exercise 3.77. Show that the operator expectation values calculated according to the Schrödinger and Heisenberg pictures [Eqs. (3.126) and (3.128), respectively] are the same.

Exercise 3.78. For the Heisenberg picture, show that the operator evolution can be written in the form (sometimes referred to as *Heisenberg's equation*)

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{A}(t) = \frac{\mathrm{i}}{\hbar}[\hat{H}, \hat{A}(t)]. \tag{3.129}$$

To see how the Heisenberg picture helps to reconcile the classical and quantum approaches, let us look at an example.

Exercise 3.79. Write the Heisenberg equations of motion (3.129) for the position and momentum of the harmonic oscillator assuming the Hamiltonian (3.83).
Answer:

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{x} = \frac{\hat{p}}{M}; \tag{3.130a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{p} = -\kappa\hat{x}. \tag{3.130b}$$

We find, quite remarkably, that the evolution of the harmonic oscillator's position and momentum observables in the Heisenberg picture are identical to the classical ones (Box 3.10). Indeed, Eq. (3.130a) is the definition of the momentum as the product of the mass and velocity, while Eq. (3.130b) is Newton's Second Law of Motion, because $F = -\kappa x$ is the spring force.

Because Eqs. (3.130) are equivalent to the classical ones (Box 3.10), so is their solution, aside from the hats on top of the observables:

$$\hat{x}(t) = \hat{x}(0)\cos\omega t + \frac{1}{M\omega}\hat{p}(0)\sin\omega t; \tag{3.131a}$$

$$\hat{p}(t) = \hat{p}(0)\cos\omega t - \frac{\kappa}{\omega}\hat{x}(0)\sin\omega t. \tag{3.131b}$$

The observed quantum-to-classical analogy may appear to be purely formal because, as one may argue, the position and momentum in the above equations are operators, abstract notions of linear algebra. But in fact there is a direct practical

connotation. To see it, we can "sandwich" both sides of Eqs. (3.131) between symbols $\langle\psi|$ and $|\psi\rangle$ associated with an arbitrary quantum state. Then these equations take the form

$$\langle x(t)\rangle = \langle x(0)\rangle \cos\omega t + \frac{1}{M\omega}\langle p(0)\rangle \sin\omega t; \tag{3.132a}$$

$$\langle p(t)\rangle = \langle p(0)\rangle \cos\omega t - \frac{\kappa}{\omega}\langle x(0)\rangle \sin\omega t. \tag{3.132b}$$

Now, rather than abstract operators, we have measurable physical quantities: mean position and momentum — and they do behave identically to their classical counterparts in any quantum state. This finding reproduces our earlier result (3.115), obtained using the Schrödinger picture.

Is this consistency with the classical behavior a unique property of the harmonic oscillator or general for all mechanical systems? A simple argument suggests the latter.

Exercise 3.80. For the Schrödinger evolution of a state of a point-like particle under the Hamiltonian (3.55), show that

$$\frac{\mathrm{d}\hat{x}}{\mathrm{d}t} = \frac{\hat{p}}{M}; \tag{3.133a}$$

$$\frac{\mathrm{d}\hat{p}}{\mathrm{d}t} = -V'(\hat{x}), \tag{3.133b}$$

where $V'(\cdot)$ is the derivative of the function $V(\cdot)$.
Hint: decompose $V(x)$ into a power series.

Equation (3.133b) corresponds once again to Newton's Second Law because, in classical mechanics, the potential energy of a conservative force is related to that force according to[11]

$$F(x) = -V'(x). \tag{3.134}$$

The relations (3.133) can be made more tangible by taking the mean values of the position and momentum of the particle in an arbitrary state. They then take the form

$$\frac{\mathrm{d}\langle x\rangle}{\mathrm{d}t} = \frac{\langle p\rangle}{M}; \tag{3.135a}$$

$$\frac{\mathrm{d}\langle p\rangle}{\mathrm{d}t} = -\left\langle \psi\right| V'(\hat{x}) \left|\psi\right\rangle = -\int_{-\infty}^{+\infty} \psi(x)V'(x)\psi^*(x)\mathrm{d}x. \tag{3.135b}$$

These relations are known as the *Ehrenfest theorem*. Importantly, it deals with expectation values of observables rather than directly with states or operators. Because these expectation values are the same in both the Schrödinger and Heisenberg pictures (Ex. 3.77), the Ehrenfest theorem is valid in both pictures as well.

[11] We specialize to one-dimensional motion.

Note that, as we know from mechanics, the classical form of Eqs. (3.133) generalizes to the Hamiltonian equations of motion:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{\partial H}{\partial p}; \quad \frac{\mathrm{d}p}{\mathrm{d}t} = -\frac{\partial H}{\partial x}. \tag{3.136}$$

In the quantum domain, these equations are replaced by Heisenberg's equation.

In this way, the Heisenberg picture makes the relationship between quantum and Newtonian mechanics quite apparent. There is, however, a trade-off: a loss of connection between the observable and its eigenstates. For example, consider the harmonic oscillator evolution (3.131) for one-quarter of the oscillation period. Denoting this duration by t_0 (so that $\omega t_0 = \pi/2$), we find $\hat{x}(t_0) = \hat{p}(0)/M\omega$. This result is easy to interpret classically: the position of a pendulum after one-quarter of its period is determined entirely by its initial velocity. But in quantum mechanics, where observables are associated with operators, the observation that the position operator at a certain moment $t = t_0$ becomes proportional to the momentum operator at $t = 0$ is much less comfortable to accept. Indeed, we have defined the position operator as an integral (3.11) over position eigenstates. Its having evolved into the momentum operator at $t = t_0$ means that it is no longer described by that integral. The very nature of the position operator changes with time as it acquires a different set of eigenstates. Moreover, the position observable at different times does not even commute with itself:

$$[\hat{x}(0), \hat{x}(t_0)] = \frac{1}{M\omega}[\hat{x}(0), \hat{p}(0)] = \frac{\mathrm{i}\hbar}{M\omega}.$$

The evolution of observables becomes even less intuitive when we are dealing with the interaction of different quantum systems. It can then happen, for example, that the position of one particle at a particular moment becomes the momentum of another particle at a different moment. Or, if we are dealing with light–atom interactions, the electric field observable associated with the electromagnetic wave transforms into an operator defining the transition between atomic levels. This feature can make the application of the Heisenberg picture to quantum problems quite confusing.

To add to this confusion, let me draw your attention to the following. The Hamiltonian (3.55) uses operators $\hat{x}$ and $\hat{p}$ whose physical nature is the position and momentum, respectively. But, as we found, the nature of these operators in the Heisenberg picture changes with time. Therefore, it seems that Eq. (3.55) represents the correct Hamiltonian only at time $t = 0$, whence the equation should be rewritten as

$$\hat{H} = V(\hat{x}(0)) + \frac{\hat{p}(0)^2}{2M}. \tag{3.137}$$

But then Heisenberg's equation (3.129) would contain a commutator between the Hamiltonian, which is a function of $\hat{x}(0)$ and $\hat{p}(0)$, and the observable $\hat{A}(t)$, which can be either $\hat{x}(t)$ or $\hat{p}(t)$. That is, we calculate the commutator between operators

associated with different moments in time. But when we solved the above exercises, we gave no thought to this, simply writing $[\hat{x},\hat{p}] = \mathrm{i}\hbar$. Wasn't this a mistake?

Exercise 3.81. Show that the Hamiltonian does not evolve in time[12], i.e., $\hat{H}(t) = \hat{H}(0)$.

Exercise 3.82. Show that the Hamiltonian (3.137) can be rewritten as

$$\hat{H} = V(\hat{x}(t)) + \frac{\hat{p}(t)^2}{2M}, \tag{3.138}$$

where t is an arbitrary moment in time and operators $\hat{x}(t)$ and $\hat{p}(t)$ are obtained from Eq. (3.127).
Hint: Use the power series decomposition of the function $V(\hat{x})$.

Remarkably, we find that, even though the position and momentum observables evolve with time, their function given by the right-hand side of Eq. (3.138) remains constant. So both components of the commutator in Eq. (3.129) can be associated with the same time t, thereby resolving our concern.

This observation can be generalized.

Exercise 3.83. Consider some operator $\hat{B}$ that at time $t = 0$ is a function of operators $\hat{A}_1, \ldots \hat{A}_m$:

$$\hat{B}(0) = f\big((\hat{A}_1(0), \ldots, \hat{A}_m(0)\big). \tag{3.139}$$

By decomposing this function into a power series, show that the above relation is preserved at an arbitrary time t, i.e.,

$$\hat{B}(t) = f\big(\hat{A}_1(t), \ldots, \hat{A}_m(t)\big). \tag{3.140}$$

As we see, the evolution in the Heisenberg picture preserves any functional relationship between operators that may have existed prior to that evolution. One consequence of this result is that the Hamiltonian has the same dependence on the position and momentum at different times [see Eqs. (3.137) and (3.138)]. Another relevant example is given in the next exercise.

Exercise 3.84. Show that the time evolution of the position and momentum observables in the Heisenberg picture does not change their commutator:

$$[\hat{x}(t), \hat{p}(t)] = \mathrm{i}\hbar. \tag{3.141}$$

Exercise 3.85. Substitute the solution (3.131) into the Hamiltonian (3.83) to check explicitly that the right-hand sides of Eqs. (3.137) and (3.138) are the same.

[12] The fact that the Hamiltonian, when not explicitly dependent on time, does not evolve under the Heisenberg picture can be viewed as the quantum counterpart of the classical energy conservation law.

3.9.2 Displacement operator

In this section, we will study in detail an example Hamiltonian that can be treated using both the Schrödinger and Heisenberg pictures.

Exercise 3.86. Solve Heisenberg's equation for the Hamiltonian

$$\hat{H} = \beta \hat{p}, \tag{3.142}$$

where β is a real constant, and show that the evolution of the position and momentum operators during time t_0 is given by

$$\hat{x}(t_0) = \hat{x}(0) + x_0; \tag{3.143a}$$
$$\hat{p}(t_0) = \hat{p}(0), \tag{3.143b}$$

where

$$x_0 = \beta t_0. \tag{3.144}$$

We see that the evolution under the Hamiltonian (3.142) leads to the displacement of the position observable by $\hat{x}_0$. Accordingly, the evolution operator

$$\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\hat{H}t} = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\hat{p}x_0} \equiv \hat{D}_x(x_0) \tag{3.145}$$

is called the *position displacement operator*. Let us now study its action in the Schrödinger picture.

Exercise 3.87. Show that the displacement operator is unitary and $\hat{D}_x^\dagger(x) = \hat{D}_x^{-1}(x) = \hat{D}_x(-x)$.

Exercise 3.88. Using the Schrödinger picture, show that

a)
$$\hat{D}_x(x_0)\,|x\rangle = |x+x_0\rangle\,; \tag{3.146}$$

b) if the wavefunction of a state $|\psi\rangle$ in the position basis is $\psi(x)$, then the wavefunction of the state $\hat{D}_x(x_0)\,|\psi\rangle$ is $\psi(x-x_0)$ (Fig. 3.12)[13];

c)
$$\hat{D}_x(x_0)\,|p\rangle = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}x_0 p}\,|p\rangle\,; \tag{3.147}$$

d) if the wavefunction of the state $|\psi\rangle$ in the momentum basis is $\tilde{\psi}(p)$, then the wavefunction of the state $\hat{D}_x(x_0)\,|\psi\rangle$ is $\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}x_0 p}\psi(p)$.

Exercise 3.89. Using both the Heisenberg and Schrödinger pictures, show that the application of the position displacement operator:

a) adds x_0 to the mean position value, but does not change the mean momentum value;

[13] Note that the displacement by *positive* x_0 corresponds to a *negative* change in the argument of the wavefunction. More on this in Sec. 3.9.3.

b) does not change the position and momentum uncertainties.

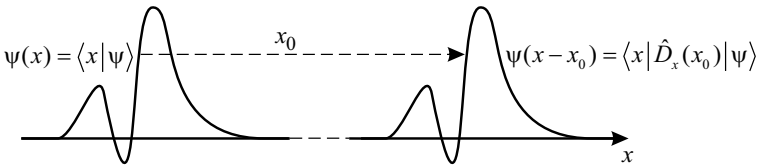


Fig. 3.12 Effect of the position displacement operator on a wavefunction.

Exercise 3.90. Show that the *momentum displacement operator* $\hat{D}_p(p_0) \equiv \mathrm{e}^{\frac{\mathrm{i}}{\hbar}p_0\hat{x}}$ has similar properties with respect to the momentum as the position displacement operator with respect to the position.

Exercise 3.91. The state $|\psi\rangle$ has wavefunction $\psi(x)$ in the position basis. For given values of x_0 and p_0, find the wavefunctions of the following states in the position basis:

a) $\hat{D}_p(p_0)|\psi\rangle$;
b) $\hat{D}_x(x_0)\hat{D}_p(p_0)|\psi\rangle$;
c) $\hat{D}_p(p_0)\hat{D}_x(x_0)|\psi\rangle$.

Answer:

a) $\langle x|\,\hat{D}_p(p_0)\,|\psi\rangle = \mathrm{e}^{\frac{\mathrm{i}}{\hbar}p_0x}\langle x|\,\psi\rangle = \mathrm{e}^{\frac{\mathrm{i}}{\hbar}p_0x}\psi(x)$;
b) $\langle x|\,\hat{D}_x(x_0)\hat{D}_p(p_0)\,|\psi\rangle = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}p_0x_0}\mathrm{e}^{\frac{\mathrm{i}}{\hbar}p_0x}\psi(x-x_0)$;
c) $\langle x|\,\hat{D}_p(p_0)\hat{D}_x(x_0)\,|\psi\rangle = \mathrm{e}^{\frac{\mathrm{i}}{\hbar}p_0x}\psi(x-x_0)$.

The wavefunctions obtained in parts (b) and (c) are not the same. This means that the effect of the position and momentum displacement operator sequence depends on their order, so the operators do not commute. However, the permutation of these operators results only in an overall phase factor $\mathrm{e}^{-\mathrm{i}p_0x_0/\hbar}$, and hence does not affect the physics of the resulting state. This is a manifestation of the Baker–Hausdorff–Campbell formula (A.54), as we shall see next.

Exercise 3.92. For the *phase-space displacement operator* $\hat{D}_{xp}(x_0,p_0) \equiv \mathrm{e}^{\frac{\mathrm{i}}{\hbar}(p_0\hat{x}-x_0\hat{p})}$, show that

$$\hat{D}_{xp}(x_0,p_0) = \mathrm{e}^{\frac{\mathrm{i}}{2\hbar}p_0x_0}\hat{D}_x(x_0)\hat{D}_p(p_0) = E^{-\frac{\mathrm{i}}{2\hbar}p_0x_0}\hat{D}_p(p_0)\hat{D}_x(x_0) \tag{3.148}$$

The above result implies that $\hat{D}_x(x_0)\hat{D}_p(p_0) = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}p_0x_0}\hat{D}_p(p_0)\hat{D}_x(x_0)$, which is consistent with the difference between the answers to parts (b) and (c) of Ex. 3.91.

Exercise 3.93. Write the Hamiltonian that would lead to the evolution corresponding to the phase-space displacement operator. Find the corresponding transformation of the position and momentum observables in the Heisenberg picture.

3.9.3 Evolution of probability densities*

We have seen that the position observable, evolving under the displacement Hamiltonian, produces an operator $\hat{x}(t)$ that is a function of the initial $\hat{x}(0)$. Situations such as this are relatively common. Here we shall investigate whether, in such a situation, we can use the information we have from the Heisenberg picture to predict the evolution of a *wavefunction* in the Schrödinger picture. In the case of position displacement, for example, the relation is straightforward (Fig. 3.12). But can it be generalized?

Throughout this section, we consider the Hamiltonian to be time-independent as usual.

Exercise 3.94. Suppose that in the Heisenberg picture the evolution of an operator $\hat{x}$ under the Hamiltonian $\hat{H}$ transforms it as follows:

$$\hat{x}(t) = \mathrm{e}^{\frac{\mathrm{i}}{\hbar}\hat{H}t}\hat{x}(0)\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\hat{H}t} = f(\hat{x}(0),t), \tag{3.149}$$

where $f(x,t)$ is a real invertible function. Show that in the Schrödinger picture an eigenstate $|x\rangle$ of operator the $\hat{x}$ with the eigenvalue x will evolve into an eigenstate of the same operator with eigenvalue $f(x,t)$.

This result can be written mathematically as

$$\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\hat{H}t}|x\rangle = K(x,t)\,|f(x,t)\rangle\,, \tag{3.150}$$

where $K(x,t)$ is some proportionality coefficient. In the case of position displacement and position eigenstates, this coefficient is equal to one as in Eq. (3.146), but this is not so in general. For example, the if we look at the effect of position displacement on the *momentum* observable, we have $f(p,t_0) = p$ [see Eq. (3.143b)], but $K(x,t_0) = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}x_0 p} \neq 1$ as in Eq. (3.147).

So there appears to be no possibility of determining the complex argument of $K(x,t)$ from the evolution of $\hat{x}$ in the Heisenberg picture. However, fortunately, we can determine its absolute value by recalling that $\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\hat{H}t}$ is a unitary operator, and hence the right-hand side of Eq. (3.150) must have the same norm as $|x\rangle$.

Exercise 3.95. Show that

$$\big\langle f(x,t)\big|\, f(x',t)\big\rangle = \frac{\delta(x-x')}{f'(x,t)}, \tag{3.151}$$

where the derivative $f'(x,t) = \frac{\partial}{\partial x}f(x,t)$ is assumed finite and nonzero.

Exercise 3.96. Show that, in Eq. (3.150),

$$|K(x,t)|^2 = |f'(x,t)|. \tag{3.152}$$

For example, let's say that for some t, $f(x,t) = 2x$, so the evolution "stretches" the position observable by a factor of two. Then Eq. (3.151) becomes, quite reasonably, $\langle 2x|\ 2x'\rangle = \frac{1}{2}\delta(x-x')$, and hence $|K(x,t)|^2 = 2$.

Exercise 3.97. Assuming Eq. (3.150) to hold, show that the wavefunction $\psi(x,t)$ of an arbitrary state $|\psi\rangle$ evolves in time according to

$$\psi(x,t) = K^*(x,-t)\,\psi(f(x,-t),0). \tag{3.153}$$

With the result of the above two exercises, we can predict the effect of the evolution on the absolute value of the wavefunction of observable x. Before we do so, let us eliminate the inconvenient negative time argument in the above equation.

Exercise 3.98. Show that

a) $f(x,-t) = f^{-1}(x,t)$;
b) $|K(x,-t)|^2 = \frac{1}{|f'(x,t)|}$.

Exercise 3.99. Synthesize the above results to obtain, for the evolution of the probability density associated with the wavefunction $\psi(x,t)$,

$$|\psi(x,t)|^2 = \frac{1}{|f'(x,t)|}|\psi(f^{-1}(x,t),0)|^2. \tag{3.154}$$

Looking again at our example where $f(x,t) = 2x$, Eq. (3.154) becomes $|\psi(x,t)|^2 = \frac{1}{2}|\psi(\frac{x}{2},0)|^2$. The probability density function stretches along the x axis and acquires a renormalization factor of $\frac{1}{2}$, in agreement with what we would expect intuitively.

Although the Heisenberg picture does not predict the evolution of the complex phase of the wavefunction, it can be used to calculate the time dependence of its absolute value, and hence the experimentally measurable probability density associated with the observable $\hat{x}$. Generally, the Heisenberg picture is as powerful as the Schrödinger picture in predicting experimental results; the choice of one or the other for a specific calculation is dictated by considerations of simplicity and, quite frequently, the personal taste of the scientist.

3.10 Transformations of harmonic oscillator states

Let us now look at a few operators that can be applied to quantum states of the harmonic oscillator and are particularly important in the context of quantum optics. We study these operators in both the Schrödinger and the Heisenberg picture, thereby acquiring additional practice and learning more about the relationship between these pictures.

Throughout this section, we will *not* be assuming that the system is under the action of the harmonic oscillator Hamiltonian. The reference to the harmonic os-

cillator is limited to the use of the rescaled position and momentum observables introduced in Sec. 3.8, the creation and annihilation operators, and also the states and relations developed in their context. These relations (except those that pertain to the energies and evolution of the states) remain valid no matter what the Hamiltonian may be, and are valid for any values of κ, M, and ω used for the rescaling.

3.10.1 Coherent state as displaced vacuum

We begin by showing that the coherent state can be written as the displaced vacuum state and reproducing some of the results of Sec. 3.8.3 in a simplified fashion.

Exercise 3.100. Show that the phase-space displacement operator in the rescaled units, $\hat{D}_{XP}(X_0,P_0) \equiv \mathrm{e}^{\mathrm{i}P_0\hat{X}-\mathrm{i}X_0\hat{P}}$, corresponds to the following transformations in the Heisenberg picture [Fig. 3.13(a)]:

$$\hat{D}^\dagger_{XP}(X_0,P_0)\hat{X}\hat{D}_{XP}(X_0,P_0) = X + X_0; \tag{3.155a}$$

$$\hat{D}^\dagger_{XP}(X_0,P_0)\hat{P}\hat{D}_{XP}(X_0,P_0) = P + P_0; \tag{3.155b}$$

$$\hat{D}^\dagger_{XP}(X_0,P_0)\hat{a}\hat{D}_{XP}(X_0,P_0) = \hat{a} + \frac{X_0 + iP_0}{\sqrt{2}}, \tag{3.155c}$$

where $\hat{a}$ is the annihilation operator.
Hint: Introduce a fictitious Hamiltonian $\hat{H} = \hbar\omega(P_0\hat{X} - X_0\hat{P})$, where $\omega = 1/t$, and study the evolution of operators $\hat{X}$, $\hat{P}$ and $\hat{a}$ under this Hamiltonian for time t.

Exercise 3.101. Check that the vector $\hat{D}_{XP}(X_\alpha,P_\alpha)|0\rangle$, where $|0\rangle$ is the vacuum state, is an eigenvector of the annihilation operator with eigenvalue $\alpha = \frac{X_\alpha + iP_\alpha}{\sqrt{2}}$. Check that the norm of this vector is 1.

Comparing this result with the definition of the coherent state (Sec. 3.8.3), we see that

$$|\alpha\rangle \propto \hat{D}_{XP}(X_\alpha,P_\alpha)|0\rangle. \tag{3.156}$$

Note that we used a proportionality sign, rather than the equality sign: coherent states $|\alpha\rangle$ follow a certain phase convention and we cannot yet be confident that the right-hand side of Eq. (3.156) has the same phase. We will determine this phase in the following.

Exercise 3.102*.

a) Show that the displacement operator can be rewritten as

$$\hat{D}_{XP}(X_\alpha,P_\alpha) = e^{\alpha\hat{a}^\dagger - \alpha^*\hat{a}}. \tag{3.157}$$

Hint: Use Eq. (3.100).

b) Transform the result of part (a) as follows:

$$\hat{D}_{XP}(X_\alpha, P_\alpha) = \mathrm{e}^{-|\alpha|^2/2}\mathrm{e}^{\alpha\hat{a}^\dagger}\mathrm{e}^{-\alpha^*\hat{a}}. \tag{3.158}$$

Hint: Use the Baker–Hausdorff–Campbell formula (A.54).

c) Show that the right-hand side of Eq. (3.156) can be rewritten as

$$\hat{D}_{XP}(X_\alpha, P_\alpha)\,|0\rangle = \mathrm{e}^{-|\alpha|^2/2}\mathrm{e}^{\alpha\hat{a}^\dagger}\,|0\rangle\,. \tag{3.159}$$

Exercise 3.103. Express the right-hand side of Eq. (3.159) in the Fock basis by expanding the exponent in a power series.

We see that the right-hand side of Eq. (3.156) has exactly the same Fock decomposition (3.122) as the coherent state (3.122). This means that, by displacing the vacuum, we obtain a state that is not just proportional, but actually equal to the coherent state:

$$|\alpha\rangle = \hat{D}_{XP}(X_\alpha, P_\alpha)\,|0\rangle\,. \tag{3.160}$$

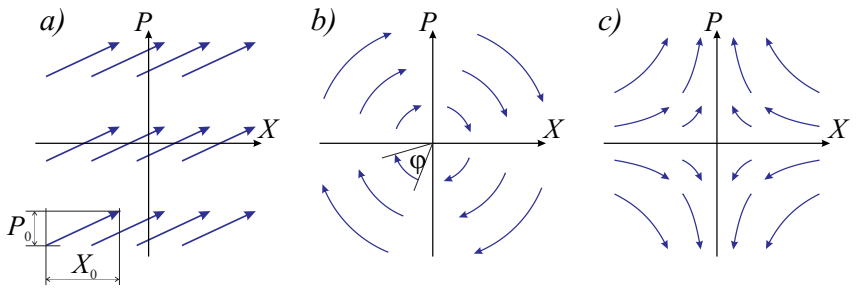


Fig. 3.13 Phase space representation of the displacement (a), phase shift (b), and squeezing (c) operators. The squeezing shown corresponds to $\mathrm{e}^{-r} = \frac{1}{2}$.

3.10.2 Phase shift

The evolution under the harmonic oscillator Hamiltonian (3.96) can be written as

$$\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\hat{H}t} = \mathrm{e}^{-\mathrm{i}\omega t\left(\hat{n}+\frac{1}{2}\right)}. \tag{3.161}$$

We found in Ex. 3.73 that this evolution transforms the coherent state $|\alpha\rangle$ into another coherent state $\mathrm{e}^{-\frac{1}{2}\mathrm{i}\omega t}\left|\alpha\mathrm{e}^{-\mathrm{i}\omega t}\right\rangle$. The coherent state acquires a coherent phase shift of ωt and, in addition, a quantum phase factor $\mathrm{e}^{-\frac{1}{2}\mathrm{i}\omega t}$ that arises from the free

term in the Hamiltonian. It is convenient to introduce the *phase-shift operator*

$$\hat{F}(\varphi) \equiv \mathrm{e}^{-\mathrm{i}\varphi\hat{n}}, \tag{3.162}$$

where φ is a real number. The action of this operator is equivalent to the evolution (3.161) for a time $t = \varphi/\omega$, but does not contain this extraneous quantum phase factor.

Exercise 3.104. Show that

a)
$$\hat{F}(\varphi)\,|n\rangle = \exp(-\mathrm{i}\varphi n)\,|n\rangle\,, \tag{3.163}$$

b)
$$\hat{F}(\varphi)\,|\alpha\rangle = \left|\alpha\mathrm{e}^{-\mathrm{i}\varphi}\right\rangle. \tag{3.164}$$

Equation (3.163) shows how the coherent phase shift works: it applies a quantum phase factor $\exp(-\mathrm{i}\varphi n)\,|n\rangle$ to each Fock component $|n\rangle$ of a state. Acting together within a superposition of Fock states, these (individually unphysical) quantum phase shifts result in a physically meaningful coherent phase shift of the coherent state.

Exercise 3.105. Show that the phase shift transforms the harmonic oscillator operators as follows [Fig. 3.13(b)]:

$$\hat{F}^\dagger(\varphi)\hat{a}\hat{F}(\varphi) = \hat{a}\mathrm{e}^{-\mathrm{i}\varphi}; \tag{3.165}$$
$$\hat{F}^\dagger(\varphi)\hat{a}^\dagger\hat{F}(\varphi) = \hat{a}^\dagger\mathrm{e}^{\mathrm{i}\varphi}; \tag{3.166}$$
$$\hat{F}^\dagger(\varphi)\hat{X}\hat{F}(\varphi) = \hat{X}\cos\varphi + \hat{P}\sin\varphi; \tag{3.167}$$
$$\hat{F}^\dagger(\varphi)\hat{P}\hat{F}(\varphi) = \hat{P}\cos\varphi - \hat{X}\sin\varphi. \tag{3.168}$$

Hint: Similarly to Ex. 3.100, introduce a fictitious Hamiltonian such that the operator transformations on the left-hand sides of the above equations can be interpreted as their evolution under this Hamiltonian in the Heisenberg picture.

We see that applying the phase shift operator (or the harmonic oscillator evolution) leads to clockwise rotation of the phase space through an angle $\varphi = \omega t$ around the origin. This reproduces our earlier result (3.115) for the evolution of the mean position and momentum under the harmonic oscillator Hamiltonian. Let us also recall that we obtained the last two of the above equations, albeit in non-rescaled variables, when we introduced the Heisenberg picture in Sec. 3.9.1.

3.10.3 Squeezing

The *single-oscillator (single-mode) squeezing operator* is given by

$$\hat{S}(r) = \mathrm{e}^{r(\hat{a}^2 - \hat{a}^{\dagger 2})/2}, \tag{3.169}$$

where the *squeezing parameter* r is a real number.

Exercise 3.106.[§] Show that the squeezing operator is unitary and $\hat{S}^\dagger(r) = \hat{S}^{-1}(r) = \hat{S}(-r)$.
Hint: See Ex. 3.87.

Exercise 3.107. Check that the squeezing operator is equivalent to the evolution operator under the Hamiltonian

$$\hat{H} = \frac{\mathrm{i}}{2}\hbar\gamma[\hat{a}^2 - (\hat{a}^\dagger)^2] = -\frac{1}{2}\hbar\gamma[\hat{X}\hat{P} + \hat{P}\hat{X}] \tag{3.170}$$

for time t with $r = \gamma t$. Show that this evolution in the Heisenberg picture transforms operators as follows:

$$\hat{S}^\dagger(r)\hat{X}\hat{S}(r) = \hat{X}\mathrm{e}^{-r}; \tag{3.171}$$

$$\hat{S}^\dagger(r)\hat{P}\hat{S}(r) = \hat{P}\mathrm{e}^{r}; \tag{3.172}$$

$$\hat{S}^\dagger(r)\hat{a}\hat{S}(r) = \hat{a}\cosh r - \hat{a}^\dagger \sinh r; \tag{3.173}$$

$$\hat{S}^\dagger(r)\hat{a}^\dagger\hat{S}(r) = \hat{a}^\dagger \cosh r - \hat{a}\sinh r. \tag{3.174}$$

Exercise 3.108. Suppose that the state $|\psi\rangle$ has mean square uncertainties in the position and momentum equal to $\langle \Delta X_0^2 \rangle$ and $\langle \Delta P_0^2 \rangle$, respectively. Show that the mean square uncertainties in these observables in the state $\hat{S}(r)|\psi\rangle$ are $\mathrm{e}^{-2r}\langle \Delta X_0^2 \rangle$ and $\mathrm{e}^{2r}\langle \Delta P_0^2 \rangle$, respectively.

These results justify the name "squeezing operator". It "shrinks" the position observable while "stretching" the momentum by a factor e^r [Fig. 3.13(c)]. This simultaneous opposite effect on the two observables ensures that the position and momentum uncertainty product is unaffected, so the uncertainty principle still holds. In particular, when the squeezing operator is applied to a vacuum or coherent state, the uncertainty product in the resulting state corresponds to the minimum value (3.95) allowed by the theory.

When we apply the squeezing operator to the vacuum state, we obtain the *squeezed vacuum* state. Its remarkable feature is that its zero-point vibration amplitude in position (for $r > 0$) or momentum (for $r < 0$) is less than that of the vacuum state — the state of lowest possible energy, with zero energy quanta. In the optical implementation of the harmonic oscillator, this zero-point vibration manifests itself as random fluctuations of the electric field around zero. So, in the squeezed vacuum state, this noise is lower than when the light is completely off!

Let us now ask ourselves what the wavefunction of the squeezed vacuum state $\hat{S}(r)|0\rangle$ looks like. Direct derivation of these wavefunctions in the Schrödinger picture is quite tedious. However, in view of our results from the Heisenberg picture, it is easy to guess that the operation of squeezing results in horizontal rescaling and renormalization of the vacuum state wavefunction (3.107a):

$$\psi_{sq}(X) = \langle X | \hat{S}(r) | 0 \rangle = \mathrm{e}^{r/2}\psi_0(X\mathrm{e}^{r}) = \frac{\mathrm{e}^{r/2}}{\pi^{1/4}}\mathrm{e}^{-X^2\mathrm{e}^{2r}/2}; \tag{3.175a}$$

$$\tilde{\psi}_{sq}(P) = \langle P | \hat{S}(r) | 0 \rangle = \mathrm{e}^{-r/2}\psi_0(P\mathrm{e}^{-r}) = \frac{\mathrm{e}^{-r/2}}{\pi^{1/4}}\mathrm{e}^{-P^2\mathrm{e}^{-2r}/2}. \tag{3.175b}$$

Exercise 3.109. Check that the wavefunctions (3.175) are

a) normalized;
b) consistent with Eq. (3.154).

The test we just performed does not tell us whether we have guessed the complex phase of the wavefunctions correctly. To perform this test, let us simply plug them into the time-dependent Schrödinger equation and check for consistency.

Exercise 3.110. Check that wavefunctions (3.175) satisfy the Schrödinger equation for the Hamiltonian (3.170) with $r = \gamma t$.

The *two-oscillator (two-mode) squeezing* operator, acting on two oscillators indicated by subscripts A and B, is given by

$$\hat{S}_2(r) = \exp[r(-\hat{a}_A\hat{a}_B + \hat{a}_A^\dagger\hat{a}_B^\dagger)], \tag{3.176}$$

where r is a real number.

Exercise 3.111. a) Check that the two-mode squeezing operator can be associated with the following fictitious Hamiltonian:

$$\hat{H} = \mathrm{i}\hbar\gamma(-\hat{a}_A\hat{a}_B + \hat{a}_A^\dagger\hat{a}_B^\dagger) = \hbar\gamma(\hat{X}_A\hat{P}_B + \hat{P}_A\hat{X}_B) \tag{3.177}$$

with $r = \gamma t$.

b) Show that two-mode squeezing is represented in the Heisenberg picture by the following operator transformation[14]:

$$\hat{S}_2^\dagger(r)\hat{X}_\pm\hat{S}_2(r) = \hat{X}_\pm\mathrm{e}^{\pm r}; \tag{3.178}$$

$$\hat{S}_2^\dagger(r)\hat{P}_\pm\hat{S}_2(r) = \hat{P}_\pm\mathrm{e}^{\mp r}; \tag{3.179}$$

$$\hat{S}_2^\dagger(r)\hat{a}_A\hat{S}_2(r) = \hat{a}_A\cosh r + \hat{a}_B^\dagger\sinh r; \tag{3.180}$$

$$\hat{S}_2^\dagger(r)\hat{a}_B\hat{S}_2(r) = \hat{a}_B\cosh r + \hat{a}_A^\dagger\sinh r;, \tag{3.181}$$

where

$$\hat{X}_\pm = \frac{\hat{X}_A \pm \hat{X}_B}{\sqrt{2}}; \quad P_\pm = \frac{\hat{P}_A \pm \hat{P}_B}{\sqrt{2}}. \tag{3.182}$$

c) Find the expectation values and uncertainties of the observables $\hat{X}_{A,B}$, $\hat{P}_{A,B}$, $\hat{X}_\pm$ and $\hat{P}_\pm$ in the two-mode squeezed vacuum state $\hat{S}_2(r)|00\rangle$.
Answer: All expectation values are zero. The mean square uncertainties are:

[14] The mode operator transformation given by Eqs. (3.173), (3.174) or Eqs. (3.180), (3.181) is called the *Bogoliubov* transformation.

$$\left\langle \Delta X_-^2 \right\rangle = \left\langle \Delta P_+^2 \right\rangle = \frac{1}{2} \mathrm{e}^{-2r}; \tag{3.183}$$

$$\left\langle \Delta X_+^2 \right\rangle = \left\langle \Delta P_-^2 \right\rangle = \frac{1}{2} \mathrm{e}^{2r}; \tag{3.184}$$

$$\langle \Delta X_A^2 \rangle = \langle \Delta X_B^2 \rangle = \langle \Delta P_A^2 \rangle = \langle \Delta P_B^2 \rangle = \frac{1}{2} \cosh 2r. \tag{3.185}$$

Exercise 3.112. By substituting into the time-dependent Schrödinger equation, check that the normalized wavefunctions of the two-mode squeezed vacuum state in the position and momentum bases are (Fig. 3.14):

$$\begin{aligned} \Psi_{sq2}(X_A, X_B) &= \left\langle X_A, X_B \right| \hat{S}_2(r) \left| 0,0 \right\rangle \qquad (3.186a) \\ &= \frac{1}{\sqrt{\pi}} \mathrm{e}^{-(X_A - X_B)^2 \mathrm{e}^{2r}/4} \mathrm{e}^{-(X_A + X_B)^2 \mathrm{e}^{-2r}/4}; \\ \tilde{\Psi}_{sq2}(P_A, P_B) &= \left\langle P_A, P_B \right| \hat{S}_2(r) \left| 0,0 \right\rangle \qquad (3.186b) \\ &= \frac{1}{\sqrt{\pi}} \mathrm{e}^{-(P_A - P_B)^2 \mathrm{e}^{-2r}/4} \mathrm{e}^{-(P_A + P_B)^2 \mathrm{e}^{2r}/4}. \end{aligned}$$

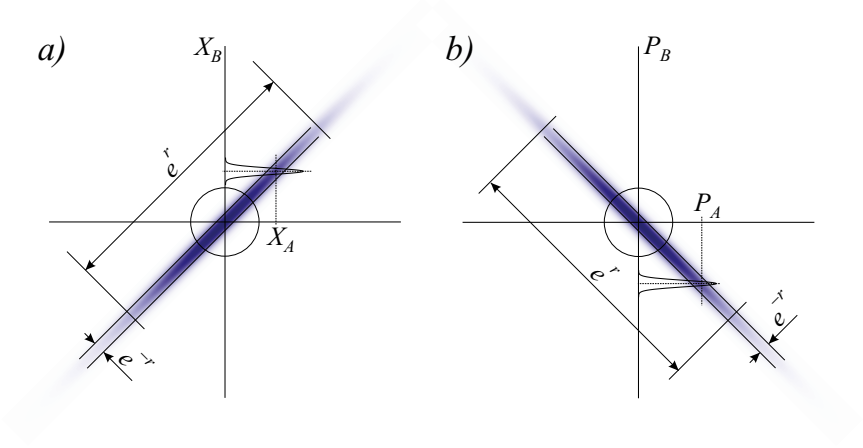


Fig. 3.14 Wavefunction of the two-oscillator squeezed vacuum state in the position (a) and momentum (b) bases. The position observables are correlated, and the momentum observables anticorrelated, so the uncertainties in their difference and sum, respectively, are below the vacuum level (represented by the circles). By measuring either the position or momentum, Alice remotely prepares a state that approximates, for large r, an eigenstate of the same observable at Bob's location.

Exercise 3.113. The two-oscillator squeezed state (3.186a) is shared between observers Alice and Bob.

a) Suppose Alice measures her particle's position and obtains the result X_A. What will the wavefunction of Bob's particle become in the position basis? What is its corresponding position uncertainty $\langle \Delta X_B^2 \rangle$?
b) Suppose Alice measures her particle's momentum and obtains the result P_A. What will the wavefunction of Bob's particle become in the momentum basis? What is its corresponding momentum uncertainty $\langle \Delta P_B^2 \rangle$?

Answer:

$$\langle \Delta X_B^2 \rangle = \langle \Delta P_B^2 \rangle = \frac{1}{2\cosh 2r}. \tag{3.187}$$

Equation (3.187) reveals a remarkable property of the two-mode squeezed vacuum. If we measure either the position or momentum of one of the two oscillators, we will know the corresponding observable of the other oscillator with an uncertainty that is less than that of the vacuum state (Fig. 3.14). In other words, we can remotely, at will, prepare the other oscillator in one of the two states for which the product of the position and momentum uncertainties is below the minimum allowed by by the uncertainty principle. This violates local realism by the same logic as does the original Einstein–Podolsky–Rosen state (Sec. 3.3.3).

This property of the two-mode squeezed vacuum, occurring for any value of the squeezing parameter r (positive or negative), is due to its entangled nature. Because it is relatively easy to prepare experimentally (Box 3.13), this state is a primary entangled resource in various quantum optical information protocols that rely on continuous-variable representations of electromagnetic oscillators.

Let us look briefly at the two-mode squeezed state in *non-rescaled* variables. What would be its wavefunction and under which circumstances would it instantiate the EPR paradox?

Exercise 3.114. Alice and Bob share a state with the wavefunction

$$\Psi(x_A, x_B) = \mathcal{N}\left[\mathrm{e}^{-\frac{(x_A - x_B)^2}{4d^2}} \mathrm{e}^{-\frac{(x_A + x_B)^2}{4D^2}} \right], \tag{3.188}$$

where x_A and x_B are non-rescaled position and momentum observables, and d and D are positive constants.

a) Find the factor ζ that rescales the position observable according to $\hat{X}_{A,B} = \zeta \hat{x}_{A,B}$ in such a way that the above wavefunction takes the form (3.186a). Show that the corresponding squeezing factor is $e^r = \sqrt{\frac{D}{d}}$.
b) Find the corresponding rescaling factor for the momentum observable such that $[\hat{X}_{A,B}, \hat{P}_{A,B}] = \mathrm{i}$.

Our result means that a two-particle Gaussian wavepacket (3.188) exhibits entanglement with *any* amount of correlation between the positions of the two particles, no matter how small. The entanglement is absent only for $d = D$, i.e., when this state becomes explicitly separable:

$$\Psi(x_A, x_B) = \mathcal{N}\left[\mathrm{e}^{-\frac{x_A^2+x_B^2}{2d^2}}\right].$$

Our final goal in the discussion of squeezing is to find the Fock basis decomposition of the single-mode and two-mode squeezed states. We will first do an approximate evaluation for small r to illustrate the concept, and then a complete calculation.

Exercise 3.115. a) Decompose the single-mode squeezing operator into a power series to first order in r and apply it to the vacuum state in the Schrödinger picture in the Fock basis. Show that the resulting state is given by

$$\hat{S}(r)\,|0\rangle \approx |0\rangle - \frac{r}{\sqrt{2}}\,|2\rangle\,. \tag{3.189}$$

Calculate the position and momentum variances in this state and show that they are consistent with the result of Ex. 3.108.

b) Decompose the two-mode squeezing operator into a power series to first order in r and apply it to the double-vacuum state $|0,0\rangle$. Show that the resulting state is given by

$$\hat{S}_2(r)\,|0,0\rangle \approx |0,0\rangle + r\,|1,1\rangle\,. \tag{3.190}$$

Calculate the variances of the observables $\hat{X}_\pm$ and $\hat{P}_\pm$ in this state and show that they are consistent with Eqs. (3.183) and (3.184).

We can see the salient features of the Fock structure of the squeezed states from this simple calculation. The Taylor decomposition of the two-mode squeezing operator contains terms with various powers of the operators $\hat{a}_A\hat{a}_B$ and $\hat{a}_A^\dagger\hat{a}_B^\dagger$. This means that $\hat{S}_2(r)$ creates and destroys energy quanta in the two oscillators strictly in pairs, so the two-mode squeezed state only contains terms with the same numbers of quanta:

$$\hat{S}_2(r)\,|0,0\rangle = \sum_{n=0}^{\infty} D_n\,|nn\rangle\,.$$

Similarly, the single-mode squeezing operator creates and annihilates quanta in a single oscillator in a strictly pairwise fashion, so the Fock decomposition of the single-mode squeezed state only contains terms with even numbers of photons:

$$\hat{S}(r)\,|0\rangle = \sum_{m=0}^{\infty} C_m\,|2m\rangle\,.$$

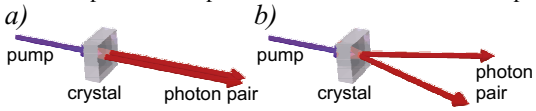
In the following, we will find the coefficients C_m and D_n. This calculation is a good exercise but relatively lengthy, so it can be skipped in the first reading.

Exercise 3.116.* Show that

$$\hat{S}(r)\,|0\rangle = \frac{1}{\sqrt{\cosh r}} \sum_{m=0}^{\infty} (-\tanh r)^m \frac{\sqrt{(2m)!}}{2^m m!}\,|2m\rangle \tag{3.191}$$

Box 3.13 Preparing and measuring the squeezed states.

In the optical realization of the harmonic oscillator, squeezed states can be produced using (you guessed it) parametric down-conversion (Box 1.6). As we know, a primary property of this phenomenon is that the photons are generated in pairs — exactly as one would expect to see in squeezed vacuum states. The down-conversion is configured differently depending on whether the single- or two-mode squeezed vacuum is to be generated: either in the *degenerate* fashion, when the two photons are emitted into the same optical mode or *non-degenerate*, when the photons in a pair are distributed between two optical channels.



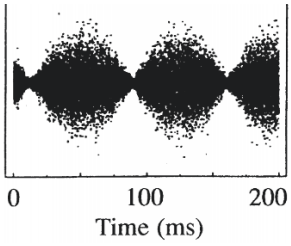
Spontaneous parametric down-conversion. a) Degenerate configuration, leading to the single-mode squeezed vacuum. b) Non-degenerate configuration, leading to the two-mode squeezed vacuum.

The non-degenerate configuration is the same as discussed in the context of heralded photon sources (Box 1.6) and entangled-pair sources (Box 2.1). However, these discussions relied on the assumption that the pumping was weak, so the probability of generating two or more pairs at a time is negligible. When we drop this assumption we obtain squeezing.

We see that the series (3.193) is a geometric progression: the amplitude of each subsequent term is a factor of $\tanh r$ times the previous one. This is what is expected from parametric down-conversion: because this process is spontaneous, the probability of producing n pairs is the nth power of the probability of producing a single pair. If this probability is significant, the squeezing factor e^{-r} (see Ex. 3.108) is substantially different from 1.

In the single-mode squeezing case (3.191), this geometric-progression relation becomes more complicated because of the interference between photons of the same pair emitted into the same mode.

Once the squeezed state has been generated, how can it be detected? One way to obtain evidence for two-mode squeezing is to measure the photon numbers in the two emission modes and check that they are correlated. However, this method does not reveal the phase relation between the photon pair components, and furthermore, it is not suitable for detecting single-mode squeezing. It is much more conclusive to perform multiple measurements of the position and momentum observables using the homodyne detector (Box 3.12) and check that their statistics behave as expected.



Multiple measurements of the observable $\hat{X}\cos\theta + \hat{P}\sin\theta$ in the single-mode position-squeezed vacuum state. The parameter θ varies with time so $\sim$10, 90, 160 ms correspond to measurements of the position observable, and those at $\sim$50, 130, 200 ms to the momentum. Reproduced from G. Breitenbach, S. Schiller, and J. Mlynek, *Measurement of the quantum states of squeezed light*, Nature **387**, 471 (1997).

by going through the following steps:

a) Calculate the inner product between $\hat{S}(r)|0\rangle$ and a coherent state $|\alpha\rangle$ (with a real amplitude α), e.g., in the position basis.
Answer:

$$\langle\alpha|\,\hat{S}(r)\,|0\rangle = \sqrt{\frac{1}{\cosh r}}\exp\left[-\frac{e^{2r}}{1+e^{2r}}\alpha^2\right]. \tag{3.192}$$

b) Decompose the coherent state on the left-hand side of Eq. (3.192) in the Fock basis and the exponent on the right-hand side of that equation into a power series in α. Equate the terms with the same power of α on the two sides of the equation to obtain Eq. (3.191).

Exercise 3.117* Show that

$$\hat{S}_2(r)\,|0,0\rangle = \frac{1}{\cosh r}\sum_{n=0}^{\infty}\tanh^n r\,|nn\rangle \tag{3.193}$$

by going through the following steps:

a) Calculate the overlap of $\hat{S}_2(r)\,|0,0\rangle$ with the tensor product $|\alpha,\alpha\rangle$ of identical coherent states in Alice's and Bob's oscillators:

$$\langle\alpha,\alpha|\,\hat{S}_2(r)\,|0,0\rangle = \frac{1}{\cosh r}\exp\left[-\frac{2}{1+e^{2r}}\alpha^2\right]. \tag{3.194}$$

b) Decomposing the coherent states on the left-hand side in the Fock basis and keeping only the terms with equal photon numbers, show that

$$\sum_{n=0}^{\infty}\langle n,n|\,\hat{S}_2(r)\,|0,0\rangle\frac{\alpha^{2n}}{n!} = \frac{1}{\cosh r}e^{\alpha^2\tanh r}. \tag{3.195}$$

c) Decompose the exponent on the right-hand side of the above equation into a power series in α to obtain Eq. (3.193).

Exercise 3.118* Find the mean and variance of the number of energy quanta

a) in the single-mode squeezed vacuum state;
b) in the two-mode squeezed state (per channel).

Hint: Write the squared norm of both states using Eqs. (3.191) and (3.193) and calculate the derivative by $\tanh r$.
Answer:

a) $\langle m\rangle = \sinh^2 r$; $\langle\Delta m^2\rangle = 2\sinh^2 r + 2\sinh^4 r$.
b) $\langle n\rangle = \sinh^2 r$; $\langle\Delta n^2\rangle = \sinh^2 r + \sinh^4 r$.

3.11 Problems

Problem 3.1. A state has wavefunction

$$\psi(x) = Ax\mathrm{e}^{-\kappa^2 x^2/2}.$$

a) Find the normalization factor A.
b) Find the wavefunction $\tilde{\psi}(p)$ in the momentum basis.
c) Check the uncertainty principle: $\langle \Delta p^2 \rangle \langle \Delta x^2 \rangle \geq \hbar^2/4$.

Hint:

$$\int_{-\infty}^{+\infty} x^2 \mathrm{e}^{-x^2} dx = \frac{\sqrt{\pi}}{2}; \quad \int_{-\infty}^{+\infty} x^4 \mathrm{e}^{-x^2} dx = \frac{3\sqrt{\pi}}{4}.$$

Problem 3.2. Find the matrix element $\langle p | \hat{A} | p' \rangle$ if the operator $\hat{A}$ is a function of position:

a) $\hat{A}(x) = A_0$;
b) $\hat{A} = \mathrm{e}^{-\hat{x}^2/b^2}$.

Problem 3.3. For the energy eigenstates of Ex. 3.40, find the uncertainties in the position and momentum and check that the uncertainty principle is satisfied.

Problem 3.4. Consider the state $\psi(x) = \begin{cases} Ax \text{ for } |x|<a/2 \\ 0 \text{ for } |x|\geq a/2 \end{cases}$ (where $A = 2\sqrt{3}/a^{3/2}$ is the norm) in the potential of Ex. 3.40. Find the energy spectrum of this state, i.e., the probabilities pr_n of observing each energy eigenstate. Show that these probabilities add up to 1.
Hint: $\sum 1/n^2 = \pi^2/6$.

Problem 3.5. Consider a particle of mass M whose initial state has the wavefunction $\psi(x)$, in an infinite potential box of width a. Show that the evolution under the Schrödinger equation will restore the initial state (possibly with a phase factor) after time $t = 4Ma^2/\pi\hbar$.

Problem 3.6. For the finite well potential (3.65):

a) find analytically the approximate corrections to the first two energy levels of an infinitely deep potential well (Ex. 3.40) when it is replaced by a finite well with $V_0 \gg E_1$, where E_1 is given by Eq. (3.69);
b) find numerically the first two energy eigenvalues for $k_0 a = 10$. Is your result consistent with the result of part (a)?

Problem 3.7. A particle is in the ground state of an infinite potential box of length a. The box suddenly expands (symmetrically) to twice its width. What is the probability of finding the particle in the ground state of the new potential?

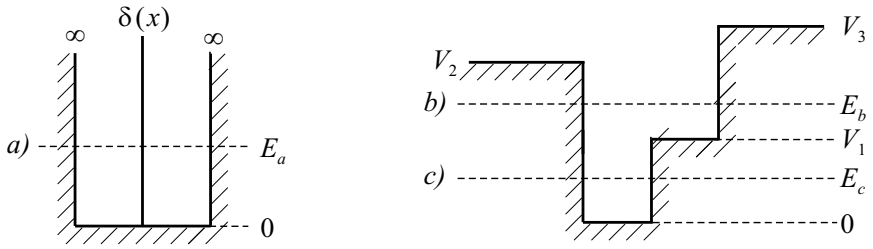


Fig. 3.15 Potential for Problem 3.8

Problem 3.8. Sketch qualitatively the real parts of the stationary wavefunctions for the potentials shown in Fig. 3.15 with energy levels as shown. Your solution should give proper attention to details, such as the relations between de Broglie wavelengths in different areas of the plot, continuity conditions, etc.

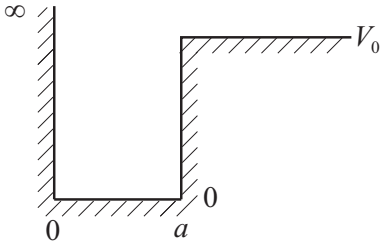


Fig. 3.16 Potential for Problem 3.9

Problem 3.9. Find the transcendental equation for the energy eigenvalues associated with the bound stationary states of the potential

$$V(x) = \begin{cases} +\infty \text{ for } x \leq 0; \\ 0 \text{ for } 0 < x \leq a; \\ V_0 \text{ for } x > a. \end{cases}$$

Compare your result with that of Ex. 3.39.

Problem 3.10. Solve Ex. 3.41 in the momentum basis. Check consistency with the position basis solution.
Hint:

$$\int_{-\infty}^{+\infty} \frac{1}{1+x^2} \mathrm{d}x = \pi; \quad \int_{-\infty}^{+\infty} \frac{1}{(1+x^2)^2} \mathrm{d}x = \pi/2 \tag{3.196}$$

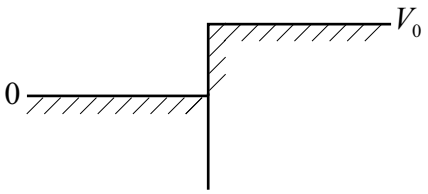


Fig. 3.17 Potential for Problem 3.11.

Problem 3.11. Find the energies and wavefunctions of all bound states associated with the potential

$$V(x) = V_0\theta(x) - W_0\delta(x),$$

where V_0 and W_0 are positive and $\theta(x)$ is the Heaviside step function (Fig. 3.17). Find the conditions for the existence of at least one bound state.

Problem 3.12. Calculate the reflection and transmission for scattering on a delta-potential $V(x) = W_0\delta(x)$, with energy $E > 0$. Compare your results with those obtained from Eqs. (3.81) for an infinitely thin and high rectangular potential barrier ($L \to 0$, $V_0 = W_0/L$).

Problem 3.13. A massive particle of mass M is attached to a spring with spring constant κ. The other end of the spring is attached to a wall, resulting in harmonic oscillatory motion.

a) Write the full set of energy eigenvalues and the corresponding normalized wavefunctions in the *non-rescaled* position basis.
b) Suppose another wall is inserted at point $x = 0$ as shown in Fig. 1, so the particle cannot go into the region $x > 0$. How should the above set be modified in order to represent the energy eigenvalues and eigenstates for the new potential?

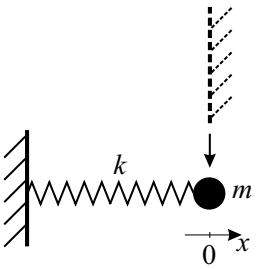


Fig. 3.18 Illustration for Problem 3.13

Problem 3.14. A massive particle of mass M is attached to a spring with spring constant κ. The other end of the spring is attached to a wall, resulting in harmonic oscillatory motion. The particle is initially in the ground energy eigenstate.

a) At time $t = 0$, an additional, position-independent force F begins to act on the particle. Find the probability of detecting the particle in the ground state of the new potential.
b) Find the expectation values of the position $\langle x(t)\rangle$ and momentum $\langle p(t)\rangle$ of the particle as a function of time.
Hint: you need not find the evolution of the wavefunction.

Problem 3.15. [15]

The *single-photon added coherent state* (SPACS) is obtained from coherent states by action of the creation operator: $|\alpha,1\rangle = \mathscr{N}\hat{a}^\dagger|\alpha\rangle$.

a) Find the normalization factor $\mathscr{N}$.
b) Find the decomposition of this state in the photon number basis (you are not required to simplify the result).
c) Find the expectation value of the position observable.
d) Find the wavefunction of the SPACS for real α.
e) Which quantum state does SPACS approach in the limit $\alpha = 0$? $\alpha \to \infty$?

Problem 3.16. Consider the state of the harmonic oscillator whose decomposition in the photon number basis has the form

$$|\psi(t=0)\rangle = \alpha|0\rangle - \beta|2\rangle,$$

where α and β are real and $\alpha^2 + \beta^2 = 1$.

a) Find the wavefunction of $|\psi(t=0)\rangle$ in the position basis.
b) Find the behavior $|\psi(t)\rangle$ of this state as a function of time in the photon number basis.
c) Find the expectation value and variance of the energy as a function of time.
d) Find the expectation value and variance of the position as a function of time.
e) For which values of α and β is the state $|\psi(t=0)\rangle$ position-squeezed, i.e., the position variance is less than that of the vacuum state?

Problem 3.17. Consider the following state of two harmonic oscillators:

$$|\psi\rangle = \alpha|0,0\rangle - \beta|1,1\rangle,$$

where α and β are real with $\alpha^2 + \beta^2 = 1$.

a) For which values of α and β does this state exhibit two-mode position-squeezing, i.e., the variance of $\hat{X}_A - \hat{X}_B$ is less than that of the double-vacuum state?
b) Answer the same question for the momentum observable.

[15] In all problems below, use the rescaled position and momentum observables, i.e., $[\hat{X},\hat{P}] = \mathrm{i}$.

Problem 3.18. Consider coherent superpositions of coherent states $|S_\pm\rangle = \mathcal{N}_\pm(|\alpha\rangle \pm |-\alpha\rangle)$, where $\mathcal{N}_\pm$ are normalization factors and the amplitude α is real and positive[16].

a) Find $\mathcal{N}_\pm$.
b) Find the matrices (wavefunctions) of these states

- in the Fock basis;
- in the position basis;
- in the momentum basis.

c) Show that, for small amplitudes α, these states can be approximated, up to the first two terms in the Fock decomposition, by the states

$$|S_+\rangle \approx \hat{S}(r_+)|0\rangle,$$
$$|S_-\rangle \approx \hat{S}(r_-)|1\rangle,$$

and find $r_\pm(\alpha)$ for which the approximation is optimal.

Problem 3.19. For the phase-space displacement operators $\hat{D}_{XP}(X_\alpha, P_\alpha)$ and $\hat{D}_{XP}(X_\beta, P_\beta)$ with $\alpha, \beta = \frac{X_{\alpha,\beta} + iP_{\alpha,\beta}}{\sqrt{2}}$:

a) express the operator $\hat{D}_{XP}(X_\beta, P_\beta)\hat{D}_{XP}(X_\alpha, P_\alpha)$ through $\hat{D}_{XP}(X_\alpha + X_\beta, P_\alpha + P_\beta)$;
b) express the state $\hat{D}_{XP}(X_\beta, P_\beta)|\alpha\rangle$ through the coherent state vector $|\alpha+\beta\rangle$.

Problem 3.20. For the position displacement operator $\hat{D}_X(X_0)$ in rescaled variables:

a) Find $\hat{D}_X^\dagger(X_0)\hat{a}\hat{D}_X(X_0)$ and $\hat{D}_X^\dagger(X_0)\hat{a}^\dagger\hat{D}_X(X_0)$.
b) Find $[a, \hat{D}_X(X_0)]$ and $[a^\dagger, \hat{D}_X(X_0)]$.
c) Find the Fock decomposition of the *displaced single-photon state* $\hat{D}_X(X_0)|1\rangle$.
Hint: $|n\rangle = (\hat{a}^\dagger)^n|0\rangle/\sqrt{n!}$.

Problem 3.21. A harmonic oscillator, initially in the vacuum state, has evolved under the Hamiltonian $\hat{H}_1 = r\left[\hat{a}^2 + (\hat{a}^\dagger)^2\right]/2$ or $\hat{H}_2 = r\hat{P}^2$, with a real and positive r, for time t_0. Perform the following calculations for the resulting state:

a) Find the mean and variance of the general *quadrature observable* $\hat{X}_\theta = \hat{X}\cos\theta + \hat{P}\sin\theta$ for an arbitrary angle θ.
b) Which angle corresponds to the highest squeezing?
c) What is the corresponding quadrature variance?

Answer these questions for both Hamiltonians $\hat{H}_1$ and $\hat{H}_2$.

[16] This state is dubbed the "Schrödinger kitten" because it is a superposition of two "classical" and potentially macroscopic coherent states, and yet it is highly nonclassical. It is a subject of intense research because, by constructing such states with increasingly high amplitudes α, we may be able to identify the limits of quantum physics — see Sec. 2.4.3.

Problem 3.22. Two harmonic oscillators, initially in the vacuum state $|0\rangle \otimes |0\rangle$, interact under the Hamiltonian

$$\hat{H} = \hbar\chi \hat{X}_A \hat{P}_B$$

with a real and positive χ.

a) Write the differential equations for the position and momentum observables $\hat{X}_{A,B}(t)$ and $\hat{P}_{A,B}(t)$ in the Heisenberg picture.
b) Solve these equations and obtain the expressions for $\hat{X}_{A,B}(t)$ and $\hat{P}_{A,B}(t)$.
c) Find the expectation values and variances of observables $\hat{X}_{A,B}$, $\hat{P}_{A,B}$, $\hat{X}_{\pm} = (\hat{X}_A \pm \hat{X}_B)/\sqrt{2}$, and $\hat{P}_{\pm} = (\hat{P}_A \pm \hat{P}_B)/\sqrt{2}$ as functions of time t.
d) For which values of t is two-mode squeezing present, i.e., one of the uncertainties of $\hat{X}_{\pm}$ or $\hat{P}_{\pm}$ is below that of the vacuum state at time $t = 0$?
e) Find the first-order approximation of the state into which the double-vacuum state evolves under Hamiltonian $\hat{H}$ in the Fock basis, in the Schrödinger picture, assuming $\chi t/\hbar \ll 1$.
f) Find the mean square value $\langle X_{\pm}^2 \rangle$ of that state. Is your result consistent with that of part (d)?

Chapter 4
Angular momentum

4.1 3D motion

Now that we've become experts in one-dimensional quantum mechanics, it is time to remember that the space we live in is three-dimensional. So in order to provide a quantum theoretical description for realistic physical objects, such as atoms, we need to generalize our results to three dimensions. A straightforward way of doing so would be to say that the Hilbert space of three-dimensional states of a point-like particle is a tensor product of Hilbert spaces associated with individual coordinates:

$$\mathbb{V}_{3D} = \mathbb{V}_x \otimes \mathbb{V}_y \otimes \mathbb{V}_z. \tag{4.1}$$

The 3D position and momentum observables are vectors[1] whose components are the position and momentum observables of individual one-dimensional spaces[2]: $\hat{\vec{r}} = (\hat{x}, \hat{y}, \hat{z})$, $\hat{\vec{p}} = (\hat{p}_x, \hat{p}_y, \hat{p}_z)$. The commutation relations between the components of the 3D position and momentum observables are $[\hat{r}_j, \hat{p}_k] = \mathrm{i}\hbar\delta_{jk}$. That is, the position and momentum do not commute if and only if they belong to the same Hilbert space.

By *eigenstates* of vector operators we understand simultaneous eigenstates of their component operators. For example, the state $|\vec{r}\rangle = |x\rangle \otimes |y\rangle \otimes |z\rangle$ simultaneously satisfies three equations:

$$\begin{aligned} \hat{x}\,|\vec{r}\rangle &= (\hat{x} \otimes \hat{\mathbf{1}} \otimes \hat{\mathbf{1}})(|x\rangle \otimes |y\rangle \otimes |z\rangle) = x\,|\vec{r}\rangle\,; \\ \hat{y}\,|\vec{r}\rangle &= (\hat{\mathbf{1}} \otimes \hat{y} \otimes \hat{\mathbf{1}})(|x\rangle \otimes |y\rangle \otimes |z\rangle) = y\,|\vec{r}\rangle\,; \\ \hat{z}\,|\vec{r}\rangle &= (\hat{\mathbf{1}} \otimes \hat{\mathbf{1}} \otimes \hat{z})(|x\rangle \otimes |y\rangle \otimes |z\rangle) = z\,|\vec{r}\rangle\,, \end{aligned} \tag{4.2}$$

so $|\vec{r}\rangle$ is an eigenstate of $\hat{\vec{r}}$.

Let me emphasize that a vector operator is *not* a tensor product operator in the sense of Sec. 2.1.3. Rather, it is just a set of three operators. This means, for example,

[1] To avoid confusion, we will not be using the term "vector" in the sense "element of the Hilbert space" in this chapter. We use this term only for observables which have x-, y- and z-components.

[2] Sometimes we will be using an alternative notation, as follows: $\hat{\vec{r}} = (\hat{r}_1, \hat{r}_2, \hat{r}_3)$, $\hat{\vec{p}} = (\hat{p}_1, \hat{p}_2, \hat{p}_3)$.

A. I. Lvovsky, *Quantum Physics*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-662-56584-1_4

that acting with operator $\hat{\vec{r}}$ upon a tensor product of position eigenstates $|\vec{r}\rangle \equiv |x\rangle \otimes |y\rangle \otimes |z\rangle$, we will obtain a set of three states $(x|\vec{r}\rangle, y|\vec{r}\rangle, z|\vec{r}\rangle)$. If $\hat{\vec{r}}$ were an operator tensor product, we would instead obtain a single state $xyz|\vec{r}\rangle$.

As in the one-dimensional case, the wavefunction of any state $|\psi\rangle$ is given by

$$\psi(\vec{r}) = \langle \vec{r}|\ \psi\rangle . \tag{4.3}$$

Exercise 4.1. Show that:

a) any state $|\psi\rangle$ is related to its wavefunction (4.3) according to

$$|\psi\rangle = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \psi(\vec{r})\,|\vec{r}\rangle\,\mathrm{d}x\mathrm{d}y\mathrm{d}z; \tag{4.4}$$

b) the inner product of two states $|\psi\rangle$ and $|\varphi\rangle$ in $\mathbb{V}_{3D}$ is given by

$$\langle\psi|\ \varphi\rangle = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \psi^*(\vec{r})\varphi(\vec{r})\mathrm{d}x\mathrm{d}y\mathrm{d}z. \tag{4.5}$$

Exercise 4.2. Write down the *three-dimensional de Broglie wave*, i.e., the inner product of states $|\vec{r}\rangle = |x\rangle \otimes |y\rangle \otimes |z\rangle$ and $|\vec{p}\rangle = |p_x\rangle \otimes |p_y\rangle \otimes |p_z\rangle$.
Answer:

$$\langle\vec{r}|\ \vec{p}\rangle = \frac{1}{(2\pi\hbar)^{3/2}} \mathrm{e}^{\frac{\mathrm{i}}{\hbar}(xp_x+yp_y+zp_z)} = \frac{1}{(2\pi\hbar)^{3/2}} \mathrm{e}^{\frac{\mathrm{i}}{\hbar}\vec{r}\cdot\vec{p}}. \tag{4.6}$$

We shall now look at the Hamiltonian governing the motion in 3D space. As in the one-dimensional case, one of our goals in this chapter will be to find wavefunctions of energy eigenstates for various potentials.

A general Hamiltonian for mechanical motion is the sum of kinetic and potential energies. In 3D, it takes the form

$$\hat{H} = \frac{\hat{p}_x^2}{2M} + \frac{\hat{p}_y^2}{2M} + \frac{\hat{p}_z^2}{2M} + V(\hat{\vec{r}}). \tag{4.7}$$

The kinetic energy observable in $\mathbb{V}_{3D}$ is a sum of the kinetic energies corresponding to individual coordinates. If this is also the case for the potential, i.e., if one can decompose $V(\hat{\vec{r}}) = V_x(x) + V_y(y) + V_z(z)$, we can look for solutions of the time-independent Schrödinger equation among separable states according to Ex. 2.26(c). A simple example of such a situation is the free-space case with $V(\vec{r}) = 0$. Indeed, the 3D de Broglie wave (4.6), which represents an eigenstate of this Hamiltonian, is a product of de Broglie waves for the individual coordinates.

Exercise 4.3. Show that the state $|\vec{p}\rangle$ is an eigenstate of the kinetic energy operator $\hat{p}^2/2M = (\hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2)/2M$ with eigenvalue $(p_x^2 + p_y^2 + p_z^2)/2M$.

Another example is given in the following exercise.

Exercise 4.4.* Find the energy eigenvalues and their degeneracy for a three-dimensional isotropic harmonic oscillator with $V(\vec{r}) = M\omega^2 r^2/2$, where $r^2 = x^2 + y^2 + z^2$.

Generally, however, the potential is not a sum of potentials for individual coordinates. As a result, the evolution under the Hamiltonian (4.7) will typically *entangle* states that were initially tensor products of vectors in $\mathbb{V}_x$, $\mathbb{V}_y$, and $\mathbb{V}_z$. The eigenstates of the Hamiltonian will also be entangled with respect to the three component spaces. To illustrate this point, let us write the time-independent Schrödinger equation for 3D motion in the position basis.

Exercise 4.5. Show that, in the position basis:

a) the action of a component of the momentum operator on an arbitrary state $|\psi\rangle$ in the position representation is $\langle\vec{r}|\,\hat{p}_i\,|\psi\rangle = -i\hbar\frac{\partial}{\partial r_i}\langle\vec{r}|\,\psi\rangle = -i\hbar\frac{\partial}{\partial r_i}\psi(\vec{r})$;

b) the action of the momentum operator *vector* in the position basis is $\langle\vec{r}|\,\hat{\vec{p}}\,|\psi\rangle = -i\hbar\vec{\nabla}\langle\vec{r}|\,\psi\rangle = -i\hbar\vec{\nabla}\psi(\vec{r})$, where $\vec{\nabla} = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z})$ (in other words, in the position basis, $\hat{\vec{p}} \simeq -i\hbar\vec{\nabla}$);

c) the time-independent Schrödinger equation takes the form

$$\left[\frac{\hat{p}^2}{2M} + V(\vec{r})\right]\psi(\vec{r}) = E\psi(\vec{r}), \tag{4.8}$$

or

$$\left[-\frac{\hbar^2}{2M}\nabla^2 + V(\vec{r})\right]\psi(\vec{r}) = E\psi(\vec{r}), \tag{4.9}$$

where $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$ is the *Laplacian*.

We have obtained a three-dimensional partial differential equation. Its solution typically cannot be written as a product of functions of individual Cartesian variables — a manifestation of the entanglement mentioned above.

Solving Eq. (4.9) in its general form is a formidable task. Fortunately, physical problems that require such an undertaking are relatively rare. Usually, the potential has some symmetries that simplify the solution. We study one such case next.

4.2 Rotationally symmetric potential

4.2.1 Spherical coordinates

Consider a rotationally invariant potential $V(\vec{r}) = V(r)$, where $r = \sqrt{x^2+y^2+z^2}$ is the length of the radius vector to the point (x,y,z) — such as the potential of the electric field that an atomic nucleus imposes on electrons. If we learn how to solve the time-independent Schrödinger equation for that potential, we will be able to calculate stationary state wavefunctions of an electron in an atom.

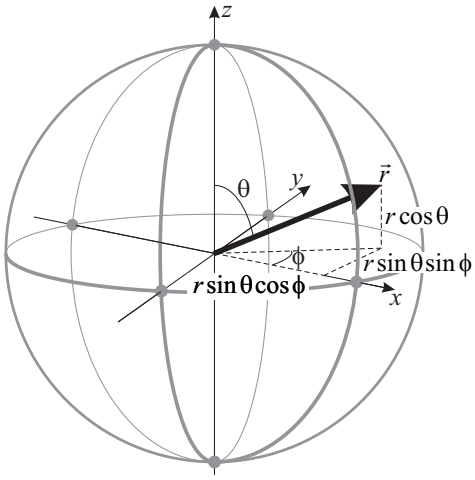


Fig. 4.1 Spherical coordinates.

How would we calculate the motion of a particle in a rotationally invariant potential classically? Perhaps we would consider the two degrees of freedom of such motion — radial and angular, and notice that they are largely decoupled from each other because the angular momentum is conserved. This decoupling would allow us to write and solve the equations of motion for each degree of freedom separately. Mathematically, this would correspond to using spherical, rather than Cartesian, coordinates — thereby significantly simplifying the calculation.

Our strategy in the quantum case will be quite similar. We start by presenting $\mathbb{V}_{3D}$ as a tensor product of Hilbert spaces associated with the spherical coordinates:

$$\mathbb{V}_{3D} = \mathbb{V}_r \otimes \mathbb{V}_\theta \otimes \mathbb{V}_\phi \tag{4.10}$$

with (Fig. 4.1)

$$x = r\sin\theta\cos\phi; \tag{4.11a}$$
$$y = r\sin\theta\sin\phi; \tag{4.11b}$$
$$z = r\cos\theta. \tag{4.11c}$$

Accordingly, the wavefunction $\psi(\vec{r})$ becomes a function of r, θ, and ϕ. The advantage of switching to spherical coordinates is that the rotationally symmetric potential is now an operator only in $\mathbb{V}_r$. The trade-off, however, is the kinetic energy. In contrast to the Cartesian case, it cannot be presented as a sum of terms each of which is local within its component Hilbert space. Nevertheless there is an advantage in using this approach, which we will see before the end of this section.

To proceed, we must introduce the rule for calculating inner products between two states whose wavefunctions are expressed in spherical coordinates. This inner

Box 4.1 Normalization in spherical coordinate Hilbert spaces

The extra factor $r^2 \sin\theta$ in Eq. (4.13) may appear strange. We derived the relationship (3.6) and its multidimensional analogue (4.5) from first principles, so one may argue that the inner product between two states expressed in *any* continuous basis must be of the same form, without any extra factors. The explanation is that Eq. (3.6) was derived using the normalization rule Eq. (3.1)(a) for the position eigenstates. The eigenstates of the three spherical observables do not have to follow this rule because they have different properties. For example, spherical coordinates can take values from limited ranges: $r \in [0, +\infty)$, $\theta \in [0, \pi]$, $\phi \in [0, 2\pi)$, in contrast to the position x, which ranges from $-\infty$ to $+\infty$.

A detailed study of this matter would lead us too deeply into the mathematical jungle, away from physics, so we will not undertake to do this. However, you can try it independently as an exercise. To this end, you would need to define the inner products of spherical coordinate eigenstates $\langle r_1 |\, r_2 \rangle, \langle \theta_1 |\, \theta_2 \rangle, \langle \phi_1 |\, \phi_2 \rangle$ and use them to obtain analogs of relations from Sec. 3.1, while making sure that they are consistent with each other and with Eq. (4.13).

product in the position basis is given by Eq. (4.5).In order to change the integration variables from Cartesian to spherical, we must include the Jacobian determinant:

$$\begin{aligned}\langle \psi |\, \varphi \rangle &\overset{(4.5)}{=} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \psi^*(\vec{r})\varphi(\vec{r})\mathrm{d}x\mathrm{d}y\mathrm{d}z \\ &= \int_0^{2\pi}\int_0^{\pi}\int_0^{\infty} \psi^*(\vec{r})\varphi(\vec{r})|J|\mathrm{d}r\mathrm{d}\theta\mathrm{d}\phi\end{aligned}$$

whose absolute value is given by

$$J = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial \phi} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial \phi} \end{vmatrix} = r^2 \sin\theta \tag{4.12}$$

For the inner product (4.5) we must therefore write

$$\langle \psi |\, \varphi \rangle = \int_0^{2\pi}\int_0^{\pi}\int_0^{\infty} \psi^*(\vec{r})\varphi(\vec{r}) r^2 \sin\theta \mathrm{d}r\mathrm{d}\theta\mathrm{d}\phi. \tag{4.13}$$

Exercise 4.6. Prove the second equality in Eq. (4.12).

It is a common convention to unite the two Hilbert spaces associated with the angular motion into a single tensor product space $\mathbb{Y} \equiv \mathbb{V}_\theta \otimes \mathbb{V}_\phi$, so that

$$\mathbb{V}_{3D} = \mathbb{V}_r \otimes \mathbb{Y} \tag{4.14}$$

Elements of the space $\mathbb{V}_r$ are represented by wavefunctions $R(r)$ of the radius, while the functions $Y_\lambda(\theta,\phi)$ of the two angles define elements in $\mathbb{Y}$.

Given Eq. (4.13), it is natural to define inner products for spaces $\mathbb{V}_r$ and $\mathbb{Y}$ as follows:

$$\langle R_1 | R_2 \rangle = \int_0^\infty R_1^*(r) R_2(r) r^2 \mathrm{d}r \tag{4.15a}$$

$$\langle Y_1 | Y_2 \rangle = \int_0^{2\pi} \int_0^{\pi} Y_1^*(\theta,\phi) Y_2(\theta,\phi) \sin\theta \mathrm{d}\theta \mathrm{d}\phi, \tag{4.15b}$$

where $R_{1,2}(r)$ and $Y_{1,2}(\theta,\phi)$ are wavefunctions of arbitrary states $|R_{1,2}\rangle$ and $|Y_{1,2}\rangle$ in $\mathbb{V}_r$ and $\mathbb{Y}$, respectively.

Exercise 4.7. Show that

a)[§] inner products (4.15) are consistent with Defn. A.9;
b) inner products (4.15) are consistent with the inner product in $\mathbb{V}_{3D}$, according to the definition (2.4) of the inner product in the tensor product space.

4.2.2 Angular momentum

Following in the footsteps of the classical treatment of motion in a rotationally invariant potential, we now introduce the quantum notion of *angular momentum* — the observable defined as

$$\hat{\vec{L}} = \hat{\vec{r}} \times \hat{\vec{p}}. \tag{4.16}$$

This is the cross (vector) product familiar from geometry and mechanics. It can be rewritten in a number of ways. We can write each component explicitly:

$$\begin{aligned} \hat{L}_x &= \hat{y}\hat{p}_z - \hat{z}\hat{p}_y; \\ \hat{L}_y &= \hat{z}\hat{p}_x - \hat{x}\hat{p}_z; \\ \hat{L}_z &= \hat{x}\hat{p}_y - \hat{y}\hat{p}_x. \end{aligned} \tag{4.17}$$

Alternatively, we can use the Levi-Civita symbol[3] to write

[3] The *Levi-Civita symbol*, also known as the *antisymmetric unit tensor of rank 3*, is defined as follows:

- For any j, k, l, the value of ε_{jkl} changes sign whenever any two indices are exchanged. Consequently, whenever any two indices are equal, $\varepsilon_{jkl} = 0$.
- $\varepsilon_{123} \equiv \varepsilon_{xyz} = 1$.

Explicitly,

$$\varepsilon_{xyz} = 1,\ \varepsilon_{xzy} = -1,\ \varepsilon_{zxy} = 1,\ \varepsilon_{zyx} = -1,\ \varepsilon_{yzx} = 1,\ \varepsilon_{yxz} = -1, \tag{4.18}$$

all other $\varepsilon_{jkl} = 0$.

$$\hat{L}_j = \varepsilon_{jkl}\hat{r}_k\hat{p}_l, \tag{4.19}$$

where we have used Einstein's convention that summation is implied over repeated indices (and we will continue to do so throughout this chapter).

Exercise 4.8. Show that the angular momentum operator is Hermitian.

Exercise 4.9.§ Show that the angular momentum operator in the position basis is represented by[4]

$$\begin{aligned}
\hat{L}_x &\simeq -\mathrm{i}\hbar\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right); \\
\hat{L}_y &\simeq -\mathrm{i}\hbar\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right); \\
\hat{L}_z &\simeq -\mathrm{i}\hbar\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right).
\end{aligned} \tag{4.20}$$

We will now proceed to derive the commutation properties of the angular momentum operator. This task is greatly simplified by using the Levi-Civita symbol. Therefore, I would recommend that you become comfortable with this symbol (unless you are already familiar with it from classical electrodynamics). In particular, we will need the identity in the following exercise.

Exercise 4.10. Show that

$$\varepsilon_{jkl}\varepsilon_{jmn} = \delta_{km}\delta_{ln} - \delta_{kn}\delta_{lm}. \tag{4.21}$$

Exercise 4.11. Check the following (for any arbitrary $j,k \in \{1,2,3\}$):

a) $[\hat{L}_j, \hat{r}_k] = \mathrm{i}\hbar\varepsilon_{jkl}\hat{r}_l$;
b) $[\hat{L}_j, \hat{p}_k] = \mathrm{i}\hbar\varepsilon_{jkl}\hat{p}_l$;
c) $[\hat{L}_j, \hat{L}_k] = \mathrm{i}\hbar\varepsilon_{jkl}\hat{L}_l$;
d) $[\hat{L}_j, \hat{r}^2] = 0$;
e) $[\hat{L}_j, \hat{p}^2] = 0$;
f) $[\hat{L}_j, \hat{L}^2] = 0$.

Exercise 4.12. Show that the definition (4.16) of angular momentum can be rewritten as $\hat{\vec{L}} = -\hat{\vec{p}} \times \hat{\vec{r}}$, in spite of the fact that the position and momentum observables do not generally commute.

---

[4] As discussed in Sec. 3.3.1 (see also Sec. A.2), the meaning of the symbol "$\simeq$" is that Eq. (4.20) applies to wavefunctions in the position basis alone. The full form of Eq. (4.20) would be

$$\langle \vec{r}|\ \hat{L}_x|\ \psi\rangle = -\mathrm{i}\hbar\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right)\psi(\vec{r}),$$

etc.

Exercise 4.13. Show that, if the potential is rotationally invariant [i.e., $V(\vec{r}) = V(r)$], then

a) each component $\hat{\vec{L}}_i$, and also the square $\hat{\vec{L}}^2$ of the angular momentum vector, commutes with the Hamiltonian (4.7);
b) in any state $|\psi\rangle$, the mean value of each component of the angular momentum is conserved: $\frac{\mathrm{d}}{\mathrm{d}t}\langle\psi|\hat{L}_i|\psi\rangle = 0$.

This result has a direct classical analogy: the angular momentum is conserved in a rotationally symmetric potential, in accordance with Noether's theorem.

Let us now incorporate the angular momentum observable into the Schrödinger equation.

Exercise 4.14. a) Show that

$$\hat{L}^2 = \hat{r}^2\hat{p}^2 - (\hat{\vec{r}}\cdot\hat{\vec{p}})^2 + \mathrm{i}\hbar\hat{\vec{r}}\cdot\hat{\vec{p}}. \tag{4.22}$$

How does this result change for the classical angular momentum?

b) Rewrite the time-independent Schrödinger equation (4.8) as

$$\left[\frac{(\hat{\vec{r}}\cdot\hat{\vec{p}})^2 - \mathrm{i}\hbar\hat{\vec{r}}\cdot\hat{\vec{p}}}{2M} + \frac{\hat{L}^2}{2M} + \hat{\vec{r}}^2 V(\vec{r})\right]\psi(\vec{r}) = \hat{\vec{r}}^2 E\psi(\vec{r}). \tag{4.23}$$

Equation (4.23) is beginning to look like the favorable case of separated variables discussed in the previous section. Indeed, each term on the left-hand side is a local operator in either $\mathbb{V}_r$ or $\mathbb{Y}$. The first term, for example, is expressed through the operator $\hat{\vec{r}}\cdot\hat{\vec{p}}$, whose classical analogue is proportional to the projection of the momentum onto the radius vector. We would expect this projection to affect only the radial degree of freedom, i.e., to be a local operator in $\mathbb{V}_r$. The second term — the angular momentum, — on the other hand, affects only the angular degree of freedom: it is local in $\mathbb{Y}$. The third term is of course local in $\mathbb{V}_r$ as long as the potential is rotationally invariant: $V(\vec{r}) = V(r)$.

In order to show this separability rigorously, we must convert the first two terms of Eq. (4.23), which are now known to us in Cartesian coordinates, into spherical coordinates. We shall do so using the chain rule for partial derivatives, known from multivariate calculus. The calculation is straightforward but quite tedious, so if you are not comfortable with it, you may just go over the solutions in the first reading.

Exercise 4.15*.

a) Show that

$$\frac{\partial}{\partial x} = \sin\theta\cos\phi\,\frac{\partial}{\partial r} + \frac{1}{r}\cos\theta\cos\phi\,\frac{\partial}{\partial\theta} - \frac{1}{r}\frac{\sin\phi}{\sin\theta}\,\frac{\partial}{\partial\phi}; \tag{4.24a}$$

$$\frac{\partial}{\partial y} = \sin\theta\sin\phi\,\frac{\partial}{\partial r} + \frac{1}{r}\cos\theta\sin\phi\,\frac{\partial}{\partial\theta} + \frac{1}{r}\frac{\cos\phi}{\sin\theta}\,\frac{\partial}{\partial\phi}; \tag{4.24b}$$

$$\frac{\partial}{\partial z} = \cos\theta\,\frac{\partial}{\partial r} - \frac{1}{r}\sin\theta\,\frac{\partial}{\partial\theta}. \tag{4.24c}$$

b) Derive the components of the angular momentum operator in spherical coordinates from those (4.20) in Cartesian coordinates:

$$\hat{L}_x \simeq \mathrm{i}\hbar\left(\sin\phi\,\frac{\partial}{\partial\theta} + \cot\theta\cos\phi\,\frac{\partial}{\partial\phi}\right), \tag{4.25a}$$

$$\hat{L}_y \simeq \mathrm{i}\hbar\left(-\cos\phi\,\frac{\partial}{\partial\theta} + \cot\theta\sin\phi\,\frac{\partial}{\partial\phi}\right), \tag{4.25b}$$

$$\hat{L}_z \simeq -\mathrm{i}\hbar\frac{\partial}{\partial\phi,} \tag{4.25c}$$

c) Show that

$$\hat{L}^2 \simeq -\hbar^2\left[\frac{1}{\sin\theta}\,\frac{\partial}{\partial\theta}\left(\sin\theta\,\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\,\frac{\partial^2}{\partial\phi^2}\right]. \tag{4.26}$$

d) Express operators $\hat{\vec{r}}\cdot\hat{\vec{p}}$ and $(\hat{\vec{r}}\cdot\hat{\vec{p}})^2$ in spherical coordinates:

$$\vec{r}\cdot\vec{p} \simeq -i\hbar r\frac{\partial}{\partial r}; \tag{4.27}$$

$$(\vec{r}\cdot\vec{p})^2 \simeq -\hbar^2\left(r^2\frac{\partial^2}{\partial r^2} + r\frac{\partial}{\partial r}\right). \tag{4.28}$$

We see that the expressions (4.25) for the angular momentum in spherical coordinates depend only on θ and ϕ, but not at all on r, while the opposite holds for the operator $\hat{\vec{r}}\cdot\hat{\vec{p}}$. This confirms our intuition: the first operator on the left-hand side of the time-independent Schrödinger equation (4.23) is local in space $\mathbb{V}_r$, while the second one is local in $\mathbb{Y}$. Let us now use this fact to solve the Schrödinger equation.

In Ex. 4.13 we found that the Hermitian operator $\hat{L}^2$ commutes with the Hamiltonian. As we know (Ex. 1.36), two commuting Hermitian operators have a common eigenbasis in which they both diagonalize. So it would appear that, in order to find energy eigenstates, it is enough to find the eigenstates of $\hat{L}^2$.

Unfortunately, this argument does not work in a straightforward fashion. The problem is, as we discussed, that $\hat{L}^2$ is local in $\mathbb{Y}$. Accordingly, the eigenstates of the counterpart operator $\hat{\mathbf{1}}\otimes\hat{L}^2$ in $\mathbb{V}_{3D}$ are given by $|R\rangle\otimes|\lambda\rangle$, where $|\lambda\rangle$ is an eigenstate of $\hat{L}^2$ in $\mathbb{Y}$, while $|R\rangle$ is an *arbitrary* state in $\mathbb{V}_r$ (Ex. 2.23).

In other words, each eigenvalue λ of the operator $\hat{\mathbf{1}} \otimes \hat{L}^2$ is massively degenerate[5], so there is no guarantee that any arbitrary state of the form $|R\rangle \otimes |\lambda\rangle$ is automatically an eigenstate of the Hamiltonian. We can only say that there *exists* an eigenbasis of the Hamiltonian such that each of its elements has the form $|R\rangle \otimes |\lambda\rangle$. So our strategy will be to to *select*, among states of the form $|R\rangle \otimes |\lambda\rangle$, those that are eigenstates of the Hamiltonian.

To perform the selection, let us write these states in the position basis

$$\psi(r,\theta,\phi) = R(r)Y_\lambda(\theta,\phi), \tag{4.29}$$

and demonstrate the following.

Exercise 4.16. Show that, in order for a wavefunction of the form (4.29) to represent an eigenstate of the Hamiltonian with eigenvalue E [i.e., satisfy the time-independent Schrödinger equation], it is necessary and sufficient that the radial part of the wavefunction (4.29) satisfy the *radial equation*

$$\left[-\frac{\hbar^2}{2Mr^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{\lambda}{2Mr^2} + V(r)\right] R(r) = ER(r). \tag{4.30}$$

We have thus divided the problem into two simpler ones: diagonalizing $\hat{L}^2$ and solving an ordinary differential equation (4.30).[6] Moreover, only the second part needs to be solved anew for each specific potential. The first does not depend on the potential at hand and hence needs to be solved only once. This will be our task in the next section.

4.3 Angular momentum eigenstates

4.3.1 Matrix representation of the angular momentum

Our task of finding the eigenvalues and eigenstates of the $\hat{L}^2$ observable is complicated by the following circumstance.

Exercise 4.17. Show that there exist degenerate eigenvalues of $\hat{L}^2$ in $\mathbb{Y}$.
Hint: Apply the fact that two observables are simultaneously diagonalizable if and only if they commute (Ex. 1.36) to operators $\hat{L}^2$, $\hat{L}_x$, and $\hat{L}_y$.

The above result means that an eigenstate of $\hat{L}^2$ may not be uniquely identified by the corresponding eigenvalue λ. As per Ex. A.70, we can say that each λ defines a *subspace* of eigenstates of $\hat{L}^2$, and that this subspace may not be of dimension one. We should find the basis and dimension for each of these subspaces.

[5] This is aside from the fact that the eigenstates of $\hat{L}^2$ are degenerate even in $\mathbb{Y}$, as we shall see in the next section.

[6] This approach is a particular case of the *variable separation* method for solving partial differential equations.

To this end, let us bring into the picture an additional observable in $\mathbb{Y}$ that commutes with $\hat{L}^2$. Then it will have a common set of eigenstates with $\hat{L}^2$ (see Ex. 1.36) and hence give rise to an orthonormal eigenbasis within each λ-subspace. With luck, this eigenbasis will be non-degenerate with respect to the eigenvalue μ of this new observable; a pair of eigenvalues λ, μ will then uniquely identify the state.

The traditionally chosen observable that satisfies this condition (as we shall see later) is $\hat{L}_z$[7]. So our task is to find common eigenstates $|\lambda\mu\rangle$ $\hat{L}^2$ and $\hat{L}_z$[8]

The wavefunctions of states $|\lambda\mu\rangle$ can, in principle, be found by solving

$$\hat{L}^2 |\lambda\mu\rangle = \lambda |\lambda\mu\rangle,$$
$$\hat{L}_z |\lambda\mu\rangle = \mu |\lambda\mu\rangle,$$

in the position basis using the differential operators (4.25c) and (4.26). However, this road would quickly lead us into an unwieldy mathematical mess. Fortunately, there is an alternative way. We may find out quite a lot about these states, the corresponding eigenvalues, and even the matrices of the components of the angular momentum operator, simply from the commutation relations between these components. Once we have these properties, we will still have to use a bit of calculus to determine the wavefunctions, but it will require much less effort than a direct calculation.

We will follow a strategy that is closely reminiscent of the technique we used with the harmonic oscillator in Sec. 3.8.2. We begin by defining the analogues of the creation and annihilation operators, the *raising and lowering operators*, as

$$\hat{L}_+ = \hat{L}_x + \mathrm{i}\hat{L}_y, \tag{4.31a}$$
$$\hat{L}_- = \hat{L}_x - \mathrm{i}\hat{L}_y. \tag{4.31b}$$

Exercise 4.18. Show that:

a) $\hat{L}_- = \hat{L}_+^\dagger$;
b) $[\hat{L}_z, \hat{L}_\pm] = \pm\hbar L_\pm$, $[\hat{L}^2, \hat{L}_\pm] = 0$, $[\hat{L}_+, \hat{L}_-] = 2\hbar L_z$;
c) $\hat{L}^2 = \hat{L}_+\hat{L}_- + \hat{L}_z^2 - \hbar\hat{L}_z = \hat{L}_-\hat{L}_+ + \hat{L}_z^2 + \hbar\hat{L}_z$.

Exercise 4.19. Suppose some state $|\lambda\mu\rangle$ is a common eigenstate of $\hat{L}^2$ and $\hat{L}_z$. Show that then

a) the state $\hat{L}_+ |\lambda\mu\rangle$ is also a common eigenstate of these operators with eigenvalues λ, $\mu + \hbar$;
b) the state $\hat{L}_- |\lambda\mu\rangle$ is also a common eigenstate of these operators with eigenvalues λ, $\mu - \hbar$.

Hint: try the same approach as in Ex. 3.61.

---

[7] We could equally well have chosen $\hat{L}_x$ or $\hat{L}_y$. We will look at a few examples to this effect later in this section.

[8] The notation $|\lambda\mu\rangle$ may mislead one into thinking that this state is tensor product. This is, of course, not the case: $|\lambda\mu\rangle$ is an element of a single Hilbert space $\mathbb{Y}$.

The above exercise shows that states $\hat{L}_+ |\lambda\mu\rangle$ and $\hat{L}_- |\lambda\mu\rangle$ are proportional to the normalized states $|\lambda, \mu+\hbar\rangle$ and $|\lambda, \mu-\hbar\rangle$, respectively. In the following, we find the proportionality coefficient.

Exercise 4.20. Show that, neglecting an arbitrary phase factor,

$$\hat{L}_+ |\lambda\mu\rangle = \sqrt{\lambda - \mu(\mu+\hbar)}\,|\lambda, \mu+\hbar\rangle; \tag{4.32a}$$

$$\hat{L}_- |\lambda\mu\rangle = \sqrt{\lambda - \mu(\mu-\hbar)}\,|\lambda, \mu-\hbar\rangle. \tag{4.32b}$$

Hint: Using Ex. 4.18(c), find $\langle\lambda\mu|\hat{L}_+\hat{L}_-|\lambda\mu\rangle$ and $\langle\lambda\mu|\hat{L}_-\hat{L}_+|\lambda\mu\rangle$ and reconcile the result with the statement of Ex. 4.19.

Exercise 4.21. Show that μ^2 cannot be greater than λ.

Exercise 4.22. Show that the condition of Ex. 4.21 can be satisfied only if $\lambda = \hbar^2 l(l+1)$ and $\mu = \hbar m$ with

- l being a nonnegative integer or half-integer ($0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$);
- for a given l, $m \in \{-l, -l+1, \ldots, l-1, l\}$.

Hint: apply the same logic as in Sec. 3.8.2 where we proved that the eigenvalues of the number operator of the harmonic oscillator must be integer.

This is one of the main results of this section. To put it another way, if we try to measure the observable $\hat{L}^2$ in some state, we can only detect values $\hbar^2 l(l+1)$, where $l \in \{0, \frac{1}{2}, 1, \frac{3}{2}, \ldots\}$. Furthermore, if we now prepare our system in a state with a given $\hat{L}^2$ (for example, by having measured it) and then perform the measurement of the observable L_z, we will detect one of the $2l+1$ possible values ranging from $-l\hbar$ to $l\hbar$ in steps of $\hbar$. This is because, as argued at the beginning of this section, the eigenvalues of $\hat{L}^2$ are indeed degenerate, and the degree of degeneracy (the number of orthogonal eigenstates corresponding to the same eigenvalue) is $2l+1$.

From now on we will use the notation $|lm\rangle$ instead of $|\lambda\mu\rangle$ to denote the common eigenstates of $\hat{L}^2$ and $\hat{L}_z$ with eigenvalues $\lambda = \hbar^2 l(l+1)$ and $\mu = \hbar m$, respectively. In the context of the motion of a point particle, the value of l is called the *orbital quantum number*[9], and m the *magnetic quantum number*.

Exercise 4.23[§] Show that Eqs. (4.32) can be rewritten as follows:

$$\hat{L}_+ |lm\rangle = \hbar\sqrt{l(l+1) - m(m+1)}\,|l, m+1\rangle = \hbar\sqrt{(l-m)(l+m+1)}\,|l, m+1\rangle; \tag{4.33a}$$

$$\hat{L}_- |lm\rangle = \hbar\sqrt{l(l+1) - m(m-1)}\,|l, m-1\rangle = \hbar\sqrt{(l+m)(l-m+1)}\,|l, m-1\rangle. \tag{4.33b}$$

Note that Ex. 4.22 establishes only the necessary conditions for the existence of common eigenstates of $\hat{L}^2$ and $\hat{L}_z$ with given eigenvalues. We do not yet know

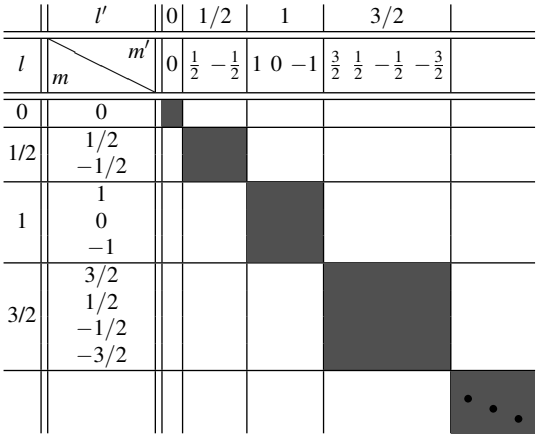
[9] Sometimes the orbital quantum number l is just called "angular momentum". This jargon term is used to suggest that the value of $\hbar l$ is the quantum counterpart to the classical absolute value of the angular momentum vector.

whether an eigenstate exists for a given pair (l,m), even if this pair satisfies the conditions, nor whether this eigenstate is unique. We will address this question in the next subsection. For now, let us just note that, if we take the uniqueness and existence of states $|lm\rangle$ for granted, it follows that they comprise an orthonormal basis in $\mathbb{Y}$ according to the spectral theorem (Ex. A.60). We will be calling the basis $\{|lm\rangle\}$ *canonical* in the context of angular momentum physics.

Exercise 4.24. Show that matrix elements $\langle lm|\,\hat{A}\,|l'm'\rangle$, where $\hat{A} = \hat{\vec{L}}^2, \hat{L}_+, \hat{L}_x, \hat{L}_y, \hat{L}_z$, vanish whenever $l \neq l'$ without calculating these matrix elements explicitly.

According to the above result, matrices of all components of $\hat{\vec{L}}$, as well as $\hat{L}^2$, have the structure shown in Table 4.1. It is a block-diagonal matrix, each block describing the angular momentum operator within the subspace of Hilbert space $\mathbb{Y}$ associated with a specific value of l. The size of each block is $(2l+1)\times(2l+1)$. In each block, values of m are traditionally listed in *decreasing* order.

Table 4.1 Structure of matrices of angular momentum operator components (Ex. 4.25). Shaded areas may contain nonzero matrix elements.



Exercise 4.25. Find the matrix elements $\langle lm|\,\hat{A}\,|l'm'\rangle$, where $\hat{A} = \hat{\vec{L}}^2, \hat{L}_+, \hat{L}_x, \hat{L}_y, \hat{L}_z$.

Exercise 4.26.§ Write the matrices of Ex. 4.25 explicitly for the Hilbert space subspaces associated with

a) $l = 1/2$,
b) $l = 1$.

Verify for both cases that the angular momentum matrices obey $\hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2 = \hat{L}^2$.

Answer:

a)

$$\hat{L}_x \simeq \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{L}_y \simeq \frac{\hbar}{2}\begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad \hat{L}_z \simeq \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \hat{L}^2 \simeq \frac{3\hbar^2}{4}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{4.34}$$

b)

$$\hat{L}_x \simeq \frac{\hbar}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \hat{L}_y \simeq \frac{\hbar}{\sqrt{2}}\begin{pmatrix} 0 & -\mathrm{i} & 0 \\ \mathrm{i} & 0 & -\mathrm{i} \\ 0 & i & 0 \end{pmatrix}, \hat{L}_z \simeq \hbar\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$
$$\hat{L}^2 \simeq 2\hbar^2\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{4.35}$$

Notice that the angular momentum matrices for the subspace $l = 1/2$ are proportional to the Pauli matrices [see Eq. (1.7)]. This identity explains the physics behind the subscripts x, y, and z that we have been assigning to these matrices throughout the course.

Exercise 4.27. Suppose you perform measurements of the x or y components of the angular momentum of a certain particle.

a) What possible values can the measurement yield if the particle is known to have been prepared in a state with

   i) $l = 1/2$,
   ii) $l = 1$?

   Answer:

   i) $\{\hbar/2, -\hbar/2\}$
   ii) $\{\hbar, 0, -\hbar\}$.

b) Find the states upon which the particle state will collapse, expressing them in the canonical basis.

The result of the last exercise — that the eigenvalues of $\hat{L}_x$ and $\hat{L}_y$ run from $-l\hbar$ to $l\hbar$ in steps of $\hbar$ — is not surprising. Although we have chosen $\hat{L}_z$ to help us in finding the basis of $\mathbb{Y}$, there is nothing unique about the z axis in terms of physical properties. The space is isotropic, so the observables $\hat{L}_x$ and $\hat{L}_y$ behave under application of quantum measurements in the same way as $\hat{L}_z$. Moreover, the same properties would be observed if we looked at the projection of the angular momentum vector onto an *arbitrary* axis.

Exercise 4.28. Consider the observable $\hat{L}_{\theta\phi}$ defined by the projection of the angular momentum onto the unit vector $\vec{R}_{\theta\phi}$ characterized by spherical angles (θ, ϕ). Restrict your analysis to the subspace with $l = 1/2$.

a) Show that the eigenvalues of $\hat{L}_{\theta\phi}$ are $\pm\hbar/2$ and find the corresponding eigenstates in the eigenbasis of $\hat{L}_z$.

Hint: Find the matrix of $\hat{L}_{\theta\phi} = \sin\theta\cos\phi\hat{L}_x + \sin\theta\sin\phi\hat{L}_y + \cos\theta\hat{L}_z$ in the eigenbasis of $\hat{L}_z$.

b) Find the mean values of $\hat{L}_x, \hat{L}_y, \hat{L}_z$ in these states and show that they are proportional to the projections of vector $\vec{R}_{\theta\phi}$ onto the corresponding coordinate axes.
Answer:

$$(\langle L_x\rangle, \langle L_y\rangle, \langle L_z\rangle) = \pm\frac{\hbar}{2}\vec{R}_{\theta\phi} \text{ for the eigenvalues } \pm\frac{\hbar}{2}. \tag{4.36}$$

Before we end our discussion of angular momentum matrices, let us take a brief look at Heisenberg's uncertainty principle.

Exercise 4.29. Find the expectation values and uncertainties of the operators $\hat{L}_x$ and $\hat{L}_y$ in the state $|lm\rangle$. Verify the uncertainty principle. Does the inequality saturate for any values of l or m?
Answer:

$$\langle L_x\rangle = \langle L_y\rangle = 0;$$
$$\langle \Delta L_x^2\rangle = \langle \Delta L_x^2\rangle = \frac{\hbar^2}{2}[l(l+1) - m^2].$$

The uncertainty principle takes the form

$$\langle \Delta L_x^2\rangle \langle \Delta L_y^2\rangle \geq \frac{\hbar^4}{4}m^2.$$

It is instructive to look at the uncertainty principle for states with $m = \pm l$, such that L_z takes its maximum possible value for a given L^2. Classically, this would imply that $L_x^2 = L_y^2 = 0$. But in the quantum case, $\langle L_z^2\rangle = l^2\hbar^2$, which is less than $\langle L^2\rangle = l(l+1)\hbar^2$. Therefore there is still "room left to play" for the x and y components of the angular momentum: $\langle \hat{L}_x^2\rangle = \langle \hat{L}_y^2\rangle = (\langle \hat{L}^2 - \hat{L}_z^2\rangle)/2 = \hbar^2[l(l+1) - l^2]/2 = \hbar^2 l/2$. This ensures that the uncertainty principle can be upheld for these components.

4.3.2 Wavefunctions of angular momentum eigenstates

Remarkably, everything we learned in the previous subsection — and we learned quite a bit — follows solely from the commutation relations between angular momentum components that we derived in Ex. 4.11. Aside from these relations, we never made direct use of the definition of that observable, nor of any of its physical properties. But now our goal is to find the wavefunctions of states $|lm\rangle$. To achieve this, we do require the explicit expressions for operators $\hat{L}^2$ and $\hat{L}_z$ in the position basis, which we calculated in Ex. 4.15.

Exercise 4.30. Show that the wavefunction of any eigenstate of operator $\hat{L}_z$ with eigenvalue m must be of the form

$$T(\theta)\mathrm{e}^{\mathrm{i}m\phi}. \tag{4.37}$$

Exercise 4.31.§ Show that the raising and lowering operators in the position basis are given by

$$\hat{L}_+ \simeq \hbar \mathrm{e}^{\mathrm{i}\phi}\left(\frac{\partial}{\partial\theta} + \mathrm{i}\cot\theta\frac{\partial}{\partial\phi}\right); \tag{4.38a}$$

$$\hat{L}_- \simeq \hbar \mathrm{e}^{-\mathrm{i}\phi}\left(-\frac{\partial}{\partial\theta} + \mathrm{i}\cot\theta\frac{\partial}{\partial\phi}\right). \tag{4.38b}$$

Hint: use Eqs. (4.25) and (4.31).

Exercise 4.32. Show, by induction, that the wavefunctions of states $|lm\rangle$ are given by the *spherical harmonics*[10]

$$Y_l^m(\theta,\phi) = \mathscr{N}_l\sqrt{\frac{(l+m)!}{(l-m)!}}\sin^{-m}\theta\frac{d^{l-m}}{d(\cos\theta)^{l-m}}\sin^{2l}\theta\mathrm{e}^{\mathrm{i}m\phi}, \tag{4.39}$$

where

$$\mathscr{N}_l = (-1)^l\sqrt{\frac{2l+1}{4\pi}}\frac{1}{2^l l!} \tag{4.40}$$

is the normalization factor[11], by applying the following steps.

a) Application of the raising operator to the state $|lm\rangle$ with $m = l$ must yield zero according to Eq. (4.33a). Check that this applies to the wavefunction $Y_l^l(\theta,\phi)$ of the state $|ll\rangle$ given by Eq. (4.39).
b) Check the correctness of the normalization factor (4.40).
Hint:
$$\int_{-1}^{1}(1-x^2)^l\mathrm{d}x = \frac{2^{2l+1}(l!)^2}{(2l+1)!}. \tag{4.41}$$
c) Apply the operator $\hat{L}^2$, which in the position basis is given by Eq. (4.26), to $Y_l^l(\theta,\phi)$, in order to check that this function represents an eigenstate of $\hat{L}^2$ with eigenvalue $l(l+1)\hbar^2$.
d) Suppose the wavefunction of the state $|lm\rangle$ is given by Eq. (4.39) for some m. Apply the lowering operator (4.38b) to show that Eq. (4.39) also expresses the wavefunction of the state $|l,m-1\rangle$.

Note that suffices to check that $Y_l^m(\theta,\phi)$ is normalized and an eigenwavefunction of $\hat{L}^2$ for $m = l$ alone, as we have done in parts (b) and (c). This is because we

[10] The standard definition of spherical harmonics uses associated Legendre polynomials. Our definition, borrowed from R. Shankar, *Principles of quantum mechanics* (Kluwer, 1990), bypasses these polynomials, and is hence less cumbersome. It is consistent with the convention most commonly used in quantum mechanics.

[11] The factor $(-1)^l$ in Eq. (4.40) is conventional.

already know, according to Eq. (4.33), that the lowering operator preserves both the eigenvalue of $\hat{L}^2$ and the normalization (with a factor $\sqrt{(l+m)(l-m+1)}$).

Exercise 4.33:§ Calculate explicitly the spherical harmonics for all possible values of m consistent with $l = 0$ and $l = 1$.
Answer:

$$Y_0^0(\theta,\phi) = \sqrt{\frac{1}{4\pi}};$$
$$Y_1^1(\theta,\phi) = -\sqrt{\frac{3}{8\pi}}\sin\theta e^{i\phi};$$
$$Y_1^0(\theta,\phi) = \sqrt{\frac{3}{4\pi}}\cos\theta;$$
$$Y_1^{-1}(\theta,\phi) = \sqrt{\frac{3}{8\pi}}\sin\theta e^{-i\phi}.$$

Absolute values of spherical harmonics up to $l = 2$ are plotted in Fig. 4.2. According to what we found in Ex. 4.30, these absolute values are independent of ϕ and hence axially symmetric.

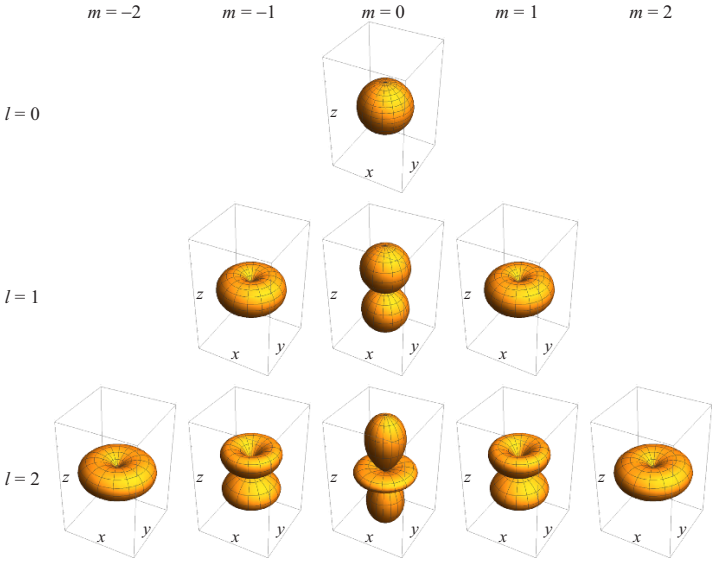


Fig. 4.2 Absolute values of spherical harmonics for the first three values of l, plotted as radii dependent on spherical angles θ and ϕ.

Earlier in this section, when we derived the conditions on physically allowed values of l and m, I mentioned that these are only necessary conditions, and not all of them may be realized in nature. By explicitly calculating the wavefunctions of states

$|lm\rangle$, we have proven the existence (and uniqueness) of these states, but only for integer l and m. Indeed, spherical harmonics contain a factor $e^{im\phi}$. For a half-integer l, m is also half-integer, and this factor results in $\psi(r,\theta,\phi) = -\psi(r,\theta,\phi+2\pi)$, which is impossible. Therefore, a point-like particle in a radially symmetric field must have an *integer* orbital quantum number.

4.3.3 Spin

Particles can nevertheless have an "intrinsic" angular momentum — the *spin* $\hat{\vec{S}}$. It can be visualized as the rotation of the particle about its own axis — in contrast to the "orbital" motion of a point particle in an external field that we studied previously. The spin degree of freedom follows the rules for angular momentum eigenstates derived in Sec. 4.3.1. In particular, possible eigenvalues of the observable $\hat{\vec{S}}^2$ are given by $s(s+1)\hbar^2$, with s a non-negative integer or half-integer number[12]. Because the spin degree of freedom has no position basis representation, s is allowed to take on half-integer values.

Wolfgang Pauli

The specific value of s is determined by the nature of the particle and cannot be controlled by external means. For example, electrons, protons, and neutrons have $s = \frac{1}{2}$, while photons have $s = 1$.

Physicists sometimes use the term "spin" to refer to this value of s — just as they use the term "angular momentum" to refer to the value of l — even though these values do not represent actual absolute values of $\hat{\vec{S}}$ or $\hat{\vec{L}}$. For example, they say that the electron has spin $\frac{1}{2}$.

Particles with half-integer spin are called *fermions*, while those with integer spin are called *bosons*. According to Wolfgang *Pauli's exclusion principle*, two identical fermions cannot be in the same quantum state. This principle is essential for many physical phenomena, for example, the periodic table of elements (Sec. 4.4.3). However, the physical reasons behind Pauli's principle require an understanding of quantum electrodynamics, which we do not study here.

The z component $\hat{S}_z$ of the spin observable has eigenvalues given by $m_s\hbar$, where $m_s \in \{-s,\dots,s\}$ is called the *spin quantum number*. In contrast to s, projections of the spin observable of a particle onto particular axes are not determined by the nature of the particle. We can prepare spin states with any values of m_s within the range allowed by the particle's spin, as well as arbitrary superpositions of them.

[12] In application to spin, the symbol s is normally used instead of l. The symbol l is reserved for the orbital angular momentum.

4.4 The hydrogen atom

4.4.1 Radial wavefunctions

In Sec. 3.5 I mentioned that one of the primary motives behind our interest in the time-independent Schrödinger equation is that it allows us to obtain the energy levels of electrons in atoms. Because transitions between energy levels are associated with absorption or emission of an optical photon, these theoretical calculations can be directly verified experimentally. Now we are fully equipped to calculate the energy levels and the corresponding wavefunctions of the hydrogen atom. The precise agreement between the results of this calculation and the experimental data on the emission spectrum of atomic hydrogen was one of the most significant triumphs of quantum mechanics (see Box 3.2).

In the hydrogen atom, an electron moves in the electrostatic potential created by a heavy nucleus:

$$V(r) = -\frac{1}{4\pi\varepsilon_0}\frac{e^2}{r}, \tag{4.42}$$

where e is the electron charge and ε_0 is the electric constant (we use the SI system of units). Hence the hydrogen atom problem is a special case of motion in a central field. Therefore we can follow our strategy developed in Sec. 4.2.2, namely, to look for an energy eigenwavefunction in the form of a product (4.29). In that product, as we now know, $\lambda = \hbar^2 l(l+1)$ and the angular component of the wavefunction $Y_\lambda(\theta,\phi) = Y_l^m(\theta,\phi)$ is one of the spherical harmonics, so we can rewrite it as

$$\psi_{Elm}(r,\theta,\phi) = R_{El}(r)Y_l^m(\theta,\phi). \tag{4.43}$$

All we need to do now is to find the radial component, which we will denote by $R_{El}(r)$.

Exercise 4.34.§ Write the radial equation (4.30) for the hydrogen atom.
Answer:

$$\left[-\frac{\hbar^2}{2M}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{\hbar^2 l(l+1)}{2Mr^2} - \frac{1}{4\pi\varepsilon_0}\frac{e^2}{r}\right] R_{El}(r) = ER_{El}(r). \tag{4.44}$$

Even though it is an ordinary differential equation, it is quite difficult to solve. The first step in simplifying it is a simple variable replacement.

Exercise 4.35. Redefine

$$R_{El}(r) = U_{El}(r)/r \tag{4.45}$$

and rewrite Eq. (4.44) for $U_{El}(r)$.
Answer:

$$\left[-\frac{\hbar^2}{2M}\frac{\partial^2}{\partial^2 r} + \frac{\hbar^2 l(l+1)}{2Mr^2} - \frac{1}{4\pi\varepsilon_0}\frac{e^2}{r}\right] U_{El}(r) = EU_{El}(r). \tag{4.46}$$

A common tactic in solving differential equations is to try guessing the general form of the solution, and then adjust its parameters to satisfy the equation. In the present case, we will look for a solution of the form

$$U_{El}(r) = \sum_{j=l+1}^{n} A_j r^j \mathrm{e}^{-\kappa r}, \tag{4.47}$$

where n is some natural number $A_{l+1} \neq 0$, and

$$\kappa = \sqrt{-2ME}/\hbar. \tag{4.48}$$

The following exercise helps to see how we arrived at this guess.

Exercise 4.36. Show that the asymptotic behavior of $U_{El}(r)$ given by the above equation is consistent with Eq. (4.46) for $r \to 0$ and $r \to \infty$.

Let us now find the coefficients A_j and the upper summation limit in Eq. (4.47).

Exercise 4.37. Show that, in order for Eq. (4.46) to hold, the following relationship must be satisfied:

$$\left[2\kappa j - \frac{2}{a}\right] A_j + [l(l+1) - j(j+1)] A_{j+1} = 0, \tag{4.49}$$

where

$$a = \frac{4\pi\varepsilon_0\hbar^2}{Me^2} \approx 0.53 \text{ Å}. \tag{4.50}$$

The latter quantity has dimensions of length and is known as the *Bohr radius*. We shall discover its physical meaning shortly.

We see from Eq. (4.49) that, for large j, $A_{j+1}/A_j \to 2\kappa/j$. If the series (4.47) with such a property were infinite ($n = \infty$), it would diverge. Indeed, in the limit $j \to \infty$ we would have $A_j \sim (2\kappa)^j/j!$ and hence for $r \to \infty$

$$U_{El}(r) \sim \sum_j \frac{(2\kappa r)^j}{j!} \mathrm{e}^{-\kappa r} \to \mathrm{e}^{2\kappa r} \mathrm{e}^{-\kappa r} = \mathrm{e}^{\kappa r},$$

where we have used the decomposition of the exponent into a Taylor series. As we know, a wavefunction that tends to infinity is unphysical.

To prevent this, we must require that the series be finite. This is fulfilled if the factor in front of A_j in Eq. (4.49) vanishes for some $j = n$. In this case

$$2\kappa n = \frac{2}{a} \tag{4.51}$$

and all A_j vanish for $j > n$.

Exercise 4.38. Calculate the radial wavefunctions $R_{nl}(r)$ of the hydrogen atom for

a) $n = 1, l = 0$;

b) $n = 2, l = 0$;
c) $n = 2, l = 1$.

Normalize the wavefunctions according to $\int |\psi(\vec{r})|^2 d^3\vec{r} = 1$.
Hint:

$$\int_0^\infty x^n \mathrm{e}^{-x} dx = n! \tag{4.52}$$

Answer (Fig. 4.3):

$$R_{10}(r) = 2a^{-3/2}\mathrm{e}^{-r/a}; \tag{4.53}$$

$$R_{20}(r) = \frac{1}{\sqrt{2}} a^{-3/2} \left(1 - \frac{r}{2a}\right) \mathrm{e}^{-r/2a}; \tag{4.54}$$

$$R_{21}(r) = \frac{1}{\sqrt{24}} a^{-5/2} r \mathrm{e}^{-r/2a}. \tag{4.55}$$

We now see the physical meaning of the Bohr radius: it determines the characteristic length scale of wave functions of energy eigenstates, as well as the approximate size of the ground state orbital.

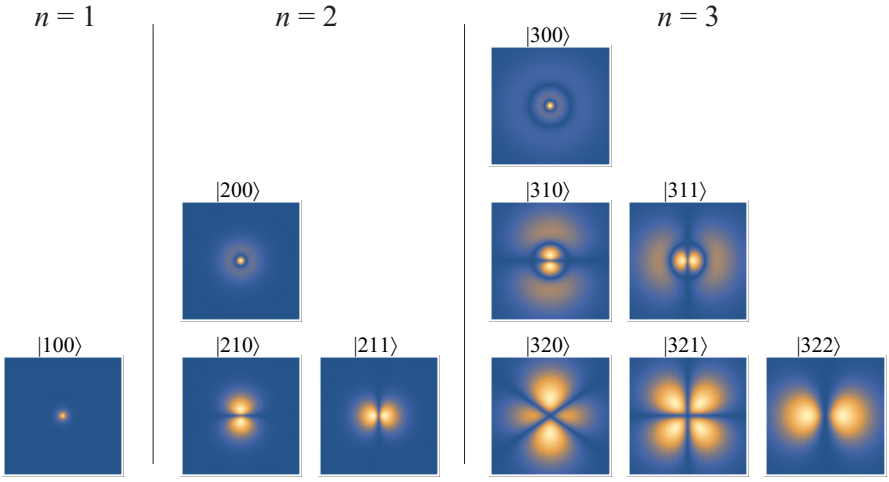


Fig. 4.3 Absolute values of the wavefunctions of a few lowest eigenstates $|nlm\rangle$ of the hydrogen atom. Cross-sections through the x-z plane are shown; the ranges are from $-20a$ to $20a$ in both dimensions.

4.4.2 Energy spectrum and transitions

Combining Eqs. (4.48) and (4.51), we find

$$E_n = -\frac{1}{2M}\left(\frac{\hbar}{na}\right)^2 = -\left(\frac{e^2}{4\pi\varepsilon_0}\right)^2 \frac{M}{2n^2\hbar^2}. \tag{4.56}$$

This result marks a major milestone: we calculated the energy spectrum of the hydrogen atom[13].

Interestingly, although the radial wavefunctions do depend on the orbital quantum number l, the energy eigenvalues (4.56) don't. Rather, they are determined by the upper limit n of the sum 4.47. For this reason, n is called the *principal quantum number*.

Each energy level, identified by the value of n, is degenerate with respect to the orbital quantum number l, which can take on any integer value from 0 to $n-1$. But the actual degeneracy of energy levels is even higher. To see this, we recall that the wavefunction (4.43) of the electron in the hydrogen atom has an angular part in addition to the radial one. The angular part of the wavefunction depends on the magnetic quantum number m, which does not affect the energy. Additionally, each electron has a spin degree of freedom, which is associated with a two-dimensional Hilbert space.

Exercise 4.39. Show that the degree of degeneracy of the energy level with principal quantum number n is $2n^2$.

Before proceeding, let us agree on a convention. Because the principal, orbital, and magnetic quantum numbers define the motional state of the electron in the atom, we will use the notation $|nlm\rangle$ to identify that state, rewriting Eq. (4.43) as follows:

$$\psi_{nlm}(r,\theta,\phi) = R_{nl}(r)Y_l^m(\theta,\phi). \tag{4.57}$$

So far, we have assumed the nucleus to be infinitely heavy, so the electron moves in the stationary potential (4.42). But it is easy to take a finite mass of the nucleus into account. As we know from classical mechanics, the two-body problem in the center-of-mass reference frame can be reduced to the motion of a single particle with the *reduced mass*

$$M = \frac{M_e}{1+M_e/M_p},$$

where M_e and M_p are in our case the rest masses of the electron and the nucleus (proton). The reduced mass is smaller than the mass of the electron by a factor of $1/1836$.

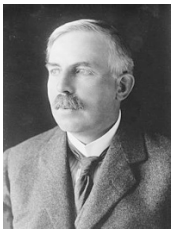
Equation (4.56) establishing the energy levels of the hydrogen atom can be expressed in the form

[13] Energies are negative, as expected for bound states.

Box 4.2 The model of the atom: a brief history

Although the idea of the atom dates back to ancient Greek philosophers (the very word "atom" is of Greek origin, meaning "uncuttable"), the first physical model thereof was proposed in 1904 by J. J. Thomson shortly after his discovery of the electron. He hypothesized that the negatively charged electrons were situated inside a blob of positively charged matter, like plums inside a pudding.

Thomson's hypothesis was refuted by experiments led by Ernest Rutherford, in which metal foil was bombarded by alpha particles. Rutherford and his colleagues found that, while most particles made it through the foil as though it were not present, a very small fraction ($\sim 1/8000$) were reflected backwards. Rutherford interpreted this observation as evidence that positive charges inside atoms are concentrated in tiny but heavy nuclei. Subsequently, in 1911, Rutherford put forward the planetary model of the atom*, according to which electrons orbit the nucleus akin to planets around the Sun. A legend says that one morning Rutherford entered his lab with a loud announcement: "I know what the atom looks like!"



Ernest Rutherford

Rutherford's model, however, had a major shortcoming that he and his colleagues immediately realized. Orbiting around the nucleus, the electron will create alternating electric and magnetic fields around itself, bringing about an electromagnetic wave that will carry the electron's energy away. As a result, the electron will fall upon the nucleus within picoseconds.

Niels Bohr

Rutherford asked his associate, a young theoretician Niels Bohr, to resolve this conundrum. Within two years, Bohr found a partial solution**. He postulated that there exists a discrete set of "stationary" orbits in which the electron can stay without radiating. Specifically, the orbit is stationary if its angular momentum equals an integer number of $\hbar$:

$$pr = n\hbar, \tag{4.58}$$

If the electron transitions between these orbits, it emits or absorbs a photon whose energy is equal to the energy difference between the levels. The spectrum of optical transitions of the hydrogen atom that Bohr calculated using his model (see Ex. 4.41) turned out to be consistent with the Rydberg formula (4.61), which was already known empirically at that time (see Box 4.3), and demonstrated excellent agreement with experiment.

A shortcoming of Bohr's model was its purely empirical nature. Although it seemed to explain experimental results, the physics behind it remained a mystery. Some light was shed on the physics by Louis de Broglie in 1924. He reconciled Bohr's model with his matter wave concept (see Box 3.2 and Ex. 4.42). In subsequent years, the model of the atom underwent a number of refinements, most notably by Wolfgang Pauli in 1926, gradually approaching the modern form that we study here.

*E. Rutherford, *The Scattering of α and β Particles by Matter and the Structure of the Atom*, Philosophical Magazine. **21**, 669 (1911).

N. Bohr, *On the Constitution of Atoms and Molecules*, Philosophical Magazine **26, 1–24 and 476–502 (1913).

$$E_n = -\frac{1}{1+M_e/M_p}\frac{\mathrm{Ry}}{n^2}, \tag{4.59}$$

where

$$\mathrm{Ry} = -\frac{e^4 M_e}{32\pi^2\varepsilon_0^2\hbar^2} \approx 2.17987217\times 10^{-18} J \approx -13.6056925\ \mathrm{eV} \tag{4.60}$$

is the *Rydberg constant*. This is one of the most significant, and most precisely measured fundamental physical constants. Because hydrogen is ubiquitous in the universe, its emission arrives on Earth from a wide variety of astronomical objects. Some of that emission was generated in the early stages of the existence of the universe. By measuring its spectrum, we can find out whether the value of the Rydberg constant, and therefore the fundamental laws of physics, have changed during the universe's lifetime.

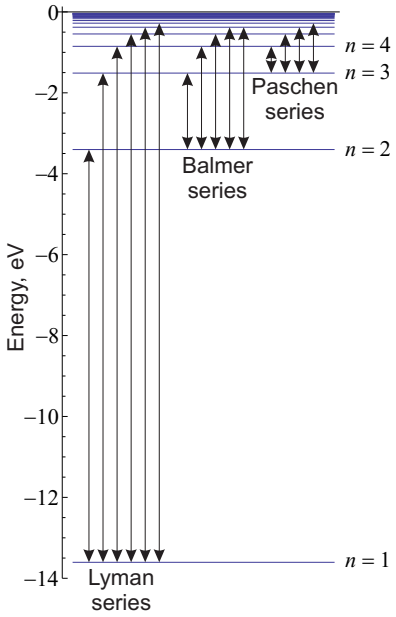


Fig. 4.4 Energy spectrum of atomic hydrogen.

Exercise 4.40. Using Bohr's postulate that transitions between atomic levels are accompanied by absorption or emission of a photon whose energy equals the difference between the energies of the levels, derive the equation (known as *Rydberg's formula*) for the wavelengths of the lines observed in the hydrogen spectrum:

$$\frac{2\pi\hbar c}{\lambda} = \frac{1}{1+M_e/M_p}\mathrm{Ry}\left|\frac{1}{n_1^2}-\frac{1}{n_2^2}\right|, \tag{4.61}$$

Box 4.3 Balmer's discovery

The discovery of Rydberg's formula is worth a separate story. The particular case for $n_1 = 2, n_2 \geq 3$ was discovered by Johann Jakob Balmer as early as 1885, almost thirty years before Bohr's model (Box 4.2). Remarkably, Balmer was not even a physicist; he was a Swiss school mathematics teacher. Apparently, as a hobby, he studied data on the solar spectrum published in 1868 by A. J. Ångström. These data included the following set of lines that were attributed to atomic hydrogen:

656.3 nm
486.1 nm
434.0 nm
410.2 nm

Johann Jakob Balmer

Driven solely by the belief that the world is governed by mathematical harmony, Balmer looked for and found a regularity behind this set of numbers. His expression for this regularity was similar to Eq. (4.61), except for the value of n_1 being equal to two. Three years later, in 1888, the Swedish physicist Johannes Rydberg became aware of Balmer's formula and generalized it to other values of n_1.

Of course, the reason why the series of lines that now carries Balmer's name was discovered first is that it lies within the visible spectrum. About twenty years later, the two series corresponding to $n_1 = 1$ and $n_1 = 3$ were measured in the ultraviolet and infrared ranges, respectively, by Theodore Lyman and Friedrich Paschen, and were found to be in excellent agreement with Rydberg's formula.

where n_1 and n_2 are positive integers.

Evaluate numerically the ranges of experimentally observable wavelengths of the transitions of the Lyman ($n_2 = 2,3,4,\ldots \to n_1 = 1$), Balmer ($n_2 = 3,4,5,\ldots \to n_1 = 2$), and Paschen ($n_2 = 4,5,6,\ldots \to n_1 = 3$) series (Fig. 4.4).

Exercise 4.41. Reproduce the result (4.56) for the hydrogen energy spectrum using the semi-classical Bohr theory (Box 4.2). Assuming the electron to be a point object moving in a circular orbit of radius r around the proton, obtain the relationship between the orbital radius and velocity by observing that the centripetal acceleration is due to the electrostatic attraction of the proton. Then reconcile this relation with Eq. (4.58) to find the parameters of the orbit as a function of n and determine the corresponding kinetic and potential energies.

Exercise 4.42. Reproduce the result (4.56) using de Broglie's equation (3.28) instead.

It may appear from the above two exercises that the full-blown quantum theory, as used in the previous subsection, is unnecessary to describe the hydrogen atom; the same results can be obtained in much simpler ways. But in fact, the approaches proposed by Bohr and de Broglie are of an *ad hoc* nature: while they yield a correct formula describing one particular observation, they cannot be used to reliably predict the results of any other experiment. Even if we only consider the hydro-

gen atom, the scope of possible questions extends far beyond merely cataloguing spectral lines. Answers to these questions can be found with the help of quantum mechanics, but not Bohr's or de Broglie's approaches.

Exercise 4.43. For the state $|n, l = n-1, m\rangle$, with an arbitrary principal quantum number n,

a) calculate the radial wavefunction;
b) calculate the mean and variance of the distance between the electron and the nucleus;
Answer: $\langle r\rangle = an\left(n+\frac{1}{2}\right)$.
c) compare your result with the one obtained from the Bohr model (Ex. 4.41).

Atoms in states with high principal quantum numbers are called *Rydberg atoms*. We see that these atoms are very large in size: the electron orbital radius scales as the second power of n. For example, the $n = 137$ state of hydrogen has an atomic radius $\sim$ 1 μm. Rydberg atoms have many interesting properties that make them a subject of intense research, particularly in application to quantum information processing.

Exercise 4.44. Find the expectation value and the uncertainty of the observables $\hat{x}$, $\hat{y}$, $\hat{z}$ in the state $|1,0,0\rangle$.

Exercise 4.45. Without calculation, determine which of the matrix elements of the observables $\hat{r}_i = \hat{x}, \hat{y}, \hat{z}$ in (a) $\langle 1,0,0|\,\hat{r}_i\,|2,0,0\rangle$, (b) $\langle 1,0,0|\,\hat{r}_i\,|2,1,0\rangle$, (c)$\langle 1,0,0|\,\hat{r}_i\,|2,1,\pm 1\rangle$ vanish.
Hint: The matrix elements have the form $\int r_i \psi_{nlm}(\vec{r})\psi_{n'l'm'}(\vec{r})\mathrm{d}^2 r$, with the wavefunctions given by Eq. (4.57). Use the symmetries of the spherical harmonics to determine whether the integrand is an even or odd function and find out how it depends on ϕ.

Exercise 4.46. Calculate the nonvanishing matrix elements from the previous exercise explicitly.

The above two exercises allow us to determine whether the transitions between the respective states in the hydrogen atom can occur due to interaction with the optical field. For example, they tell us whether an atom in state $|1,0,0\rangle$ can be excited into state $|2,1,1\rangle$ by a resonant laser that is polarized along the x axis or, on the contrary, whether an atom in state $|2,1,1\rangle$ can emit an x-polarized photon and transition into $|1,0,0\rangle$. This is because the mechanism of the light–atom interaction occurs through the coupling of the electric field with the atomic electric dipole, which has the form $\hat{\vec{d}} = e\hat{\vec{r}}$. The strength of the optical coupling is determined by the magnitude of the dipole moment matrix element associated with the relevant transition.

4.4.3 The periodic table

The periodic law, discovered by Dmitri Mendeleev in 1869, states that chemical properties of elements exhibit periodic dependence on the charge of their atomic

Table 4.2 Ground-state electron configurations for chemical elements up to $Z = 36$. For each element, the populations of energy levels defined by quantum numbers n and l are listed.

	n	1	2		3			4			
element	Z \ l	0	0	1	0	1	2	0	1	2	3
Hydrogen	1	1									
Helium	2	2									
Lithium	3	2	1								
Beryllium	4	2	2								
Boron	5	2	2	1							
Carbon	6	2	2	2							
Nitrogen	7	2	2	3							
Oxygen	8	2	2	4							
Fluorine	9	2	2	5							
Neon	10	2	2	6							
Sodium	11	2	2	6	1						
Magnesium	12	2	2	6	2						
Aluminum	13	2	2	6	2	1					
Silicon	14	2	2	6	2	2					
Phosphorus	15	2	2	6	2	3					
Sulfur	16	2	2	6	2	4					
Chlorine	17	2	2	6	2	5					
Argon	18	2	2	6	2	6					
Potassium	19	2	2	6	2	6		1			
Calcium	20	2	2	6	2	6		2			
Scandium	21	2	2	6	2	6	1	2			
Titanium	22	2	2	6	2	6	2	2			
Vanadium	23	2	2	6	2	6	3	2			
Chromium	24	2	2	6	2	6	5	1			
Manganese	25	2	2	6	2	6	5	2			
Iron	26	2	2	6	2	6	6	2			
Cobalt	27	2	2	6	2	6	7	2			
Nickel	28	2	2	6	2	6	8	2			
Copper	29	2	2	6	2	6	10	1			
Zinc	30	2	2	6	2	6	10	2			
Gallium	31	2	2	6	2	6	10	2	1		
Germanium	32	2	2	6	2	6	10	2	2		
Arsenic	33	2	2	6	2	6	10	2	3		
Selenium	34	2	2	6	2	6	10	2	4		
Bromine	35	2	2	6	2	6	10	2	5		
Krypton	36	2	2	6	2	6	10	2	6		

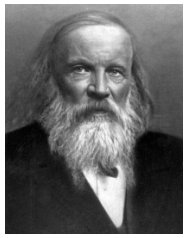
nucleus[14]. We can get some understanding of the periodic law by extending the physics of the hydrogen atom to other elements.

Atoms are normally neutral, so they have as many electrons as protons. Hydrogen has one proton and one electron, helium two of each, lithium three and so on. When the number of electrons in the atom exceeds one, they start interacting with each

[14] Mendeleev's original formulation stated that this periodic dependence is on the element's atomic weight, because the nucleus had not yet been discovered at that time.

other, and the problem of calculating their wavefunctions and energy levels becomes intractable. To deal with this complication, we shall first assume that the electrons do not interact with each other. This is, of course, a gross oversimplification, yet it will allow us to set up the "zero order approximation" as a basis for the discussion.

There are two fundamental principles we must take into account. The first is the principle of minimum energy. That is, the electrons will normally occupy the state (or one of the states) with the lowest possible energy (the *ground state*). This is justified because, if the atom is in thermal equilibrium with the environment, it follows from Boltzmann statistics: the probability of finding the atom in a state with energy E is proportional to $e^{-E/kT}$, where k is the Boltzmann constant and T the temperature of the environment. As long as $kT \ll E_1 - E_0$ (where $E_1 - E_0$ is the energy difference between the first higher energy state and the ground state), the probability for the atom to become excited is low.



Dmitri Mendeleev

Exercise 4.47. Estimate the probability for the hydrogen atom to become excited to a state with $n = 2$ at room temperature.
Hint: Don't forget to take into account the degeneracy of energy levels.

If multielectron atoms were governed entirely by the principle of minimum energy, all electrons would be found in the energy level with $n = 1$. However, this is prevented by the Pauli exclusion principle. As we found in Ex. 4.39, the $n = 1$ energy level (or *shell*, as chemists say) has room for only two electrons. If the atom has more than two electrons, the remaining ones will be pushed into the $n = 2$ shell, which has room for eight electrons, $n = 3$ with room for eighteen, and so on. The higher the atomic number, the more shells are occupied.

Let us now bring the interactions between electrons into the picture. The quantum many-body treatment can be simplified by observing that the electrons in different shells tend to interact only weakly with each other. This is because, as is evident from Fig. 4.3, lower shell electrons are, on average, much closer to the nucleus. Spatial overlaps of wavefunctions associated with different shells are relatively low, so the electrons spend little time in direct proximity with each other. From the point of view of the outer shell electrons, the inner shell electrons effectively act as a solid negatively charged sphere (hence the name "shell") around the nucleus, offsetting its attractive potential by their negative charge.

Chemical properties of an element are primarily determined by the electrons in the outermost occupied shell, the *valence shell*. This is because they have the highest energies (Fig. 4.4), and hence are more likely to enter chemical reactions. The crucial factor is the number of electrons in that shell. If it is *filled* (the Pauli principle does not allow additional electrons to join it), the ground state lies low, so the atom is reluctant to enter chemical reactions with other atoms — the element is then a noble gas. As we can see in Table 4.2, this is the case with helium (atomic number $Z = 2$) and neon ($Z = 2 + 8 = 10$). Note that the next noble gas, argon, has atomic number $Z = 18$, rather than $2 + 8 + 18 = 28$, so it does not follow this rule. I will explain this in a moment.

If the valence shell has only one electron (lithium $Z = 2 + 1 = 3$, sodium $Z = 10 + 1 = 11$, potassium $Z = 18 + 1 = 19$ etc.), this electron interacts weakly with those of inner shells, so it behaves as if it were the only electron in the atom. Such atoms are called alkali metals. They are likely to enter chemical reactions by giving away their single valence electron and becoming positively ionized. This is because the bound state energy of the outer electron is close to zero.

In halogens (fluorine $Z = 10 - 1 = 9$, chlorine $Z = 18 - 1 = 17$, etc.), on the other hand, the valence shell is missing a single electron, so it is beneficial for the atom to "steal" an electron and thereby fill its outer shell, yielding a low-lying energy eigenstate. This is why alkali metals and halogens tend to react strongly with each other and form stable substances such as table salt (NaCl).

In elements starting with potassium ($Z = 19$) in Table 4.2, the $n = 4$ shell starts to fill before those states with $n = 3, l = 2$. The reason is as follows. We found earlier that in the hydrogen atom, states with the same principal quantum number n but different orbital quantum numbers l have the same energy. As it turns out, this is a unique property of atoms or ions that have only one electron. Electrons with higher angular momenta are located, on average, further away from the nucleus. Therefore, in a multielectron atom, an electron in a state with a large l is shielded from the field of the nucleus by other electrons, and therefore has a higher energy than its peer with the same n, but lower l[15]. This property is especially prominent for high values of n and l. In particular, states with $n = 3, l = 2$ have higher energy than those with $n = 4, l = 0$. Therefore after argon, in which states with $n = 3$ and $l = 0, 1$ are filled, the fourth shell begins to fill even though there are still vacancies in the third shell. This explains why argon behaves like a noble gas.

Of course, the third shell would have to fill up, too, at some point. This happens for $Z = 21$ to 30, scandium to zinc. Because all these elements (except chromium and copper) have two electrons in the outer shell, they have largely similar chemical properties.

4.5 The Bloch sphere

In previous sections, we found the eigenstates of the operators associated with the projections of the angular momentum vector onto various axes. Let us now pose a converse problem. Can an arbitrary element of the Hilbert space be viewed as an eigenstate of the projection of the angular momentum onto a particular axis? In other words, can one associate an angular momentum vector of a certain direction with any motional state, as one does in classical physics? The answer turns out to be affirmative, but only for the subspace associated with $l = \frac{1}{2}$.

[15] The magnetic quantum number m, on the other hand, does not affect the energy, even in multielectron atoms.

Exercise 4.48. Consider an arbitrary normalized spin state $|\psi\rangle = \psi_\uparrow|\uparrow\rangle + \psi_\downarrow|\downarrow\rangle$. Without loss of generality, we define the overall quantum phase of that state such that $\psi_\uparrow$ is real and non-negative.

a) Show that, for any state $|\psi\rangle$, we can define a unique set of angles $\theta \in [0,\pi]$ and $\phi \in [0,2\pi)$ such that

$$\psi_\uparrow = \cos\frac{\theta}{2}; \tag{4.62a}$$

$$\psi_\downarrow = \sin\frac{\theta}{2}e^{i\phi}. \tag{4.62b}$$

b) Show that the state $|\psi\rangle$ is the eigenstate of the projection of the angular momentum $\hat{\vec{S}}$[16] onto the vector $\vec{R}_{\theta\phi}$ directed along the spherical angles θ,ϕ with the eigenvalue $\hbar/2$.

c) Show that the Cartesian coordinates of the tip of $\vec{R}_{\theta\phi}$ equal the mean values of the observables $\hat{\sigma}_x$, $\hat{\sigma}_y$, $\hat{\sigma}_z$ for the corresponding state $|\psi\rangle$.

Hint: recall Ex. 4.28.

Exercise 4.48 tells us that for every spin state $|\psi\rangle$, one can define a vector such that the spin in that state "points in the direction" of that vector. This vector is called the *Bloch vector* of state $|\psi\rangle$, and the full set of such vectors is called the *Bloch sphere*.

Exercise 4.49. Explain why a similar correspondence cannot be established for subspaces with angular momentum $l > \frac{1}{2}$.

Exercise 4.50.§ Check that the eigenstates of operators $\hat{S}_x$, $\hat{S}_y$, $\hat{S}_z$ correspond to the points on the Bloch sphere as shown in Fig. 4.5.

Exercise 4.51. Show that any two states that are represented by opposite points on the Bloch sphere are orthogonal.

The Hilbert space associated with a spin-$\frac{1}{2}$ particle is a qubit. Indeed, its basis consists of two elements: "spin-up" $|\uparrow\rangle$ and "spin-down" $|\downarrow\rangle$. This means that we can establish a one-to-one correspondence (isomorphism[17]) between spin states and those of any other qubit — for example, photon polarization states, mapping any spin state $\alpha|\uparrow\rangle + \beta|\downarrow\rangle$ onto a polarization state $\alpha|H\rangle + \beta|V\rangle$. Then the eigenstates

[16] We use the symbol $\hat{\vec{S}}$, rather than $\hat{\vec{L}}$, to emphasize that the subspace $l = 2$ can correspond only to a spin degree of freedom.

[17] An *isomorphism* $f(\cdot)$ between linear spaces $\mathbb{V}$ and $\mathbb{W}$ is a one-to-one mapping $|a\rangle \in \mathbb{V} \mapsto f(|a\rangle) \in \mathbb{W}$ such that, for any $|a\rangle, |b\rangle \in \mathbb{V}$ and number λ,

$$\begin{aligned} f(|a\rangle + |b\rangle) &= f(|a\rangle) + f(|b\rangle); \\ f(\lambda|a\rangle) &= \lambda f(|a\rangle). \end{aligned} \tag{4.63}$$

Note the difference between an isomorphism and a linear operator (Defn. A.15). A linear operator is a map within a single linear space, while an isomorphism may be between two different linear spaces. In addition, a linear operator does not have to be a one-to-one map.

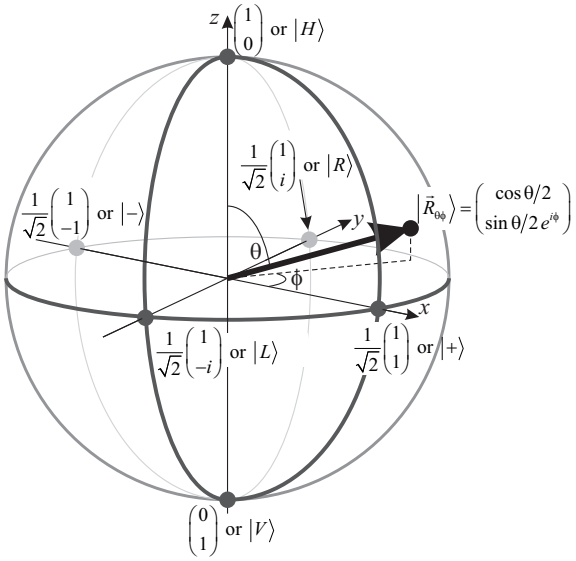


Fig. 4.5 The Bloch sphere.

of $\hat{S}_x$ will be mapped onto the diagonal polarization states $|+\rangle$ and $|-\rangle$, and the eigenstates of $\hat{S}_y$ onto the circular polarization states $|R\rangle$ and $|L\rangle$.

Accordingly, we can represent the polarization states by points on the Bloch sphere (Fig. 4.5). Note that the linear polarization states $|\alpha\rangle = \cos\alpha\,|H\rangle + \sin\alpha\,|V\rangle$ (where α is the polarization angle) can at the same time be written according to Eq. (4.62) as $|\alpha\rangle = \cos\frac{\theta}{2}\,|H\rangle + \sin\frac{\theta}{2}\,|V\rangle$ (where θ is the polar angle on the Bloch sphere). This means that this polar angle is *twice* the polarization angle. For example, as evidenced by Fig. 4.5, states $|H\rangle$ and $|V\rangle$ are separated on the Bloch sphere by a 180° arc, and states $|H\rangle$ and $|\pm\rangle$ by a 90°.

Note the different logic we used in dealing with the Pauli operators and their eigenvectors when studying photon polarization in Chapter 1 and spin in this chapter. In the former case, we first introduced the three polarization bases and then defined the Pauli operators as the observables associated with these bases in Ex. 1.29. Here, on the other hand, we first obtained the Pauli operators from the physics of angular momentum in Ex. 4.26, and then calculated their eigenstates.

Exercise 4.52. A horizontally polarized photon propagates through a

a) half-wave plate;
b)* quarter-wave plate

with its optic axis oriented at angle α with respect to the horizontal. Plot the locus of the resulting polarization states on the Bloch sphere for all possible values of α.
Hint: refer to Ex. 1.24. Part (b) can be solved numerically.

Exercise 4.53. A pair of electrons, shared between Alice and Bob, are prepared in an entangled spin state

$$\left|\Psi^{-}\right\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\right).$$

Alice measures the projection of her electron's spin onto vector $\vec{R}_{\theta,\phi}$ defined by spherical angles (θ,ϕ). Find the probability of each possible outcome of this measurement and the resulting state of Bob's electron. What is the location of that state, and Alice's corresponding measurement result, on the Bloch sphere?

4.6 Magnetic moment and magnetic field

4.6.1 Angular momentum and magnetic moment

Many elementary particles are charged, so their having an angular momentum implies that there is an electric charge moving in a loopwise fashion. This motion gives rise to a magnetic moment, which can interact with externally applied magnetic fields (Box 4.4). This interaction has a wide range of applications, from quantum information processing to biological imaging.

Exercise 4.54. For the classical motion of a point-like particle of mass M and charge e in a circular orbit with angular momentum $\vec{L}$, show that the gyromagnetic ratio[18] is given by

$$\gamma = \frac{e}{2M}. \tag{4.70}$$

Although we derived this result classically, it remains valid in the quantum domain, but the quantum gyromagnetic ratio includes a dimensionless factor called the *Landé factor*:

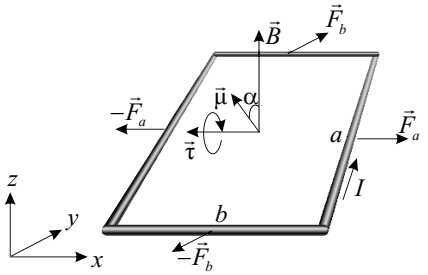
$$\gamma = g\frac{e}{2M}. \tag{4.71}$$

This factor depends on the nature of the motion. If the angular momentum is entirely due to the orbital motion, $g = 1$ (so the quantum expression is the same as the classical one). For the spin of the electron it is 2.0023, for the proton 5.5857.

The Landé factor for the spin can be derived theoretically using methods of relativistic quantum electrodynamics. A way to visualize this factor is to picture a spinning electron which is not exactly pointlike but has a finite size. The distributions of mass and charge over the electron's volume are different: while the mass is more concentrated in the particle's center, the charge is spread over the periphery. As a result, the ratio between the magnetic moment and the mechanical angular momentum is higher than that expected from a particle with identical distributions of mass and charge.

[18] See Box 4.4 for the definition of the gyromagnetic ratio.

Box 4.4 Magnetic moment in a magnetic field: a classical summary



Forces on an electric current loop placed in a magnetic field

Suppose a rectangular loop of dimensions $a \times b$ carrying current I is placed in a magnetic field $\vec{B}$ oriented along the z axis. The normal to the loop is at angle α to the z axis as shown in the figure. Each side of the loop will experience a magnetic force for which the general expression is $\vec{F} = I\vec{l} \times \vec{B}$, where $\vec{l}$ is the length vector of that side. The forces acting on the sides of length a will balance each other, but the forces acting on the sides of length b (whose magnitudes are $F_b = IbB$) will produce a torque of magnitude $\tau = 2F_b \times (a/2) \sin\alpha = IBab \sin\alpha = IBA \sin\alpha$, where A is the area of the loop.

The magnetic moment $\vec{\mu}$ carried by the loop is a vector of magnitude

$$\mu = Iab = IA \tag{4.64}$$

and the direction is perpendicular to the plane of the loop. The torque acting on the loop is therefore

$$\vec{\tau} = \vec{\mu} \times \vec{B}. \tag{4.65}$$

In this form, the relation is quite general and valid for loops of any shape.

Being acted upon by the magnetic forces, each of the wires carries potential energy. Let us calculate the full potential energy of the loop as a function of the angle α, assuming that the loop can rotate around the axis that coincides with one of the sides of length b, and that $\alpha = \pi/2$ corresponds to the zero energy position. Rotating the loop from that position to another α means displacing the other side of length b by the distance $\pm a\cos\alpha$ in the y direction, thereby performing the work $W = -F_b a\cos\alpha = -IBab\cos\alpha = -\mu B\cos\alpha$. Hence the potential energy is given by

$$U = -\vec{\mu} \cdot \vec{B}. \tag{4.66}$$

Once again, the latter expression does not depend on the shape of the loop or the location of the axis. The potential energy of a magnetic dipole in a magnetic field is minimized when the two are collinear.

In addition to the current, the charged particles moving in the loop carry mass, so their motion is associated with angular momentum $\vec{L}$. The magnetic moment is proportional to the angular momentum,

$$\vec{\mu} = \gamma\vec{L}, \tag{4.67}$$

where the proportionality coefficient is the *gyromagnetic ratio* (see also Ex. 4.54).

The effect of the torque on that angular momentum is $\dot{\vec{L}} = \vec{\tau}$. Using Eqs. (4.65) and (4.67) we obtain

$$\dot{\vec{L}} = \gamma\vec{L} \times \vec{B}. \tag{4.68}$$

As we know from classical mechanics, the solution of the differential equation (4.68) is a *precession* of the loop around the direction of the magnetic field with angular frequency

$$\Omega_L = \gamma B, \tag{4.69}$$

known as the *Larmor frequency*.

Exercise 4.55. For a charged particle with an orbital or spin angular momentum, show that

a) the projection of the magnetic moment onto the z axis quantizes according to

$$\mu_z = \hbar\gamma m; \tag{4.72}$$

b) under the action of a constant magnetic field B, the energy eigenvalues are

$$E_m = -\hbar\Omega_L = -\hbar\gamma B m, \tag{4.73}$$

where m is the corresponding magnetic or spin quantum number and Ω_L is the Larmor frequency (4.69).

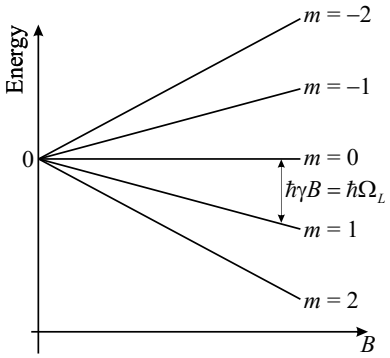


Fig. 4.6 Zeeman splitting of the energy in a magnetic field. The example in the figure has $L = 2$. The electric charge of the spinning particle, and hence the gyromagnetic ratio γ, are assumed positive.

The energy level splitting in the magnetic field that we found in part (b) is called the *Zeeman effect* (Fig. 4.6). It is ubiquitous in atomic and nuclear physics.

If the angular momentum in the above exercise is orbital, then we can see from Eq. (4.70) that the quantum of the projection of the magnetic moment onto the z axis is

$$\mu_B = \frac{e}{2M}\hbar. \tag{4.74}$$

For the electron ($M = M_e$), this quantity is called the *Bohr magneton*. It is equal to 5.8×10^{-9} eV/Gauss $= 9.3 \times 10^{-24}$ J/T.

Exercise 4.56.§ Verify that the data in the last column of Table 4.3 are consistent with those in the other columns.

Table 4.3 Magnetic dipole properties of some elementary particles.

Particle	Mass, kg	Charge, C	Spin	Landé factor	Larmor frequency, MHz/T
Electron	9.10938×10^{-31}	1.60218×10^{-19}	1/2	2.0023	28025
Proton	1.67262×10^{-27}			5.5857	42.5781
muon	1.883532×10^{-28}			2.0023	135.539

4.6.2 Stern–Gerlach apparatus

A particle with a magnetic moment placed in an external magnetic field has a potential energy given by Eq. (4.66). If the magnetic field varies as a function of position, this potential energy has a gradient, which manifests itself as a force: $\vec{F} = -\vec{\nabla}U$. Using Eq. (4.66), we can rewrite this as $\vec{F} = \vec{\nabla}(\hat{\vec{\mu}} \cdot \vec{B})$. If we define the z axis to be along the magnetic field, this result simplifies to

$$\vec{F} = (\vec{\nabla}B)\hat{\mu}_z. \tag{4.75}$$

The magnitude of the force is proportional to the projection of its magnetic moment onto the field direction.

This observation can be used to measure components of the quantum angular momentum vector. The *Stern–Gerlach apparatus*[19] contains a permanent magnet of such a shape that the field it produces is significantly non-uniform. When a particle is traveling through this field, it experiences a force and deviates from its original path. The path taken by the particle is revealed by a sensitive screen placed behind the magnet (Fig. 4.7).

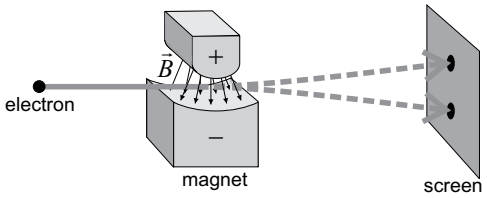


Fig. 4.7 The Stern–Gerlach apparatus

Because the magnetic moment is proportional to the angular momentum, the Stern–Gerlach apparatus effectively measures the component of the angular mo-

[19] W. Gerlach and O. Stern, *Der experimentelle Nachweis der Richtungsquantelung im Magnetfeld*, Zeitschrift für Physik **9**, 349–352 (1922); W. Gerlach and O. Stern, *Das magnetische Moment des Silberatoms*, Zeitschrift für Physik **9**, 353–355 (1922); W. Gerlach and O. Stern, *Der experimentelle Nachweis des magnetischen Moments des Silberatoms*, Zeitschrift für Physik **8**, 110–111 (1922).

mentum along the direction of the field. Because the values of this component are quantized, the particle will land at discrete spots on the target screen. For example, a free electron can end up at two spots corresponding to $m_s = \pm\frac{1}{2}$. In the context of the spin-polarization isomorphism (Sec. 4.5), the Stern–Gerlach measurement of the z projection of the electron spin is equivalent to measuring the photon's polarization in the canonical basis using a polarizing beam splitter (Sec. 1.4).

Exercise 4.57. An electron prepared in an eigenstate of the spin component along the vector with polar coordinates (θ, ϕ) and with the eigenvalue $\frac{\hbar}{2}$ passes through a Stern–Gerlach apparatus with the field vector along the z-axis. What are the probabilities that the electron will end up on each of the two spots on the screen?

Exercise 4.58. In a Stern–Gerlach apparatus, the directions of the field and its gradient may be different. Which of these two directions determines the measurement basis?

Exercise 4.59. A beam of particles with the spin $s = 1$ in the eigenstate of $\hat{s}_x$ with zero eigenvalue passes through a Stern–Gerlach apparatus with the field vector along the y-axis. How many spots will be formed on the target screen and in which proportion will the particles divide among these spots?

Exercise 4.60. A beam of electrons prepared with the spins pointing in the negative z direction passes through a Stern–Gerlach apparatus with the field vector oriented in the x-z plane, at angle θ_0 to the z-axis. What is the splitting proportion?

4.6.3 Evolution of magnetic states

From classical physics (Box 4.4), we know that a magnetic moment placed in a magnetic field will precess around that field. Do we also expect an effect of this kind in the quantum domain? To answer this question, we need to study the evolution of our quantum system under the Hamiltonian (4.66). Taking Eq. (4.71) into account, we rewrite this Hamiltonian as

$$\hat{H} = -\hat{\vec{\mu}} \cdot \vec{B} = -\gamma \hat{\vec{L}} \cdot \vec{B}. \tag{4.76}$$

Note that we treat the macroscopic magnetic field as a classical vector rather than an operator.

Exercise 4.61. By writing the differential equation for the evolution of the components of the angular momentum vector observable in the Heisenberg picture, reproduce the classical result (4.68).

We see that the Heisenberg picture behavior of the quantum magnetic moment in a field is similar to the classical one: it precesses around the field with Larmor frequency $\Omega_L = \gamma B$ (Fig. 4.8). Of course, if we are interested in the *mean* values of the angular momentum vector operator, this result applies independently of whether we use the Heisenberg or Schrödinger picture for the calculation. In the case of the

two-level system, for example, the Bloch vector [whose components $R_{x,y,z} = \langle \sigma_{x,y,z} \rangle$ as per Ex. 4.48(c)] evolves according to

$$\dot{\vec{R}} = \gamma \vec{R} \times \vec{B}. \tag{4.77}$$

This important result illustrates the utility of the Heisenberg picture: obtaining it in the Schrödinger picture would be much more difficult. We do this in the next exercise for a few special cases.

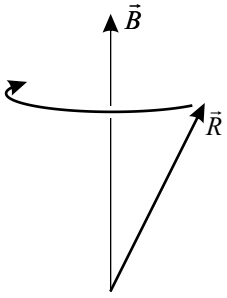


Fig. 4.8 Precession of the Bloch vector around the magnetic field. The gyromagnetic ratio is assumed positive.

Exercise 4.62. Find the evolution in the Schrödinger picture of the spin state of a free electron under the action of a constant magnetic field $\vec{B}$ given the following conditions:

a) the initial state is represented by an arbitrary point (θ_0, ϕ_0) on the Bloch sphere and the magnetic field is along the z axis;
b) the initial state corresponds to the spin pointing along the z axis and the magnetic field is along the y axis;
c) the initial state corresponds to the spin pointing along the z axis and the magnetic field is along a vector with polar angles $(\theta_0, 0)$.

Present your solutions in the matrix form in the canonical basis and as trajectories on the Bloch sphere. Check that your result is consistent with Eq. (4.77). For each answer, find the splitting ratio that would be observed in a Stern–Gerlach measurement with the magnetic field in the z direction.
Hint: See Ex. 1.47 for a related problem.

Exercise 4.63. A photon and an electron are prepared in the entangled state

$$\left|\Psi^-\right\rangle = \frac{1}{\sqrt{2}}(|H\downarrow\rangle - |V\uparrow\rangle) \tag{4.78}$$

and distributed between Alice and Bob, who use them to perform quantum teleportation of another photon in state $|\chi\rangle = \alpha|H\rangle + \beta|V\rangle$ onto the spin of Bob's electron.

To this end, Alice performs a Bell measurement on her two photons. For each possible output of this measurement, find the direction and magnitude of the magnetic field $\vec{B}$ that Bob would need to apply to his electron for a given time τ in order to bring its spin to the state $\alpha|\uparrow\rangle + \beta|\downarrow\rangle$.

4.7 Magnetic resonance

4.7.1 Rotating basis

Suppose a spin-$\frac{1}{2}$ particle is placed in a constant magnetic field B_0 directed along the z axis. As discussed previously [Ex. 4.55(b)], the states $|\uparrow\rangle$ and $|\downarrow\rangle$ are the eigenstates of the Hamiltonian with energies $E_{\uparrow,\downarrow} = \mp\frac{\hbar}{2}\Omega_0$, where $\Omega_0 = \gamma B_0$ is the Larmor frequency[20], with γ the particle's gyromagnetic ratio. Our goal in this section is to study the phenomena that occur if, in addition, a relatively weak magnetic field, *oscillating* at a frequency ω close to Ω_0[21], is applied along the x axis:

$$\vec{B} = B_0\hat{k} + B_{\rm rf}\cos\omega t\,\hat{i}. \tag{4.79}$$

In other words, we would like to know what happens if this oscillating field is close to *resonance* with the *two-level system* formed by states $|\uparrow\rangle$ and $|\downarrow\rangle$ (Fig. 4.9).

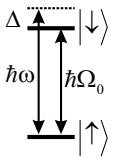


Fig. 4.9 Magnetic resonance in a two-level system.

Exercise 4.64. Write down the Hamiltonian and the differential equations for the Schödinger evolution of the particle's spin state $|\psi(t)\rangle$ in the $(|\uparrow\rangle, |\downarrow\rangle)$ basis.
Answer:

$$H = -\frac{\hbar}{2}\gamma(B_0\hat{\sigma}_z + B_{\rm rf}\hat{\sigma}_x\cos\omega t) \simeq \frac{\hbar}{2}\begin{pmatrix} -\Omega_0 & -\gamma B_{\rm rf}\cos\omega t \\ -\gamma B_{\rm rf}\cos\omega t & \Omega_0 \end{pmatrix}; \tag{4.80}$$

[20] We will be using the symbol Ω_0 rather than Ω_L for the Larmor frequency in this section.

[21] This magnetic field is usually referred to as the *radio frequency (rf)* field, because ω is typically in the range where radio and TV transmissions occur. The field B_0 is referred to as the *dc (direct current)* field.

$$\dot{\psi}_\uparrow = \frac{\mathrm{i}}{2}\Omega_0\psi_\uparrow + \frac{\mathrm{i}\gamma}{2}B_{\mathrm{rf}}\cos\omega t\,\psi_\downarrow; \tag{4.81a}$$

$$\dot{\psi}_\downarrow = -\frac{\mathrm{i}}{2}\Omega_0\psi_\downarrow + \frac{\mathrm{i}\gamma}{2}B_{\mathrm{rf}}\cos\omega t\,\psi_\uparrow, \tag{4.81b}$$

where $\psi_\uparrow(t) = \langle\uparrow \mid \psi(t)\rangle$, $\psi_\downarrow(t) = \langle\downarrow \mid \psi(t)\rangle$.

Equations (4.81) are similar to the ones we deal with when we study quantum evolution in any two-dimensional Hilbert space (see, for example, Ex. 1.47). However, the coefficients on the right-hand side are now time-dependent. This complicates the calculation quite a bit. However, for $B_{\mathrm{rf}} \ll B_0$ and near the resonance, an elegant approximate solution exists. As the first step in finding it, let us define a new, time-dependent, basis in our Hilbert space:

$$\left|\tilde{\uparrow}\right\rangle = |\uparrow\rangle\,\mathrm{e}^{\frac{\mathrm{i}}{2}\omega t}; \tag{4.82a}$$

$$\left|\tilde{\downarrow}\right\rangle = |\downarrow\rangle\,\mathrm{e}^{-\frac{\mathrm{i}}{2}\omega t}. \tag{4.82b}$$

For a reason that will become evident in the next exercise, the new basis is called the *rotating basis*. We denote the decomposition coefficients of the state $|\psi\rangle$ in the rotating basis by

$$\tilde{\psi}_\uparrow(t) = \left\langle\tilde{\uparrow}\right|\,\psi(t)\rangle = \psi_\uparrow(t)\mathrm{e}^{-\frac{\mathrm{i}}{2}\omega t}; \tag{4.83a}$$

$$\tilde{\psi}_\downarrow(t) = \left\langle\tilde{\downarrow}\right|\,\psi(t)\rangle = \psi_\downarrow(t)\mathrm{e}^{\frac{\mathrm{i}}{2}\omega t}. \tag{4.83b}$$

We will be referring to the original canonical basis $\{|\uparrow\rangle, |\downarrow\rangle\}$ as *stationary*.

Exercise 4.65. Show that the Bloch vectors in the stationary and rotating bases[22] are related by rotation through an angle ωt around the z axis.

We know that, in the absence of the rf field, the Bloch vector in the stationary basis precesses around the magnetic field with the Larmor frequency Ω_0. In the rotating basis, the Bloch vector precesses much more slowly, with angular velocity $\Omega_0 - \omega$.

Exercise 4.66. Show that Eqs. 4.81, rewritten in terms of $\dot{\tilde{\psi}}_\uparrow$ and $\dot{\tilde{\psi}}_\downarrow$, take the form

$$\dot{\tilde{\psi}}_\uparrow = -\frac{\mathrm{i}}{2}\Delta\,\tilde{\psi}_\uparrow + \frac{\mathrm{i}\gamma}{4}B_{\mathrm{rf}}(1+\mathrm{e}^{-2\mathrm{i}\omega t})\,\tilde{\psi}_\downarrow; \tag{4.84a}$$

$$\dot{\tilde{\psi}}_\downarrow = \frac{\mathrm{i}}{2}\Delta\,\tilde{\psi}_\downarrow + \frac{\mathrm{i}\gamma}{4}B_{\mathrm{rf}}(1+\mathrm{e}^{2\mathrm{i}\omega t})\,\tilde{\psi}_\uparrow, \tag{4.84b}$$

where $\Delta = \omega - \Omega_0$ is the *detuning* of the rf field from the resonance.

Up to now, our calculations have been precise. Next, we perform an important trick known as the *rotating-wave approximation*. We will neglect the quickly oscillating terms involving $\mathrm{e}^{\pm 2\mathrm{i}\omega t}$ on the right-hand side of Eqs. (4.84). The argument for

[22] The Bloch vector in the new basis is obtained by substituting $(\tilde{\psi}_\uparrow, \tilde{\psi}_\downarrow)$ into Eq. (4.62) instead of $(\psi_\uparrow, \psi_\downarrow)$.

Box 4.5 Unphysical nature of the rotating-wave Hamiltonian

Given that the stationary and rotating bases are related to each other by a complex phase shift (4.82), we would expect, for example, that $\langle \uparrow | \hat{H} | \uparrow \rangle = \langle \tilde{\uparrow} | \hat{H} | \tilde{\uparrow} \rangle$. But this equality seems inconsistent with the matrices of the stationary and rotating-wave Hamiltonians given by Eqs. (4.80) and (4.85), respectively: $H_{\uparrow\uparrow} = -\hbar\Omega_0/2$, while $(H_{\text{RWA}})_{\tilde{\uparrow}\tilde{\uparrow}} = \hbar\Delta/2$. But where does this discrepancy come from? The rotating-wave approximation cannot be the answer, because it affects only the off-diagonal elements of the Hamiltonian, not the diagonal ones.

The reason is that we can express the Schrödinger equation in a matrix form, such as Eq. (1.32), only for a time-independent basis. Only then can we write, for example, that $\langle\uparrow| \left(\frac{\mathrm{d}}{\mathrm{d}t}|\psi\rangle\right) = \frac{\mathrm{d}}{\mathrm{d}t}\langle\uparrow|\psi\rangle$. If the basis is time-dependent, we would also need to take into account the time derivative of the basis element, so the above equality would not be valid. But in deriving the rotating-wave Hamiltonian matrix (4.85) from the evolution (4.84), we neglected this, handling the rotating basis as if it were time-independent.

As a result, the rotating-wave Hamiltonian is unphysical, or *fictitious*: it does not represent the actual energy observable*. In particular, the element $(H_{\text{RWA}})_{\tilde{\uparrow}\tilde{\uparrow}}$ of its matrix is not equal to the expectation value $\langle \tilde{\uparrow} | \hat{H} | \tilde{\uparrow} \rangle$ of the full Hamiltonian $\hat{H}$. Nevertheless, H_{RWA} gives a correct mathematical description (4.84) of the spin state evolution. If our goal is to find that evolution, we may as well not worry about the physics of the rotating-wave Hamiltonian, and just use it as a formal tool for a theoretical treatment.

* In fact, Eq. (4.85) does correctly represent the Hamiltonian of the system in the so-called *interaction picture*, which we do not study here.

doing so is that these terms average out during an oscillation period $2\pi/\omega$, so their effect becomes negligible in comparison to the other terms, which do not oscillate. This approximation is valid as long as $\Delta \ll \omega, \Omega_0$ and $B_{\text{rf}} \ll B_0$.

Exercise 4.67. Show that, under the rotating-wave approximation, the evolution defined by Eqs. (4.84) is the same as under the Hamiltonian

$$\hat{H}_{\text{RWA}} \simeq \frac{\hbar}{2}\begin{pmatrix} \Delta & -\Omega \\ -\Omega & -\Delta \end{pmatrix}, \tag{4.85}$$

where $\Omega = \gamma B_{\text{rf}}/2$ is called the *Rabi frequency*.

We see that, in the rotating basis and under the rotating-wave approximation, the evolution due to a *time-varying* field is described by a *constant* Hamiltonian, and this greatly simplifies the calculations. In addition, this Hamiltonian provides us with a way to visualize things, which we shall examine next.

Exercise 4.68. Show that a Hamiltonian with a matrix identical to (4.85) is obtained if the spin is placed in a constant magnetic field $\vec{B}$ of magnitude

$$B = \frac{\sqrt{\Omega^2 + \Delta^2}}{\gamma}, \tag{4.86}$$

with components

Box 4.6 Rabi oscillations and photoelectric effect

The photoelectric effect is the emission of free electrons from a surface illuminated by light. Its salient properties, as determined experimentally, are as follows.

- The kinetic energy of the ejected electrons depends on the wavelength of the light but not its intensity.
- Electrons are emitted only if the wavelength is below a certain threshold value.

These properties, inconsistent with classical physics, were explained by Einstein in 1905 using the notion of the photon. According to this explanation, the energy $\hbar\omega$ of a photon absorbed by the surface is partially used up to overcome the potential U that binds the electron to the host substance; the remainder ($K = \hbar\omega - U$) becomes the photoelectron's kinetic energy. Therefore only light with $\hbar\omega \geq U$ can produce photoelectrons.

The intuitive nature of Einstein's explanation and its excellent agreement with the experimental data were instrumental in bringing about universal acceptance of the quantum theory by the physics community. It was primarily for this discovery that Einstein was awarded his Nobel prize in 1921.

The quantum physics of two-level systems studied here allows for an alternative explanation of the photoelectric effect. Transitions between energy levels in matter due to the action of resonant electromagnetic fields are governed by the same laws as in magnetic resonance. When a (classical) wave of frequency ω is in resonance with the transition between a bound state of energy $-U$ and a continuous-spectrum free electron state of energy K, a Rabi oscillation between these states ensues. Once the electron is in a superposition of the bound and unbound states, it may be observed in the unbound state and collapse onto that state, thereby manifesting the photoelectric effect.

In summary, it is not necessary to invoke photons to explain the photoelectric effect. It suffices to treat the host matter quantum-mechanically and the electromagnetic wave classically.

$$B_x = \frac{\Omega}{\gamma},\ B_y = 0,\ B_z = -\frac{\Delta}{\gamma}. \tag{4.87}$$

We find that the rotating-wave Hamiltonian can be interpreted to arise due to a constant magnetic field oriented at a certain angle. Of course, being a consequence of a fictitious Hamiltonian (Box 4.5), this field itself is unphysical; it has nothing to do with the actual field (4.79). But it is nevertheless supremely convenient, because it allows direct application of the results for the quantum evolution of a spin in a constant magnetic field, as obtained in the previous section, to the magnetic resonance problem.

4.7.2 Evolution under the rotating-wave approximation

As we found in Ex. 4.61, the behavior of the Bloch vector in a magnetic field is identical to the classical behavior. This means that the evolution of the Bloch vector in the rotating basis under a Hamiltonian (4.85) consists in precession around the fictitious field (4.87), as shown in Fig. 4.10(a).

In the case of exact resonance, $\Delta = 0$, the fictitious field $\vec{B}$ is of magnitude Ω/γ and directed along the x axis, so the trajectory of the Bloch vector is a meridian crossing the y axis. The precession occurs with angular velocity $\gamma B = \Omega$. Accordingly, the populations[23] of the spin-up and spin-down states will oscillate sinusoidally with the Rabi frequency. This phenomenon is known as *Rabi oscillations*.

Detuning the rf field from the resonance (so that $\Delta \neq 0$) has two effects [Fig. 4.10(a,b)]. First, the frequency of the Rabi oscillations will increase because of the term Δ^2 in the fictitious field magnitude (4.86). Second, the direction of that field is no longer horizontal. If a trajectory starts in the spin-up state, it will no longer reach the south pole of the Bloch sphere, so we will never observe the spin-down state with probability one.

Exercise 4.69. Find the highest probability $\mathrm{pr}_{\downarrow\max}$ of observing the spin-down state during the Rabi cycle as a function of the detuning Δ. The cycle starts in the spin-up state.

Hint: Although the problem can be solved by calculating the Schrödinger evolution under the Hamiltonian (4.85) (and we shall do it in the next exercise), the question is much easier to answer by just looking at the geometry of the Bloch sphere.

Answer:

$$\mathrm{pr}_{\downarrow\max} = \frac{\Omega^2}{\Omega^2 + \Delta^2}. \tag{4.88}$$

We can now see why this phenomenon is called "resonance". The Lorentzian shape of the curve (4.88) [Fig. 4.10(c)] is quite similar to the response of a mechanical harmonic oscillator or an LC circuit to a periodic driving force. Note an important difference, though: in the case of a harmonic oscillator, the resonance width is determined by the damping constant, but is independent of the driving field. The width of the magnetic resonance, in contrast, is proportional to the Rabi frequency, i.e., to the amplitude of the rf field. This phenomenon is called *power broadening* and is characteristic of two-level systems.

The two-level system has limited energy: its highest energy eigenstate is that of the spin-down state. However high the applied rf power is, it cannot further increase the system energy; the system *saturates*. The harmonic oscillator, on the other hand, has infinitely many energy levels and therefore does not saturate: when we drive it more strongly, it will respond by going into increasingly higher energy states. Accordingly, it will not exhibit any power broadening[24].

Exercise 4.70. Find the evolution of the spin state $|\psi(t)\rangle$ under the Hamiltonian (4.85), starting with the initial state $|\psi(0)\rangle = |\uparrow\rangle$. Find the probabilities of the spin-up and spin-down states as a function of time, Rabi frequency, and detuning. Reconcile the result with that of Ex. 4.69 and Fig. 4.10(b).

Hint: use Ex. 4.62(c).

---

[23] The *population* of a quantum state is the total number of particles in that state. In our case, the spin-up and spin-down populations are, respectively, $n\,\mathrm{pr}_{\uparrow}$ and $n\,\mathrm{pr}_{\downarrow}$, where n is the total number of electrons in the sample.

[24] See a related discussion in Sec. 3.8.2.

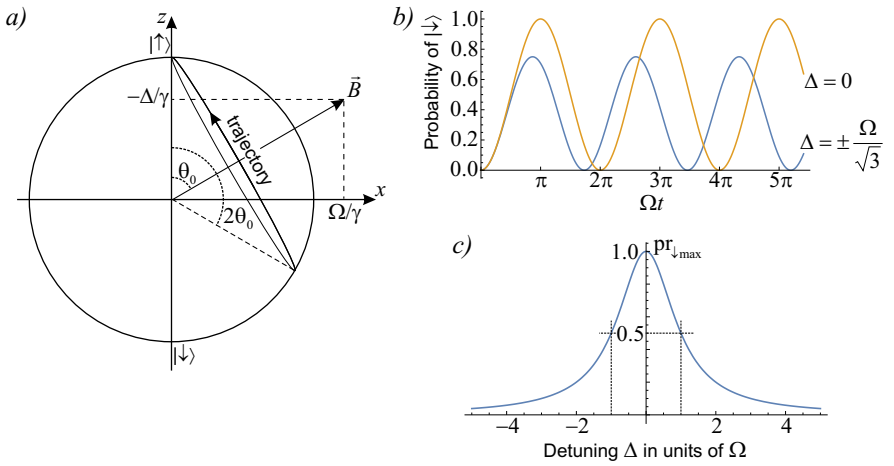


Fig. 4.10 Evolution in the rotating-wave Hamiltonian (4.85), with the particle initially in the spin-up state. a) Trajectory on the Bloch sphere, specialized to $\Delta = -\Omega/\sqrt{3}$ and viewed in the x-z plane. b) Probability of detecting the spin-down state as a function of time. c) Resonance curve (4.88). Dashed lines show its width at half-maximum.

Exercise 4.71.* Find the rotating-wave Hamiltonian for the setting in which the rf field is given by $B_{\rm rf}\cos(\omega t+\beta)$, where β is an arbitrary phase, and is directed

a) along the x axis;
b) along the y axis.

Find the coordinates of the corresponding fictitious magnetic field vector. Show that, if the rf frequency is resonant with the two-level transition, this field is always horizontal.

Answer: The fictitious magnetic field is

a) $B = (\frac{\Omega}{\gamma}\cos\beta, -\frac{\Omega}{\gamma}\sin\beta, -\frac{\Delta}{\gamma})$;
b) $B = (\frac{\Omega}{\gamma}\sin\beta, \frac{\Omega}{\gamma}\cos\beta, -\frac{\Delta}{\gamma})$.

Exercise 4.72. Write the Schrödinger equation in the stationary basis for the rf field directed along the z axis. Show that no transitions between the spin-up and spin-down states will occur in this case.

4.7.3 Pulse area

We have seen in the previous section that a resonant rf field whose Rabi frequency is Ω, acting for time t, will rotate the Bloch vector through the angle Ωt. In many practical applications, the resonant rf field is applied in a pulsed manner, so that its amplitude and hence the Rabi frequency depends on time: $\Omega = \Omega(t)$. Such a pulse

will rotate the Bloch vector through the angle $\int \Omega(t)\mathrm{d}t$. This quantity is known as the *pulse area*[25]. The notion of the pulse area is convenient because it is a single parameter that fully describes the effect of the pulse on the spin; there is no need to know the exact shape of the pulse as long as we know its integral.

For example, applying a pulse of area $\pi/2$ to the spin-up state will transform it into the state with the spin pointing along the y axis, $\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ \mathrm{i}\end{pmatrix}$. If we apply another $\pi/2$ pulse to that state, we will obtain the spin-down state. Together, these two pulses comprise a pulse of area π, whose effect is to "flip" the Bloch vector around the x axis[26].

If the rf field is applied in a pulsed manner (as is typically the case in magnetic resonance applications, see Box 4.7), achieving a macroscopic pulse area requires a relatively high Rabi frequency. Then we need not worry about precise tuning of the rf field as long as $\Omega \gg \Delta$ for most of the pulse duration (but we must still observe $\Omega \ll \Omega_0$). If this is the case, the fictitious magnetic field (4.87) is almost horizontal, and the effect of detuning negligible.

Exercise 4.73. A particle is initially in the spin-up state. It is subjected to a $\pi/2$ pulse, followed by another $\pi/2$ pulse, in which the phase of the rf field is shifted through an angle β. Find the final population of the spin-down state as a function of β. Interpret your results for $\beta = 0$ and $\beta = \pi$.

4.7.4 Applications of magnetic resonance

Suppose we have a large set (*ensemble*) of spin-$\frac{1}{2}$ particles, initially prepared in the state $|\uparrow\rangle$, along the dc magnetic field. If we apply a pulse of area $\pi/2$ to that ensemble, the spins will rotate to a horizontal position. After the rf pulse, if the dc field is still present, the spins will precess around the z axis at the frequency Ω_0.

Exercise 4.74. A short pulse of area $\pi/2$ is applied to a particle initially in the spin-up state, ending at $t = 0$. Calculate the mean values of the three Cartesian components of the magnetic moment observable at $t > 0$

a) in the rotating basis;
b) in the stationary basis.

A precessing magnetic moment will emit an electromagnetic field at the frequency of the precession. This field, which is proportional to the horizontal component of the Bloch vector, can be detected by a radio receiver and provides us with important insights into the substance that contain the spins. Here we discuss the properties of this emission.

---

[25] In reference to the integral being the "area underneath the curve".

[26] The π pulse corresponds to a logical "not" operation on a spin-based qubit: it transforms $|0\rangle = |\uparrow\rangle$ into $|1\rangle = |\downarrow\rangle$ and vice versa.

The signal obtained in response to a single pulse is known as the *free induction decay*. This name refers to the fact that the signal quickly loses its strength with time as a result of various damping and decoherence mechanisms. The primary mechanism causing the damping is the slight variation (inhomogeneity) in the dc magnetic field with the spatial position. This results in *inhomogeneous broadening* of the resonance: each spin will have a different Δ within a certain range Δ_0 known as the *inhomogeneous width*. The Bloch vectors with different detunings will precess around the z axis with different angular velocities, and spread across the equator of the Bloch sphere on a time scale of Δ_0^{-1} [Fig. 4.11(a), panel 2]. Then the fields emitted by different spins will acquire different phases and cancel each other.

Exercise 4.75. An ensemble of spins is inhomogeneously broadened in such a way that its detunings are distributed as follows:

$$p(\Delta) = \frac{1}{\sqrt{\pi}\Delta_0} \mathrm{e}^{-(\Delta/\Delta_0)^2}.$$

Under the conditions of Ex. 4.74, calculate the average magnetic moment vector of a spin in this ensemble as a function of time $t > 0$ in the rotating basis.
Hint: Use Ex. D.9(c).
Answer:

$$\overline{\langle \vec{\mu} \rangle} = \left(0, \frac{\hbar\gamma}{2} \mathrm{e}^{-\frac{(\Delta_0 t)^2}{4}}, 0\right). \tag{4.89}$$

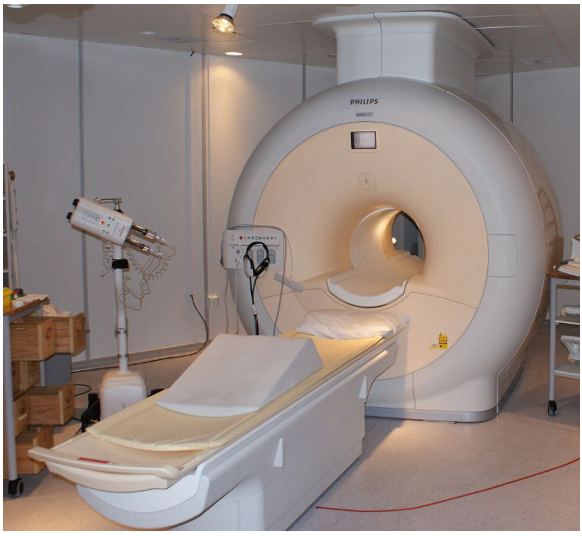
The horizontal bar over $\langle \vec{\mu} \rangle$ means that quantum averaging is followed by statistical averaging over the ensemble. Note also that the mean direction of the spin in the rotating basis is constant along the y axis; in the stationary basis, this corresponds to precession at the frequency ω, as per Ex. 4.65.

Inhomogeneous broadening is frequently the dominant limitation on the free induction decay lifetime. As such, it prevents one from measuring the time constants associated with other mechanisms that degrade the spin state — decoherence and thermalization — collectively known as the *homogeneous dephasing (relaxation)*. But for applications such as imaging (Box 4.7), it is these latter time constants that are of interest, as they are characteristic of the sample substance.

Fortunately, the effect of inhomogeneous broadening can be reversed by means of an elegant technique known as the *spin echo*. After the free induction decay has ended, one applies an additional π pulse to flip all the Bloch vectors around the x axis. As is evident from Fig. 4.11, this will invert the positions of the Bloch vectors with respect to the ensemble average. As a result, the spreading will reverse, even though each individual spin will continue to evolve at the same pace as before. The spins will reunite to point in the same direction at time $t = 2t_0$, generating a strong electromagnetic field — the echo pulse.

Exercise 4.76. Under the conditions of Ex. 4.75, after time $t_0 \gg 1/\Delta_0$, the ensemble is subjected to a very short π pulse. Calculate the mean magnetic moment in the rotating basis as a function of time $t > t_0$. Neglect relaxation.
Answer:

Box 4.7 Magnetic resonance imaging



A medical magnetic resonance scanner. The toroidal structure is a superconducting solenoid that produces the dc field. Source: *Wikipedia*.

Magnetic resonance imaging relies on detecting the spin echo signal from protons (hydrogen nuclei) that are contained in water molecules within the patient's body. This signal is analyzed to determine the characteristic relaxation time of these spins, which is then mapped relative to the source location to form a 3D image. Because the dephasing time depends on the substance embedding the emitter spins, this 3D image reflects the organ and tissue structure, as well as their pathologies. For example, gray matter and white matter in the human brain differ in the dephasing time by about 30%.

In order to implement this kind of imaging, we must know which point each echo signal is coming from. This is achieved by applying the dc field with a gradient, so that the resonance frequency is position-dependent. In this way, only the protons located within a narrow slice of the patient's body will respond to the rf field of a specific frequency. Three-dimensional imaging utilizes complex sequences of pulses, for each of which the dc field has a gradient in a different direction. This results in a spin echo signal with a temporal pattern that carries information about the location of its source.

One of the many issues arising in magnetic-resonance imaging is that it is difficult to prepare all spins in the same initial state. Before the rf pulses are applied, the proton spins are in a thermal equilibrium with the environment, which means that there is only a small difference between the densities of the spin-up and spin-down photons (Ex. 5.54). During the evolution, the Bloch vectors of these groups of protons will be oriented oppositely, and the signals they emit will largely cancel each other. This is in contrast to atomic physics (Ex. 4.47), in which the energy difference between levels is much higher, and so is the population contrast.

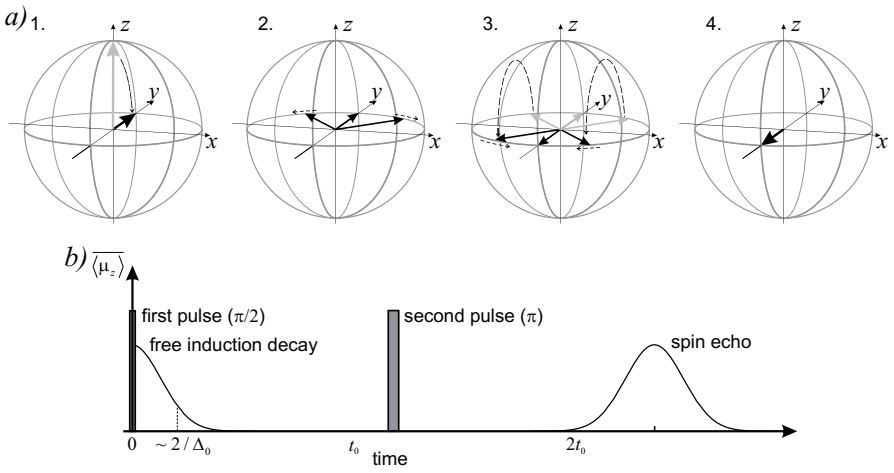


Fig. 4.11 Free induction decay and spin echo. a) Bloch sphere visualization. 1. The $\pi/2$ pulse at $t = 0$ rotates the Bloch vectors of all spins to point in the positive y direction. 2. Spins with different detunings spread across the equator of the Bloch sphere. 3. The π pulse at t_0 flips all Bloch vectors around the x axis. 4. At time $2t_0$, the Bloch vectors come back together. b) Field emitted by the ensemble (proportional to the mean y component of the spin) as a function of time.

$$\overline{\langle \vec{\mu} \rangle} = \left(0, -\frac{\hbar\gamma}{2} \mathrm{e}^{-\frac{[\Delta_0(t-2t_0)]^2}{4}}, 0\right). \tag{4.90}$$

Relaxation phenomena, neglected in the above calculation, lead to a degradation of the echo signal with t_0. By measuring the effect of changing t_0 on the echo strength, one can measure the characteristic relaxation time.

Another major application of magnetic resonance is in time metrology. Suppose we need to know precisely whether our rf field is resonant with the spin transition. This goal can be achieved using the technique known as *Ramsey spectroscopy*.

Exercise 4.77. Consider the following procedure performed on a spin initially in the $|\uparrow\rangle$ state.

1. A short pulse of area $\pi/2$ is applied. The Rabi frequency is chosen such that $\Omega \gg \Delta$, so that we can neglect detuning during the pulse and assume the pulse area to be precisely $\pi/2$.
2. The rf field is turned off for time t, so the atom evolves freely.
3. Another pulse of area $\pi/2$ is applied.
4. The populations of states $|\uparrow\rangle$ and $|\downarrow\rangle$ are measured.

Show that the final probability of detecting the particle in the state $|\downarrow\rangle$ behaves as $|\psi_\downarrow|^2 = \cos^2 \Delta t/2$. Solve the problem in the rotating basis using two different methods:

1. using geometry to trace the Bloch vector behavior as a function of time;

2. calculating the evolution operator matrices associated with the two pulses and the free evolution period.

The advantage of the Ramsey method is that the two-level system can be left alone during the free evolution time. This allows one to minimize its disturbance during that time, thereby optimizing its precision as a frequency standard (Box 4.8).

The phenomenon of Ramsey fringes may appear paradoxical. The process that leads to the dependence of the final population on Δt is the free evolution of the atom, while the rf field is off. How can the detuning of a field that is turned off have any effect on an experimentally measurable quantity?

The answer is that the detuning of the rf field determines the phase difference of the two $\pi/2$ pulses with respect to each other. As we found in Ex. 4.73, this difference has a crucial effect on the final population of the energy levels. When solving Ex. 4.77, we used the same operator for the two pulses, thereby quietly assuming that their phases are consistent with Eq. (4.79). In other words, the two phases are tied to a single "clock" $\cos\omega t$, which keeps running throughout the experiment. The detuning of the clock frequency will affect the relative phase of the two pulses, and hence the populations observed in the final measurement.

4.8 Problems

Problem 4.1. Find the general form of the commutator $[\hat{L}_j,[\hat{L}_k,\hat{r}_l]]$. Check your answer by specific examples: $[\hat{L}_x,[\hat{L}_y,\hat{r}_z]]$, $[\hat{L}_x,[\hat{L}_x,\hat{r}_z]]$ and $[\hat{L}_x,[\hat{L}_z,\hat{r}_x]]$.

Problem 4.2. Derive the differential operator (4.26) for the angular momentum squared from the expression for the Laplacian in spherical coordinates that is known from calculus:

$$\hat{\nabla}^2 = \frac{1}{r^2}\left[\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial^2\phi}\right] \tag{4.91}$$

as well as Eqs. (4.22), (4.27) and (4.28).

Problem 4.3. From the expressions (4.25) for the angular momentum components in spherical coordinates, derive these components in Cartesian coordinates [Eq. (4.20)].

Problem 4.4. Show that $[\hat{L}_x,\hat{L}_y] = i\hbar\hat{L}_z$ for the angular momentum components expressed as differential operators

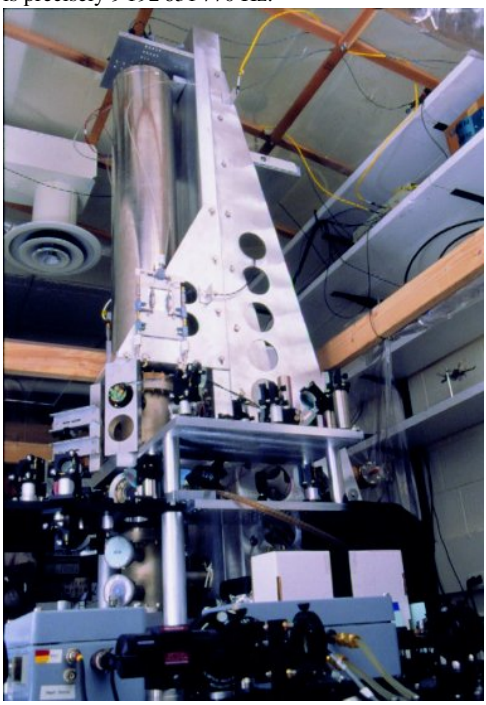
a) in Cartesian coordinates;
b) in spherical coordinates.

Problem 4.5.* Solve Ex. 4.4 in spherical coordinates and check consistency with the solution in Cartesian coordinates.

Problem 4.6. For $l = 3/2$:

Box 4.8 Atomic clock

The atomic clock employs a narrowband, stable, and reproducible atomic transition as the "pendulum". In fact, the very definition of the second is linked to the frequency corresponding to the transition between the two hyperfine levels of the ground state of the cesium 133 atom. The second is defined so that the transition frequency $\Delta E/2\pi\hbar$, with ΔE the energy difference between the levels, is precisely 9 192 631 770 Hz.



The photograph above (source: *Wikipedia*) shows the NIST F1 cesium fountain clock in Colorado, the United States' primary time and frequency standard. Its relative uncertainty is 3.1×10^{-16}, which corresponds to about a second in a hundred million years. The clock utilizes Ramsey spectroscopy. The cesium atoms are collected and cooled to microkelvin temperatures in a magneto-optical trap, and then "tossed" upwards by means of a laser beam. During the free fall, they are subjected to two Ramsey pulses separated by a free fall time of 0.56 s. This free fall operation ensures that the atomic level energies are not perturbed during the experimental cycle. No dc field is necessary because the energy level splitting is present naturally.

After the second Ramsey pulse, the populations of the two atomic levels are measured. The measurement data shows by how much the frequency of the rf field generator that produces the Ramsey pulses has deviated from the atomic transition. That frequency is then adjusted by means of a feedback mechanism.

a) find the matrices of $\hat{L}_x$, $\hat{L}_y$, $\hat{L}_z$, $\hat{L}_\pm$, and $\hat{L}^2$ explicitly;
b) check that these matrices obey $\hat{L}_x^2+\hat{L}_y^2+\hat{L}_z^2=\hat{L}^2$;
c) determine the commutators $[\hat{L}_i,\hat{L}_j]$ in the matrix form and check that they are consistent with the known commutation relations for the angular momentum components.

Problem 4.7. Generalize Ex. 4.28 to a subspace with arbitrary l. Consider the eigenstate $|lm_{\theta\phi}\rangle$ of the observable $\hat{L}_{\vec{R}_{\theta\phi}}$ (which is the projection of $\hat{\vec{L}}$ onto the vector $\vec{R}_{\theta\phi}$) with eigenvalue $m\hbar$. Find the mean values of $\hat{L}_x$, $\hat{L}_y$, $\hat{L}_z$ in this state and show that they are proportional to the projections of the vector $\vec{R}_{\theta\phi}$ onto the corresponding coordinate axes.
Hint: Change the reference frame to (x',y',z'), where the new axis z' is parallel to $\vec{R}_{\theta\phi}$, and express $\hat{L}_x$, $\hat{L}_y$ and $\hat{L}_z$ through $\hat{L}_{x'}$, $\hat{L}_{y'}$ and $\hat{L}_{z'}$.

Problem 4.8. Assuming the radius of the proton to be $r_p\sim 10^{-15}$ m, estimate the fraction of time the electron in the state $|1,0,0\rangle$ spends inside the nucleus. How will your answer change if the electron is replaced by a muon (the muon has the same charge as the electron and the mass $M_\mu=207M_e$)? Why are muonium atoms considered useful for studying nuclear structure?

Problem 4.9. Consider two objects whose angular momentum states are $|l_1,m_1=l_1\rangle$ and $|l_2,m_2=l_2\rangle$. Show that the tensor product state $|l_1,m_1=l_1\rangle\otimes|l_2,m_2=l_2\rangle$ is an eigenstate of the operators $\hat{L}^2$ and $\hat{L}_z$ (where $\hat{\vec{L}}=\hat{\vec{L}}_1+\hat{\vec{L}}_2$) with the eigenvalues corresponding to $l=m=l_1+l_2$.
Hint: Express $\hat{L}_x$ and $\hat{L}_y$ through $L_{\pm,1}$ and $L_{\pm,2}$.

Problem 4.10. As we know, the raising and lowering operators $\hat{L}_\pm$ respectively increase and decrease the eigenvalue $\hat{L}_z$ by $\hbar$. Construct the analogous operators $\hat{L}_\pm^x$ that would raise and lower the eigenstates of $\hat{L}_x$. Specializing to $l=1$,

a) find the matrices of $\hat{L}_\pm^x$ in the canonical basis;
b) find the eigenstates of $\hat{L}_x$ in the matrix form;
c) apply $\hat{L}_\pm^x$ to these eigenstates and check that their action is analogous to the action of $\hat{L}_\pm$ on the eigenstates of $\hat{L}_z$ (up to an arbitrary phase factor that may arise from the randomness in defining the eigenstates of $\hat{L}_x$).

Problem 4.11. The electron in a hydrogen atom is prepared in a state that is a simultaneous eigenstate of the following observables:

- energy with eigenvalue $\sim-(13.6/4)$ eV,
- orbital angular momentum squared with eigenvalue $2\hbar^2$,
- projection of the orbital angular momentum onto the x axis with eigenvalue $\hbar$.

Write the wavefunction of that state.

Problem 4.12. Find the expectation value and the uncertainty of the observables $\hat{x}$, $\hat{y}$, $\hat{z}$ in the states

a) $|2,1,0\rangle$,
b) $|2,1,1\rangle$

of the hydrogen atom.

Problem 4.13. Treating the world as the Bloch sphere, write in the canonical basis the spin state corresponding to your city. The Greenwich meridian corresponds to $\phi = 0$.

Problem 4.14. For an arbitrary spin state $\psi_\uparrow |\uparrow\rangle + \psi_\downarrow |\downarrow\rangle$, express the Cartesian components of the corresponding Bloch vector in terms of $\psi_\uparrow$ and $\psi_\downarrow$.

Problem 4.15. Linearly polarized photons with different polarization angles α pass through a quarter-wave plate with its optic axis oriented

a) horizontally;
b) at 45°.

Find the locus of the resulting states on the Bloch sphere.

Problem 4.16. Consider the evolution of the spin state of a spin-1 particle under the action of a constant magnetic field $\vec{B}$ oriented along the x axis. The initial state is $|\psi(0)\rangle = |m_s = 1\rangle$.

a) Find the spin state $|\psi(t)\rangle$ as a function of time in the matrix form, in the eigenbasis of $\hat{S}_z$.
b) Find the mean values $\langle \hat{S}_x(t) \rangle, \langle \hat{S}_y(t) \rangle, \langle \hat{S}_z(t) \rangle$ and check that they are consistent with what is expected classically.
c) The state $|\psi(t)\rangle$ is measured using a Stern–Gerlach apparatus with the magnetic field along the y axis. Find the probability for the particle to end up in each of the three spots. Are the values found at one-quarter and three-quarters of the Larmor period consistent with what you would expect from part (b)?

Problem 4.17. Two spin-$\frac{1}{2}$ particles interact via the Hamiltonian

$$\hat{H} = \chi \hat{S}_A \hat{S}_B.$$

a) Find the 4×4 matrix of the Hamiltonian in the canonical basis.
b) Find the matrix of the evolution operator.
c) Show that the evolution for the time $t = \pi/\hbar\chi$ will swap the states of the particles, i.e., transform any state $|\psi\rangle \otimes |\varphi\rangle$ into $|\varphi\rangle \otimes |\psi\rangle$.

Problem 4.18. An electron is placed in a harmonic oscillator potential and prepared in a state in which its spin and motional degrees of freedom are in an entangled state

$$|\Psi\rangle = \mathcal{N} \left(|\uparrow\rangle |\alpha\rangle + |\downarrow\rangle |-\alpha\rangle \right),$$

where $|\alpha\rangle$ is a coherent state.

a) Find the normalization factor $\mathcal{N}$.

b) The number of vibrational quanta n is measured. For each n, find the probability of the corresponding result and the direction of the spin after the measurement.
c) The projection of the spin onto a vector $\vec{R}_{\theta,\phi}$ is measured. Find the probability of each possible result and the wavefunction of the electron after the measurement in the position basis.

Problem 4.19. Solve Ex. 4.74(a), 4.75, and 4.76 using the Schrödinger evolution of the spin state in the matrix form, without appealing to the geometry of the Bloch vector.

Problem 4.20. In a spin echo experiment, instead of the standard excitation pulse sequence $\left(\frac{\pi}{2}, \pi\right)$, the sequence

a)* $\left(\frac{\pi}{2}, \theta\right)$;
b) (θ, π)

is applied. Calculate the amplitude of the echo signal in comparison with the one obtained under the standard pulse sequence.

Problem 4.21. In a Ramsey spectroscopy experiment, instead of the standard excitation pulse sequence $\left(\frac{\pi}{2}, \pi/2\right)$, the sequence

a) $\left(\frac{\pi}{2}, \theta\right)$;
b) $\left(\theta, \frac{\pi}{2}\right)$;
c) (θ, θ)

is applied. Calculate the populations of the states $|\uparrow\rangle$ and $|\downarrow\rangle$ as a function of θ and Δt, where Δ is the detuning of the rf field and t the duration of the experiment.

Chapter 5
Quantum physics of complex systems

5.1 The density operator

5.1.1 Pure and mixed states

In many practical cases we may not have complete knowledge of the state of a quantum system. Our knowledge could be of the form of a *statistical ensemble, or mixture*: that our system is in state $|\psi_1\rangle$ with a probability p_1, in state $|\psi_2\rangle$ with a probability p_2, etc., with $\sum_i p_i = 1$. The states $|\psi_i\rangle$ are all assumed to be normalized, but may not necessarily be orthogonal; their number does not have to be equal to the dimension of the Hilbert space.

Situations of such limited knowledge occur very often. One case in point is the mixed state that is produced when we lose a part of an entangled state, as discussed in Sec. 2.2.4. Another example is if we have a large collection of particles in different states and we are interested in a value of an observable that is averaged over all these particles, such as in the case of inhomogeneously broadened ensembles in magnetic resonance (Sec. 4.7.4).

The first thing we need to do is to invent a convenient mathematical formulation of the information we have about the ensemble. While listing all possible states and their probabilities would work in principle, it is verbose and difficult to handle. There exists a description that is much more concise, yet sufficient for all practical purposes. It is the operator

$$\hat{\rho} = \sum_i p_i \, |\psi_i\rangle\langle\psi_i| \,, \tag{5.1}$$

which is called the *density operator* of the ensemble. The matrix of the density operator $\rho_{jk} = \langle v_j | \, \hat{\rho} \, | v_k \rangle$ in any orthonormal basis $\{|v_j\rangle\}$ is called the *density matrix*[1].

[1] The density operator formalism was proposed independently by John von Neumann and Lev Landau in 1927. The terms "density matrix" and "density operator" are often used interchangeably. We will do so in this book, too.

A. I. Lvovsky, *Quantum Physics*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-662-56584-1_5

Exercise 5.1. For the following ensembles within the Hilbert space of polarization states of a single photon, write the density operators in the Dirac notation and the density matrices in the canonical basis:

a) $|H\rangle$;
b) $\psi_H |H\rangle + \psi_V |V\rangle$;
c) $|+45^\circ\rangle$ with a probability 1/2, $|-45^\circ\rangle$ with a probability 1/2;
d) $(|H\rangle + |V\rangle)/\sqrt{2}$ with a probability 1/2, $|H\rangle$ with a probability 1/4, $|V\rangle$ with a probability 1/4.

Exercise 5.2. Suppose an ensemble is measured in the basis $\{|v_m\rangle\}$ $(1 \le m \le N = \dim \mathbb{V})$. Show that the probability of detecting a specific basis element $|v_m\rangle$ is the corresponding diagonal element of the density matrix in that basis:

$$\mathrm{pr}_m = \langle v_m | \hat{\rho} | v_m \rangle . \tag{5.2}$$

Hint: you may find it useful to learn about conditional probabilities (see Sec. B.2).

Physical properties of a quantum state manifest themselves through measurements. Exercise 5.2 shows that the density operator can be used to calculate probabilities of any measurement result as precisely as the full verbal description of a statistical ensemble. So the density operator contains complete information about the ensemble's measurable physical properties. This is what I meant by saying above that the density operator is "sufficient for all practical purposes".

Equation (5.2) is the extension of Born's rule, which we studied in the context of the Measurement Postulate, to statistical ensembles.

Exercise 5.3. A photon's polarization is described by a density matrix $\hat{\rho}$. The polarization is measured in the

a) canonical,
b) diagonal,
c) circular bases.

Express the probability associated with each measurement outcome through the elements of the matrix of $\hat{\rho}$ in the canonical basis.

Exercise 5.4. Show that the density operator of an ensemble of unnormalized states $\{|\psi_i\rangle\}$ is given by $\hat{\rho} = \sum_i |\psi_i\rangle\langle\psi_i|$.

A given density operator does not necessarily represent a unique ensemble, as is evident from the following.

Exercise 5.5. Show that the following statistical ensembles are associated with the same density operator:

- $|H\rangle$ with a probability 1/2, $|V\rangle$ with a probability 1/2;
- $|+\rangle$ with a probability 1/2, $|-\rangle$ with a probability 1/2;
- $|R\rangle$ with a probability 1/2, $|L\rangle$ with a probability 1/2;
- $|\theta\rangle$ with a probability 1/2, $|\pi/2+\theta\rangle$ with a probability 1/2.

Different ensembles described by the same density operator (such as those above) exhibit identical physical behavior, so it is fundamentally impossible to determine by measurement which of the ensembles we are dealing with. Therefore, at least some of the information contained in the description of an ensemble as a list of states and probabilities is redundant. This is an additional argument in favor of using the density matrix instead.

From now on, we will use the term "state" for both *pure* states that can be associated with a specific element $|\psi\rangle$ of the Hilbert space and statistical ensembles identified by a density operator. If the state is not pure — so its density operator cannot be written in the form $\hat{\rho} = |\psi\rangle\langle\psi|$ — we shall call it *mixed*.

Exercise 5.6. Show that an ensemble (5.1) with two or more nonzero terms with unequal $|\psi_i\rangle$ cannot correspond to a pure state.

Exercise 5.7. Which of the states of Ex. 5.1 are pure?

A special status among mixed states belongs to the *fully mixed* state, whose density operator is $\hat{\rho} = \hat{\mathbf{1}}/N$ (where N is the dimension of the Hilbert space). As is evident from the following exercise, if a system is in the fully mixed state, this means that no information whatsoever is available about the quantum system.

Exercise 5.8. Show that, if a fully mixed state is measured in any orthonormal basis, the probability for each result is $1/N$.

Exercise 5.9. Show that all states in Ex. 5.5 are fully mixed.

Exercise 5.10. For the subspace corresponding to the orbital quantum number $l = 1$, find the density matrix of each of the eigenstates of observable $\hat{L}_x$ with eigenvalues $\hbar$, 0, and $-\hbar$. Then find the density matrix of the mixture of these states with probability $\frac{1}{3}$ each. Show that the result is a fully mixed state.
Hint: Use the result of Ex. 4.27.

5.1.2 Diagonal and off-diagonal elements

Exercise 5.11. Show that the diagonal elements of the density matrix of a physical state in any basis

a) are real and nonnegative;
b) add up to one.

Exercise 5.12* For each off-diagonal element ρ_{mn} of a density matrix, show that

a)
$$|\rho_{mn}|^2 \leq \rho_{mm}\rho_{nn}, \tag{5.3}$$

b) the inequality (5.3) saturates for all elements of a density matrix if and only if the corresponding state is pure.

The last exercise, together with Ex. 5.2, reveals the different roles played by the diagonal and off-diagonal elements of the density matrix. The diagonal elements show the *probabilities* of detecting the system in the associated basis states. Off-diagonal elements, on the other hand, show the extent to which the relevant basis elements are in a superposition state or a statistical mixture — in other words, the degree of *coherence* between these elements (see Sec. 2.4.2). Here is an example.

Exercise 5.13§ Find the density matrices of the following states of an electron's spin in the canonical spin basis.

a) $\frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle)$;
b) $\frac{1}{\sqrt{2}}(|\uparrow\rangle - |\downarrow\rangle)$;
c) an equal probability mixture of the states in (a) and (b).

Answer:

a)
$$\frac{1}{2}(|\uparrow\rangle + |\downarrow\rangle)(\langle\uparrow| + \langle\downarrow|) \simeq \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix};$$

b)
$$\frac{1}{2}(|\uparrow\rangle - |\downarrow\rangle)(\langle\uparrow| - \langle\downarrow|) \simeq \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix};$$

c)
$$\frac{1}{2}(|\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow|) \simeq \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

All these states contain equal fractions of the spin-up and spin-down components — hence in all three cases the diagonal elements of the density matrix are the same and equal $1/2$. However, the first two states above are pure while the third is fully mixed. Accordingly, the first two states have significant off-diagonal elements while the third has none.

Exercise 5.14§ For a spin-$\frac{3}{2}$ particle, find the density matrices of states

a) $|\psi\rangle = \frac{1}{\sqrt{2}}(\left|\frac{3}{2}\right\rangle + \left|\frac{1}{2}\right\rangle)$;
b) $|\phi\rangle = \frac{1}{\sqrt{2}}(\left|-\frac{1}{2}\right\rangle + \left|-\frac{3}{2}\right\rangle)$;
c) $\frac{1}{\sqrt{2}}(|\psi\rangle + |\phi\rangle)$;
d) the equal-probability mixture of $|\psi\rangle$ and $|\phi\rangle$.

Answer:

$$\text{a)}\frac{1}{2}\begin{pmatrix} 1 & 1 & & \\ 1 & 1 & & \\ & & & \\ & & & \end{pmatrix};\ \text{b)}\frac{1}{2}\begin{pmatrix} & & & \\ & 1 & 1 & \\ & 1 & 1 & \\ & & & \end{pmatrix};\ \text{c)}\frac{1}{4}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix};\ \text{d)}\frac{1}{4}\begin{pmatrix} 1 & 1 & & \\ 1 & 1 & & \\ & & 1 & 1 \\ & & 1 & 1 \end{pmatrix}.$$

This is a somewhat more sophisticated example. Here, comparing cases (c) and (d), we see that the off-diagonal terms that are responsible for the coherence between states $|\psi\rangle$ and $|\phi\rangle$ are present in the density matrix of the superposition, but absent in that of the mixture. On the other hand, the off-diagonal elements $\rho_{12}, \rho_{21}, \rho_{34}, \rho_{43}$ that arise due to coherences within individual states $|\psi\rangle$ and $|\phi\rangle$ are not erased in the density matrix (d) in spite of that state being a mixture. In case (d), the inequality (5.3) saturates for some, but not all, elements of the off-diagonal elements of $\hat{\rho}$.

Exercise 5.15. Show that the density operator is Hermitian.

Exercise 5.16. Show that, for a given density operator, there exists a spectral decomposition in the form[2]

$$\hat{\rho} = \sum_{i=1}^{N} q_i |v_i\rangle\langle v_i|, \tag{5.4}$$

where $\{|v_i\rangle\}$ is an orthonormal basis, all $q_i \geq 0$, and $\sum_i q_i = 1$.

The above spectral decomposition, which diagonalizes the density matrix, is useful in a number of ways. It can immediately tell us, for example, if the state is pure or mixed (see the next exercise). Additionally, the absence of the off-diagonal terms means there is no quantum coherence between different elements of the diagonalizing basis, which means that the state is a probabilistic mixture of these elements.

Exercise 5.17. Find the spectral decomposition of the density operators in Ex. 5.1.

Exercise 5.18. How many nonzero elements can the diagonalized density matrix of a pure state contain?

Exercise 5.19. Show that the density operator is non-negative.

Let us now define the analog of the density matrix for continuous bases, such as the position and momentum. As discussed in Chapter 3 [see Eq. (3.13)], operators in such bases are represented by two-variable functions rather than matrices. In particular, the density operator Eq. (5.1) is represented by

$$\rho(x,x') = \langle x|\,\hat{\rho}\,|x'\rangle = \sum_i p_i \psi_i(x)\psi_i^*(x'), \tag{5.5}$$

where $\psi_i(x)$ are the wavefunctions of the statistical ensemble components.

[2] Note that the existence of a spectral decomposition Eq. (5.4) does not trivially follow from the definition (5.1) of the density matrix. The two expressions are quite similar, but the elements of the sum in Eq. (5.4) comprise an orthonormal basis, while in Eq. (5.1) they are arbitrary states.

Exercise 5.20. Express the density operator of the state $a|0\rangle + b|1\rangle$ of a harmonic oscillator

a) in the Fock basis;
b) in the position basis.

Exercise 5.21. For a normalized density operator $\hat{\rho}$, show that

a) $\hat{\rho}$ cannot be unitary for any Hilbert space of dimension greater than one;
b) equality $\hat{\rho} = \hat{\rho}^2$ holds if and only if $\hat{\rho}$ represents a pure state.

Exercise 5.22. Consider a mixture of states that are themselves statistical ensembles: state $\hat{\rho}_1$ occurring with probability p_1, $\hat{\rho}_2$ with probability p_2, *etc.*, with $\sum_i p_i = 1$.

a) Show that such an ensemble is described by the density operator

$$\hat{\rho} = \sum_i p_i \hat{\rho}_i. \tag{5.6}$$

b) Show that this ensemble cannot be pure if at least one of its terms is mixed.

5.1.3 Evolution

Exercise 5.23. Use the Schrödinger equation to show that

a) the differential equation for the evolution of the density matrix in time is

$$\frac{d\hat{\rho}}{dt} = -\frac{\mathrm{i}}{\hbar}[\hat{H}, \hat{\rho}]; \tag{5.7}$$

b) the evolution of the density operator can be written as

$$\hat{\rho}(t) = \hat{U}\hat{\rho}(0)\hat{U}^\dagger, \tag{5.8}$$

where $\hat{U} = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\hat{H}t}$.

Differential equations for the evolution of density operators, such as Eq. (5.7), are frequently referred to as *quantum master equations*.

Notice the opposite sign in Eqs. (5.7) and (5.8) as compared with the otherwise similar Eqs. (3.129) and (3.127), respectively. This difference may appear questionable: why is the evolution of the density matrix opposite to that of other operators? The answer is that the equations in Sec. 3.9 are in the Heisenberg picture, where we assume that quantum states are stationary and the operators corresponding to physical observables evolve. Here, on the other hand, we work in the Schrödinger picture, where the evolution is associated with the states, and hence with the density matrix that expresses the state. Therefore the observable operators in Sec. 3.9 and the density operator in this section are of different nature, and there is no reason to expect their evolution to obey the same equations.

Exercise 5.24. For the state which at time $t = 0$ is

a) a superposition $(|E_1\rangle + |E_2\rangle)/\sqrt{2}$,
b) a statistical mixture $(|E_1\rangle\langle E_1| + |E_2\rangle\langle E_2|)/2$

of energy eigenstates, write the density matrix as a function of time in the energy eigenbasis.
Answer:

a) $\hat{\rho}(t) = \hat{\rho}(0) = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$;

b) $\hat{\rho}(t) = \frac{1}{2}\begin{pmatrix} 1 & e^{\frac{i}{\hbar}(E_2-E_1)t} \\ e^{-\frac{i}{\hbar}(E_2-E_1)t} & 1 \end{pmatrix}$.

Generalizing Ex. 5.24(a), we see that, if an ensemble is a statistical mixture of energy eigenstates, its density operator does not change under the Schrödinger evolution. This result may appear surprising at first. We have learned that states with the energy E acquire a quantum phase $\mathrm{e}^{-\mathrm{i}Et/\hbar}$ while evolving. States associated with different energies should acquire different phases — so how come we do not see that in the evolution of the density matrix?

The answer is, when we are dealing with a *statistical mixture* of states, their phases are unphysical: they cannot be observed in a measurement. A mixture of states $|E_1\rangle$ and $|E_2\rangle$ behaves in an experiment exactly the same way as a mixture of $|E_1\rangle\,\mathrm{e}^{-\mathrm{i}E_1t/\hbar}$ and $|E_2\rangle\,\mathrm{e}^{-\mathrm{i}E_2t/\hbar}$. As discussed earlier (Sec. 5.1.1), the purpose of the density matrix is to describe, as succinctly as possible, the state's physical properties. Two states that have the same properties will have the same density matrix.

In contrast, if we have a *coherent superposition* of two states with different energies [Ex. 5.24(b)], its density matrix (specifically, the off-diagonal elements) does evolve, reflecting the change in the physical properties of that state with time.

Exercise 5.25. For a state that is initially a mixture of $|\uparrow\rangle$ with probability $3/4$ and $|\downarrow\rangle$ with probability $1/4$, get practice finding the evolution $\hat{\rho}(t)$ of the density matrix in a magnetic field B pointing along the x axis by using three methods:

a) calculating the evolution of each pure state component separately and subsequently obtaining the density matrix of the ensemble;
b) first calculating the density matrix of the initial ensemble and subsequently letting it evolve according to Eq. (5.8);
c) solving Eq. (5.7) in matrix form.

5.2 Trace

The *trace* of an operator $\hat{A}$ is the sum of the diagonal elements of its matrix:

$$\mathrm{Tr}\hat{A} = \sum_{m=1}^{n} A_{nn}. \tag{5.9}$$

Traces play an important role because they express the effects of measurements on quantum states when these states are written in the form of density matrices. Before we discover these effects, let us recall some of the trace's salient properties, as known from linear algebra, and derive a few new ones relevant to quantum physics.

Exercise 5.26. Show that the trace of an operator is the same in all orthonormal bases.

This explains why we say "trace of an operator" rather than "trace of a matrix". The same operator will have different matrices in different orthonormal bases, but all these matrices will have the same sum of diagonal elements.

Exercise 5.27. Show that the trace of a density operator representing a physical state equals 1.

Exercise 5.28.§ Operators $\hat{A}$ and $\hat{B}$ have matrices A_{ij} and B_{ij}, respectively, in the same orthonormal basis. Show that

$$\mathrm{Tr}(\hat{A}\hat{B}) = \sum_{ij} A_{ij}B_{ji}. \tag{5.10}$$

Exercise 5.29. Show that, for any operators,

a) $\mathrm{Tr}(\hat{A}\hat{B}) = \mathrm{Tr}(\hat{B}\hat{A})$;
b) $\mathrm{Tr}(\hat{A}_1 \ldots \hat{A}_k) = \mathrm{Tr}(\hat{A}_k\hat{A}_1 \ldots \hat{A}_{k-1})$ (*chain rule*).

Exercise 5.30. Find an example showing that, generally, $\mathrm{Tr}(\hat{A}\hat{B}\hat{C}) \neq \mathrm{Tr}(\hat{B}\hat{A}\hat{C})$.

Exercise 5.31. For an operator $\hat{A}$ and vectors $|\psi\rangle$ and $|\varphi\rangle$, show that

$$\langle \psi | \hat{A} | \varphi \rangle = \mathrm{Tr}(\hat{A}\,|\varphi\rangle\langle\psi|). \tag{5.11}$$

Exercise 5.32. Show that the trace of the squared density matrix is useful as a measure of a state's degree of purity. Specifically, for a physical state $\hat{\rho}$, show that $1/N \leq \mathrm{Tr}(\hat{\rho}^2) \leq 1$, with the first inequality saturating if and only if $\hat{\rho}$ represents the completely mixed state, and the second inequality saturating if and only if $\hat{\rho}$ describes a pure state.

Let us now restate the Measurement Postulate of quantum mechanics in terms of density matrices.

Exercise 5.33. Suppose a projective measurement in the basis $\{|v_m\rangle\}$ is performed on an ensemble $\hat{\rho}$ and yields some result $|v_m\rangle$. Show that:

a) the (unnormalized) ensemble after the measurement is given by

$$\hat{\Pi}_m\hat{\rho}\hat{\Pi}_m, \tag{5.12}$$

where $\hat{\Pi}_m = |v_m\rangle\langle v_m|$ is the projection operator;
b) the probability of obtaining the result $|v_m\rangle$ is

$$\mathrm{pr}_m = \mathrm{Tr}(\hat{\Pi}_m\hat{\rho}) = \mathrm{Tr}(\hat{\rho}\hat{\Pi}_m). \tag{5.13}$$

Exercise 5.34. Apply Eq. (5.12) to determine the probability of detecting a $(+45^\circ)$-polarization in a photon described by each of the ensembles of Ex. 5.1. Check that your findings are consistent with the probabilities expected by treating each state as a statistical ensemble of pure states.

Exercise 5.35. A state is represented in the basis $\{|v_m\rangle\}$ by the matrix

$$\hat{\rho} \simeq \begin{pmatrix} \rho_{11} & \cdots & \rho_{1N} \\ \vdots & \ddots & \vdots \\ \rho_{N1} & \cdots & \rho_{NN} \end{pmatrix}. \tag{5.14}$$

Suppose this state is measured in the same basis $\{|v_m\rangle\}$. The measurement is non-destructive, but its outcome is unknown to us. Show that the density matrix after the measurement will be of the form

$$\hat{\rho}_{\text{after}} = \begin{pmatrix} \rho_{11} & & \\ & \ddots & \\ & & \rho_{NN} \end{pmatrix}. \tag{5.15}$$

That is, the off-diagonal elements of the density matrix will vanish after the measurement, but the diagonal elements will remain intact.

I emphasize that this simple rule applies only if the density matrix is written in the same basis in which the measurement is performed. Let me illustrate this by an example.

Exercise 5.36. A photon polarized at $+45^\circ$ is measured in the canonical basis. Find the density matrix before and after the measurement

a) in the canonical basis,
b) in the diagonal basis.

Exercise 5.37. Show that the expectation value of any observable $\hat{V}$ in the state $\hat{\rho}$ is

$$\langle V \rangle = \mathrm{Tr}(\hat{\rho}\hat{V}) = \mathrm{Tr}(\hat{V}\hat{\rho}). \tag{5.16}$$

Exercise 5.38. Using the density matrix formalism in the Schrödinger picture, specifically Eqs. (5.7) and (5.16), reproduce the Heisenberg equation of motion (3.129) for the mean value of an arbitrary observable:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle V \rangle = \frac{\mathrm{i}}{\hbar}\left\langle [\hat{H}, \hat{V}] \right\rangle. \tag{5.17}$$

5.3 Partial trace

Let us now return to the question we asked in Chapter 2. Suppose Alice and Bob share a state $\hat{\rho}_{AB}$ that is a density matrix over the tensor product Hilbert space. Alice

either loses her portion of the state, or measures it in some basis, but does not tell Bob the result. What does Bob's portion of the state become? Or, to formulate the question in the language we just learned, what will be the density operator of Bob's state[3]?

A *partial trace* of a bipartite state $\hat{\rho}_{AB}$ over Hilbert space $\mathbb{V}_A$ is the operator in Hilbert space $\mathbb{V}_B$ defined by

$$\mathrm{Tr}_A(\hat{\rho}_{AB}) = \sum_{m=1_A}^{N} \langle v_m | \hat{\rho}_{AB} | v_m \rangle_A , \tag{5.18}$$

where $\{|v_m\rangle\}$ is an orthonormal basis in $\mathbb{V}_A$. The procedure for calculating the partial trace is sometimes referred to as *tracing over* the space $\mathbb{V}_A$.

Exercise 5.39. Alice and Bob share the state $\hat{\rho}_{AB}$. Alice performs a local measurement in the basis $\{|v_m\rangle\}$ on her part of the ensemble. Show that:

a) if Alice's measurement result is known to be a specific $|v_m\rangle$, the resulting (unnormalized) bipartite state is described by $\hat{\Pi}_{A,m}\hat{\rho}_{AB}\hat{\Pi}_{A,m} = |v_m\rangle\langle v_m| \otimes \langle v_m| \hat{\rho}_{AB} |v_m\rangle$, and Bob's portion of this state is

$$\hat{\rho}_{B,m} = \langle v_m| \hat{\rho}_{AB} |v_m\rangle = \mathrm{Tr}_A(\hat{\Pi}_{A,m}\hat{\rho}_{AB}) = \mathrm{Tr}_A(\hat{\Pi}_{A,m}\hat{\rho}_{AB}\Pi_{A,m}); \tag{5.19}$$

b) if Alice's measurement result is unknown, the reduced density operator of Bob's state is the partial trace

$$\hat{\rho}_B = \mathrm{Tr}_A(\hat{\rho}_{AB}).$$

To make this theory a bit less abstract, let us look at a couple of examples.

Exercise 5.40. Perform the following calculations for the setting of Ex. 2.45.

a) In that exercise, we found the ensembles describing the states of Bob's photon for Alice performing her measurement in the canonical and diagonal bases. For each of these ensembles, find the corresponding density matrix in the canonical basis. Check that the density matrix does not depend on Alice's basis.
b) Find the reduced density matrices of Bob's photon in the canonical basis using the partial trace formalism. Check consistency with part (a).

Exercise 5.41. For each of the four Bell states, find the reduced density operator associated with Alice's and Bob's qubits.

Bob's reduced density operator must be the same no matter which basis Alice chooses for her measurement. If this were not the case, Alice would be able to instantly transfer information to Bob just by choosing a particular basis, or simply choosing whether or not to throw her portion of the state away (see Ex. 2.43). Let us now show this rigorously in the language of density operators.

Exercise 5.42. Show that the partial trace is independent of the choice for Alice's basis in which it is calculated.

[3] It is sometimes called Bob's *reduced density operator*.

Exercise 5.43. Show that $\mathrm{Tr}_A(\hat{\rho}_{AB})$ has trace 1 if $\hat{\rho}_{AB}$ is a physical state.

Exercise 5.44. Suppose Alice and Bob share a bipartite state. Show that

a) if the bipartite ensemble is in a pure, separable (non-entangled) state, then both Alice's and Bob's reduced density operators are also pure states;
b) the reduced density operator of an entangled state is always a mixed state.
Hint: use Eq. (2.15).

The partial trace formalism allows us to reproduce our previous result regarding the effect of a measurement on the density matrix (Ex. 5.35), but now analyzing the measurement at a deeper level, using the von Neumann model.

Exercise 5.45. Let the initial state of a quantum system be described in some basis $\{|v_i\rangle\}$ by the density operator (5.14). A measurement is performed on this system in the same basis $\{|v_i\rangle\}$. This measurement will entangle the system with the apparatus according to Eq. (2.33). Show that, if the apparatus is removed from the entangled state, the reduced density matrix of the system will only have diagonal elements, as in Eq. (5.15).

This result has important implications for decoherence, which, as we discussed in Sec. 2.4.2, can be interpreted as an "inadvertent" von Neumann measurement of the system by the environment in the decoherence-preferred basis. The subsequent loss of information about the environment corresponds to the partial trace operation. The corresponding effect on the system's density matrix (written in the decoherence-preferred basis) is to strip away its off-diagonal elements. We will study a few examples to this effect in Sec. 5.5.

Taking the partial trace is an irreversible operation: it is impossible to get $\hat{\rho}_{AB}$ back from $\mathrm{Tr}_A(\hat{\rho}_{AB})$. This is the mathematical reason why decoherence, in contrast to unitary quantum evolution, is an irreversible process.

5.4 Density matrix and Bloch vector

In Section 4.5 we associated any state of a qubit with a vector on the Bloch sphere. If the physical system associated with the qubit is the spin-$\frac{1}{2}$ particle, the coordinates of the Bloch vector equal the mean values of the corresponding projections of the angular momentum [Ex. 4.48(c)]. Here I would like to extend the notion of the Bloch vector to density matrices.

This extension is straightforward. For any ensemble

$$\hat{\rho} = \sum_i p_i \, |\psi_i\rangle\langle\psi_i| \, ,$$

the Bloch vector is defined as

$$\vec{R}_{\hat{\rho}} = \sum_i p_i \vec{R}_i, \tag{5.20}$$

where each $\vec{R}_i$ is the Bloch vector of the corresponding state $|\psi_i\rangle$. That is, the Bloch vector of an ensemble is a weighted average of its components.

Exercise 5.46. Show that the Cartesian coordinates of the Bloch vector $\vec{R}_{\hat{\rho}}$ defined by Eq. (5.20) equal the mean values of the observables $\hat{\sigma}_x$, $\hat{\sigma}_y$, $\hat{\sigma}_z$ for the corresponding state $\hat{\rho}$.
Hint: According to Eq. (5.16), you need to show that

$$\vec{R}_{\hat{\rho}} = \mathrm{Tr}[\hat{\rho}\hat{\vec{\sigma}}]. \tag{5.21}$$

Exercise 5.47.§ Calculate the Bloch vector explicitly in terms of the elements of the density matrix

$$\hat{\rho} \simeq \begin{pmatrix} \rho_{\uparrow\uparrow} & \rho_{\uparrow\downarrow} \\ \rho_{\downarrow\uparrow} & \rho_{\downarrow\downarrow} \end{pmatrix}$$

Answer:

$$R_x = \langle\sigma_x\rangle = \rho_{\uparrow\downarrow} + \rho_{\downarrow\uparrow}; \tag{5.22a}$$
$$R_y = \langle\sigma_x\rangle = \mathrm{i}\rho_{\uparrow\downarrow} - \mathrm{i}\rho_{\downarrow\uparrow}; \tag{5.22b}$$
$$R_z = \langle\sigma_x\rangle = \rho_{\uparrow\uparrow} - \rho_{\downarrow\downarrow}. \tag{5.22c}$$

Exercise 5.48. Show that

a) the length of the Bloch vector of a mixed state is less than one;
b) the Bloch vector of the fully mixed state is zero.

Exercise 5.49. We have shown previously [see Eq. (4.77)] that the Bloch vector of a spin-$\frac{1}{2}$ particle in a pure state precesses in a magnetic field in the same way as a classical magnetic moment. Show that this result also applies to states described by density operators.

Exercise 5.50. Calculate the trajectory of the Bloch vector from the time-dependent density matrix obtained in Ex. 5.25 and show that it precesses around the magnetic field as expected classically [Eq. (4.77)].

Exercise 5.51. Show that the length of the Bloch vector is related to the purity factor of the corresponding state (Ex. 5.32) according to

$$\mathrm{Tr}\hat{\rho}^2 = \frac{|\vec{R}_{\hat{\rho}}|^2}{2} + \frac{1}{2}. \tag{5.23}$$

Hint: Suppose the state $\hat{\rho}$ has the spectral decomposition $\hat{\rho} = p|v_1\rangle\langle v_1| + (1-p)|v_2\rangle\langle v_2|$. Relate both $|\vec{R}_{\hat{\rho}}|$ and $\mathrm{Tr}\hat{\rho}^2$ to p.

Exercise 5.52. Show that any Bloch vector of length $|\vec{R}_{\hat{\rho}}| \le 1$ uniquely identifies the corresponding density matrix.

Let us summarize the above results. Similarly to pure states, the Bloch vector of a mixed state corresponds to the quantum mean value of the spin vector operator in

that state. There is a one-to-one correspondence between states (pure or mixed) and Bloch vectors. However, Bloch vectors of mixed states terminate *inside* the Bloch sphere rather than on its surface. The more mixed the state, the shorter the Bloch vector; the fully mixed state corresponds to a zero vector in the center of the Bloch sphere.

5.5 Density matrix and magnetic resonance

In Chapter 4, we studied the basics of magnetic resonance. However, the pure state formalism we used there did not allow us to account for interaction between the spins and the environment, or *homogeneous dephasing (relaxation)*, which is an essential part of this phenomenon. Because relaxation is associated with the loss of a state's purity, its analysis requires the use of density operators.

There are two primary relaxation mechanisms: decoherence and thermalization.

5.5.1 Decoherence

The decoherence of spin states is brought about by their mutual interaction; for this reason, this mechanism is referred to as the *spin–spin relaxation*. As is usually the case for internal degrees of freedom (Sec. 2.4.2), the energy eigenbasis is the decoherence-preferred basis. When the particles interact with each other, the populations of the energy levels do not change, but their energy eigenstates accumulate random phases, which leads to loss of coherence between them.

We will be studying relaxation in the absence of the rf field, assuming that this field is applied in a pulsed manner, so that the decoherence during the pulses is insignificant. We choose the z axis along the dc field $\vec{B}_0$, so the Hamiltonian (4.76) $\hat{H} = -\hat{\vec{\mu}} \cdot \vec{B}_0 = -\gamma \hat{S}_z B_0$. Then the basis consisting of the eigenstates of $\hat{S}_z$ is also, conveniently, the Hamiltonian's eigenbasis, and hence also the decoherence-preferred basis[4].

We found in Sec. 5.3 that decoherence removes the off-diagonal elements of the density matrix. However, this result was obtained for a single copy of the decohering object. In our case, the density matrix represents a large ensemble of particles, and not all of them decohere at once. Hence the density matrix is affected by the decoherence in a more complex way.

Let us adopt the following model. We assume that each particle, if it happens to interact with the environment, does decohere very rapidly — effectively instantly. This brings about the loss of the off-diagonal elements of the density matrix associated with that specific particle. However, the probability for this event to happen

[4] This is the case for both the stationary and the rotating basis, because they both consist of eigenstates of $\hat{S}_z$.

for each particle within a certain small time interval is finite, and proportional to the duration of that interval. Then, when we average over the many particles making up the ensemble, the off-diagonal elements of the density matrix will decay gradually.

Exercise 5.53. Suppose the probability for an individual particle to decohere within a small time interval Δt is $\Delta t/T_2$, where T_2 is a constant known as the characteristic decoherence time.

a) Show that, in the absence of the Hamiltonian evolution, the elements of the density matrix decay according to the differential equation

$$\left[\frac{\mathrm{d}}{\mathrm{d}t}\rho_{ij}(t)\right]_{\text{decoh}} = \begin{cases} 0, & i = j \\ -\rho_{ij}(t)/T_2, & i \neq j \end{cases}, \tag{5.24}$$

where the subscript "decoh" indicates that the decay is due to the decoherence mechanism.

b)[§] Show that the solution of the above equation is

$$\begin{pmatrix} \rho_{\uparrow\uparrow}(t) & \rho_{\uparrow\downarrow}(t) \\ \rho_{\downarrow\uparrow}(t) & \rho_{\downarrow\downarrow}(t) \end{pmatrix} = \begin{pmatrix} \rho_{\uparrow\uparrow}(0) & \rho_{\uparrow\downarrow}(0)\mathrm{e}^{-t/T_2} \\ \rho_{\downarrow\uparrow}(0)\mathrm{e}^{-t/T_2} & \rho_{\downarrow\downarrow}(0) \end{pmatrix}. \tag{5.25}$$

This behavior — that the diagonal elements of the density matrix are constant, but the off-diagonal elements decay exponentially — is characteristic of decoherence, not only for spin ensembles, but for a wide variety of physical situations.

5.5.2 Thermalization

The second mechanism is the *spin–lattice relaxation*, associated with the thermal translational motion of nuclei. This mechanism tends to bring the spin state into thermal equilibrium with the environment — that is, into the state with density matrix

$$\hat{\rho}_0 \simeq \begin{pmatrix} \rho_{\uparrow\uparrow,0} & 0 \\ 0 & \rho_{\downarrow\downarrow,0} \end{pmatrix}, \tag{5.26}$$

where the populations of the upper and lower energy levels are related according to the Boltzmann distribution

$$\frac{\rho_{\downarrow\downarrow,0}}{\rho_{\uparrow\uparrow,0}} = \frac{\exp(-E_\downarrow/kT)}{\exp(-E_\uparrow/kT)} \overset{(4.73)}{=} \exp\left(-\frac{\gamma\hbar B_0}{kT}\right),$$

with no coherence present between these levels.

Exercise 5.54. The field in a medical MRI scanner, which exploits proton spins, is 1.5 tesla. Calculate the mean difference between the fractions of protons in the spin-up and spin-down states at room temperature under thermal equilibrium.

Exercise 5.55. Find the magnitude and direction of the Bloch vector $\vec{R}_0$ corresponding to Eq. (5.26).
Answer (in Cartesian coordinates):

$$\vec{R}_0 = \left(0, 0, \tanh\frac{\gamma\hbar B_0}{2kT}\right) \overset{\gamma\hbar B_0/2kT \ll 1}{\approx} \left(0, 0, \frac{\gamma\hbar B_0}{2kT}\right). \tag{5.27}$$

By the same logic as above, we assume that the diagonal elements decay exponentially to their thermal values, i.e.,

$$\rho_{\uparrow\uparrow}(t) - \rho_{\uparrow\uparrow,0} = [\rho_{\uparrow\uparrow}(0) - \rho_{\uparrow\uparrow,0}]\mathrm{e}^{-t/T_1}; \tag{5.28a}$$

$$\rho_{\downarrow\downarrow}(t) - \rho_{\downarrow\downarrow,0} = [\rho_{\downarrow\downarrow}(0) - \rho_{\downarrow\downarrow,0}]\mathrm{e}^{-t/T_1}, \tag{5.28b}$$

where T_1 is the characteristic time of thermalization.

Exercise 5.56.§ Show that the decay (5.28) corresponds to the following differential equations:

$$\left[\frac{\mathrm{d}}{\mathrm{d}t}\rho_{\uparrow\uparrow}(t)\right]_{\text{therm}} = -[\rho_{\uparrow\uparrow}(t) - \rho_{\uparrow\uparrow,0}]/T_1; \tag{5.29a}$$

$$\left[\frac{\mathrm{d}}{\mathrm{d}t}\rho_{\downarrow\downarrow}(t)\right]_{\text{therm}} = -[\rho_{\downarrow\downarrow}(t) - \rho_{\downarrow\downarrow,0}]/T_1. \tag{5.29b}$$

Let us now introduce a convention. Thermalization, of course, affects not only the diagonal elements of the density matrix, but the off-diagonal ones as well, causing them to decay exponentially. However, this decay is considered to be a part of the decoherence process, so Eq. (5.24) *incorporates* the contribution of thermalization to the decay of the off-diagonal elements. We will therefore write the differential equation for the density matrix thermalization as follows:

$$\left[\frac{\mathrm{d}}{\mathrm{d}t}\rho_{ij}(t)\right]_{\text{therm}} = \begin{cases} -[\rho_{ii}(t) - \rho_{ii,0}]/T_1, & i = j \\ 0, & i \neq j \end{cases}, \tag{5.30}$$

keeping in mind that the thermalization of the off-diagonal elements is accounted for in the equation for the decoherence.

An apparent consequence of this convention is that T_2 cannot exceed T_1: the off-diagonal elements decay due to both decoherence and thermalization, and the diagonal ones due to thermalization alone. In fact, spins typically decohere much faster than they thermalize, so $T_2 \ll T_1$. Human brain tissues, for example, have $T_1 \sim 1$ s and $T_2 \sim 0.1$ s.

In other physical settings, however, T_2 can be as high as $2T_1$. This is possible if the thermalization mechanism differs from the decoherence mechanism, i.e., if it cannot be modeled as a gradual admixture of the thermal state with the spin ensemble. Such situations are common, for example, in two-level systems corresponding to optical transitions in atoms and molecules. We shall show in Ex. 5.60 that the condition

$T_2 \leq 2T_1$ must hold universally, otherwise the evolution will result in an unphysical density operator.

5.5.3 Relaxation and the Bloch vector

The overall evolution of the density matrix is the result of cumulative action of the Hamiltonian evolution and relaxation. It is given by

$$\frac{\mathrm{d}\hat{\rho}}{dt} = -\frac{\mathrm{i}}{\hbar}[\hat{H},\hat{\rho}] + \left[\frac{\mathrm{d}\hat{\rho}}{\mathrm{d}t}\right]_{\text{relax}}, \tag{5.31}$$

where the first term corresponds to the Schrödinger equation (5.7) and the second and third to the decoherence and thermalization terms, (5.24) and (5.30), respectively. Let us now apply this result to the evolution of the Bloch vector.

Exercise 5.57. Show that the behavior of the Bloch vector components corresponding to Eq. (5.31) is

$$\frac{\mathrm{d}\vec{R}}{\mathrm{d}t} = \gamma\vec{R}\times\vec{B} + \left[\frac{\mathrm{d}\vec{R}}{\mathrm{d}t}\right]_{\text{relax}}, \tag{5.32}$$

where

$$\left[\frac{\mathrm{d}R}{\mathrm{d}t}\right]_{\text{relax}} = \left(-\frac{R_x}{T_2}, -\frac{R_y}{T_2}, -\frac{R_z - R_{0z}}{T_1}\right) \tag{5.33}$$

and $\vec{R}_0$ is the Bloch vector (5.27) of the thermal state.

Exercise 5.58. Show that the following solution satisfies Eq. (5.32) for the rotating-wave approximation Hamiltonian (4.85) in the absence of the rf field, with the spin detuned by Δ from the rotating wave frequency.

$$\begin{aligned} R_x(t) &= [R_x(0)\cos\Delta t - R_y(0)\sin\Delta t]\mathrm{e}^{-t/T_2};\\ R_y(t) &= [R_y(0)\cos\Delta t + R_x(0)\sin\Delta t]\mathrm{e}^{-t/T_2};\\ R_z(t) &= R_0 + [R_z(0) - R_0]\mathrm{e}^{-t/T_1}. \end{aligned} \tag{5.34}$$

Exercise 5.59.[§] Plot the trajectory of the Bloch vector tip under the condition of Ex. 5.58 for

a) $\Delta \neq 0$, $T_1 = 0$, $T_2 = 0$;
b) $\Delta = 0$, $T_2 = T_1/10$;
c) $\Delta = 0$, $T_2 = 2T_1$;
d) $\Delta = 5T_1^{-1}$, $T_2 = 2T_1$.

Assume that the temperature $T = 0$. The initial state corresponds to the spin pointing along the x axis.
Answer: see Fig. 5.1.

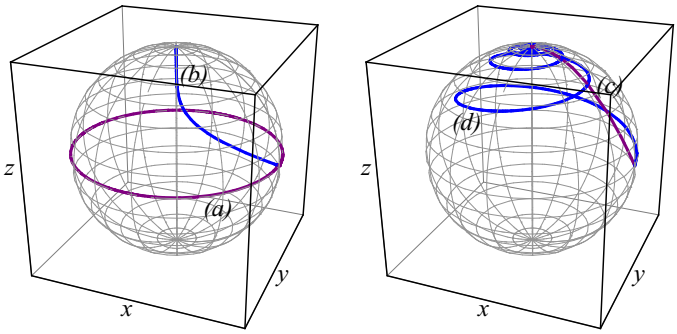


Fig. 5.1 Bloch vector tip trajectories of Ex. 5.59.

We see that decoherence causes the horizontal (x and y) components of the Bloch vector to shrink exponentially, while the vertical (z) component thereof tends to the value corresponding to thermal equilibrium. For this reason, historically, the thermalization of the diagonal elements of the density matrix is sometimes called *longitudinal* relaxation, while the loss of the off-diagonal elements due to decoherence is known as *transverse* relaxation. We can see that this terminology is not quite appropriate; better terms would be "vertical" and "horizontal", respectively.

Exercise 5.60.* Show that T_2 cannot be greater than $2T_1$.
Hint: assume absolute zero temperature. Apply infinitesimal evolution (5.32) in the rotating basis to the Bloch vector with polar coordinates $(\theta,0)$ such that $\theta \ll 1$.

Now that we understand the treatment of relaxation, we are ready to get back to the subject considered at the end of Chapter 4: measurement of relaxation times. As discussed back then, this measurement is important for magnetic resonance imaging applications, because it allows one to distinguish different body tissues from each other. However, the homogeneous relaxation is often obscured by inhomogeneous broadening, which occurs on much faster time scales.

Spin echo is the method of choice for evaluating the transverse relaxation time. In Sec. 4.7.4, we made preliminary calculations to understand the physical principles behind the reversal of the inhomogeneous dephasing that gives rise to the echo. Our next task is to take homogeneous relaxation effects into account.

Exercise 5.61. For an inhomogeneously broadened spin ensemble with inhomogeneous width Δ_0 much greater than the inverse relaxation times T_1^{-1}, T_2^{-1}, show that the mean magnetic moment of an ensemble element (Ex. 4.76) at zero temperature is given by

$$\overline{\langle\vec{\mu}\rangle} = \frac{\hbar\gamma}{2}\left(0, -\mathrm{e}^{-\frac{[\Delta_0(t-2t_0)]^2}{4}}\mathrm{e}^{-t/T_2}, 1-2\mathrm{e}^{-(t-t_0)/T_1}+\mathrm{e}^{-t/T_1}\right). \tag{5.35}$$

In the stationary basis, this magnetic moment will precess around the z axis. Therefore the magnitude of the echo signal is determined entirely by its horizontal component, which decays with the characteristic time T_2.

You may have noticed a subtlety when solving this exercise. In order to calculate the spin echo signal, we had to average the Bloch vector over the ensemble comprising all detunings. But the state associated with each detuning is itself non-pure (because of the homogeneous relaxation), which means that it also represents an ensemble, as discussed earlier in this chapter.

We treated these ensembles quite differently. For decoherence and thermalization, we average continuously over the ensemble throughout the evolution (see Ex. 5.53), thereby accounting in real time for the effects of these phenomena on the spin state. But for the inhomogeneously broadened ensemble, the averaging is performed only once, at the end of the calculation. Why is there this difference?

The reason is the different physics that gives rise to the two types of ensembles. Homogeneous relaxation occurs due to an entangling interaction between the system and the environment. Because the environment is beyond our control, we can trace over it without losing any valuable information; so the system state becomes irreversibly mixed. Inhomogeneous broadening, in contrast, is caused, not by entanglement, but by a slight variation in the physical conditions (and Hamiltonians) under which each spin evolves. Moreover, these conditions do not change with time. The evolution of each individual member of the ensemble is therefore completely predictable and reversible. We must keep track of this evolution, without premature averaging, in order to be able to predict the rephasing of the spins and the echo.

Let us now turn to the longitudinal relaxation time. It can be measured, for example, using the *zero crossing method*. Remarkably, this method does not require the inversion of inhomogeneous dephasing in order to work. The idea is to first flip the Bloch vector of the thermalized ensemble using a π pulse. The ensemble will then gradually re-thermalize. The Bloch vector will relax from pointing downward to pointing upward, so at some point in time its length will be exactly zero.

To measure the length of the Bloch vector after it has relaxed for some time t_0, we apply a $\pi/2$ pulse. The Bloch vector will then become horizontal and start precessing around the dc field, generating the free induction decay signal that is proportional to its length. But if the second pulse is applied at the moment the Bloch vector tip passes through the origin, this signal will vanish.

Exercise 5.62. Show that, in the zero crossing measurement, the free induction decay signal will vanish for $t_0 = T_1 \ln 2$.

5.6 Generalized measurements*

The density operator formalism generalizes the Hilbert Space Postulate of quantum mechanics by accounting for the possibility that we may not have full knowledge of

a quantum state. The Measurement Postulate can be extended in a similar manner, to take into account realistic quantum measurement devices.

5.6.1 A realistic detector

Consider, for example, a polarization measurement setup shown in Fig. 1.2(a). Ideally, we would expect it to perform a measurement of the photon polarization in the canonical basis. Suppose, however, that the beam splitter is not perfect: it may transmit some of the vertical polarization and reflect some of the horizontal. To account for this feature, we introduce the notion of the *output states* of the measurement device — macroscopic (classical) indications that the device can display. In the case of the polarization measurement, assuming perfect detectors, there would be two output states:

- detector in the transmitted channel clicks;
- detector in the reflected channel clicks.

Then, we model our device as an ideal projective measurement in some basis $\{|v_i\rangle\}$, followed by a "scrambler" (Fig. 5.2). The scrambler is a classical device that functions as follows: for each output $|v_i\rangle$ of the quantum measurement, it randomly, with probability μ_{ji}, chooses the jth output state. This state is then "displayed" by the detector.

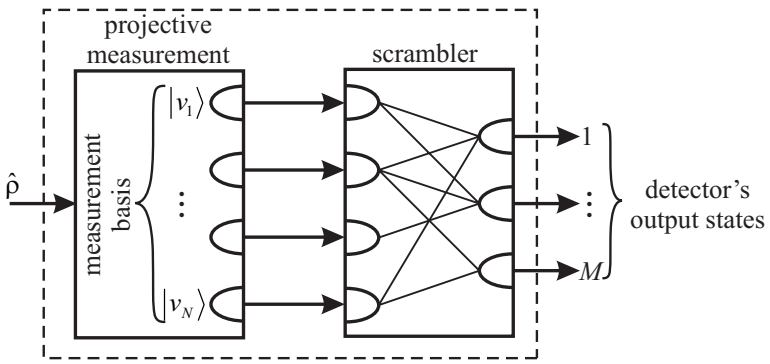


Fig. 5.2 Model of a realistic detector described by a POVM.

Exercise 5.63. Consider a realistic polarization detector that consists of an ideal polarization projective measurement in a canonical basis, followed by a scrambler that maps the measurement results onto the output states marked H and V. The scrambler works as follows:

- if the input state is $|H\rangle$, it will display H with probability $3/4$ and V with probability $1/4$;

- if the input state is $|V\rangle$, it will display V with probability $2/3$ and H with probability $1/3$.

The quantum efficiency is unity and the dark counts are negligible. Find the scrambler matrix of this detector.

Exercise 5.64. Show that, for any scrambler matrix, $\sum_{j=1}^{M} \mu_{ji} = 1$, where M is the total number of the detector's output states.

The number of detector output states may not be equal to the Hilbert space dimension. As an example, let us consider a non-discriminating photon detector (Box 1.2). This detector has two output states: "click" and "no click". On the other hand, the dimension of the Hilbert space associated with these quantum measurements is infinite: it is spanned by photon number states from zero to infinity[5].

Exercise 5.65. A non-discriminating detector has the following properties:

- There are no dark events.
- Each incoming photon generates an avalanche with probability η (the detector's quantum efficiency). If at least one avalanche is present, the detector's circuit produces a "click".

Model this detector as a projective measurement in the photon number basis, followed by a scrambler, and calculate the scrambler matrix.

5.6.2 Positive operator-valued measure (POVM)

The basis $\{|v_i\rangle\}$ of the ideal measurement, combined with the scrambler matrix μ_{ji}, completely describes any detector modeled by Fig. 5.2. However, as on many other occasions we have encountered in this book, quantum theorists prefer a more compact description that we discuss next. For a detector modeled by Fig. 5.2, the set of operators

$$\hat{F}_j = \sum_i \mu_{ji} \Pi_i, \tag{5.36}$$

each associated with the jth output state of the detector, with $\Pi_i = |v_i\rangle\langle v_i|$, is called the *positive-operator valued measure (POVM)* of that detector. A measurement described by a POVM is called a *generalized measurement*.

Exercise 5.66. Show that each element of a POVM is a non-negative Hermitian operator.

Exercise 5.67. Determine the POVMs of the detectors described in

a) Ex. 5.63;

[5] As discussed in Sec. 3.8, the quantum description of an electromagnetic field mode is equivalent to that of the harmonic oscillator.

b) Ex. 5.65.
Answer:

$$\hat{F}_{\text{no click}} = \sum_n (1-\eta)^n |n\rangle\langle n| , \tag{5.37a}$$

$$\hat{F}_{\text{click}} = \sum_n [1-(1-\eta)^n] |n\rangle\langle n| . \tag{5.37b}$$

Exercise 5.68. Show that, for the POVM of a detector modeled by Fig. 5.2,

$$\sum_{j=1}^{M} \hat{F}_j = \hat{\mathbf{1}}, \tag{5.38}$$

where M is the number of POVM elements.

Exercise 5.69. Show the following:

a) When a quantum state $\hat{\rho}$ is measured by a detector described by some POVM $\{\hat{F}_j\}$, the probability of the jth outcome is

$$\text{pr}_j(\hat{\rho}) = \text{Tr}(\hat{F}_j \hat{\rho}) \tag{5.39}$$

(this is the extension of Born's rule to generalized measurements).

b) When Alice's part of a bipartite quantum state $\hat{\rho}_{AB}$ is measured by a detector described by POVM $\{\hat{F}_j\}$ and the jth outcome occurs, the (unnormalized) state of Bob's channel becomes

$$\hat{\rho}_{B,j} = \text{Tr}_A(\hat{F}_j \hat{\rho}_{AB}). \tag{5.40}$$

Exercise 5.70. Alice and Bob share two photons in a mixture of the states $|\Psi_1\rangle = (|HH\rangle + |HV\rangle + 2|VV\rangle)/\sqrt{6}$ with probability $3/5$ and $|\Psi_2\rangle = |HV\rangle$ with probability $2/5$. Alice measures her photon by means of the detector described in Ex. 5.63 and obtains

a) result H;
b) result V;
c) an unknown result.

Find the resulting state of Bob's photon

- using the pure state and projection measurement formalism (express your answer in the form of a statistical ensemble);
- using the density matrix and generalized measurement formalism (express your answer in the form of an unnormalized density matrix).

Check that the answers are mutually consistent.

These results show how useful the POVM is. Comparing Eqs. (5.39) and (5.40) with Eqs. (5.13) and (5.19), we see that in many situations the POVM replaces the set of projectors in the mathematical description of a detector.

However, there is an important caveat. The POVM can fully replace projectors only for measurements that destroy the measured quantum system (as is the case, for example, with traditional photon detectors), or if we are not interested in the state of the system after the measurement. But if the system is not destroyed, its state after a generalized measurement is not equal to $\hat{F}_j\hat{\rho}\hat{F}_j$, in contrast to projective measurements, where the post-measurement state (5.12) is $\hat{\Pi}_j\hat{\rho}\hat{\Pi}_j$. We shall see this in the next exercise.

Exercise 5.71. a) Determine the density operator of the post-measurement state in the case of the jth result of a measurement shown in Fig. 5.2. The answer should be expressed in terms of the scrambler matrix and the projection operators defining the quantum part of the detector.

b) Apply the result of part (a) to the state $\hat{\rho} = |+\rangle\langle+|$ measured by the detector described in Ex. 5.63. Find the post-measurement state for each outcome. Check that they are not equal to $\hat{F}_j\hat{\rho}\hat{F}_j$.

Another difference between generalized measurements and projective measurements is that the former are not repeatable. If we once again subject the state $\hat{\Pi}_j\hat{\rho}\hat{\Pi}_j$, obtained as a result of a projective measurement, to the same measurement, we will obtain $\hat{\Pi}_j\hat{\Pi}_j\hat{\rho}\hat{\Pi}_j\hat{\Pi}_j = \hat{\Pi}_j\hat{\rho}\hat{\Pi}_j$, so the state will not change. But in the case of generalized measurement, the situation is different.

Exercise 5.72. Suppose a photon in the initial state $\hat{\rho} = |+\rangle\langle+|$ is measured in a non-destructive manner by the detector described in Ex. 5.63 and the result H obtains. Apply the same measurement once again to the post-measurement state and find the resulting state as well as the probability of each outcome.

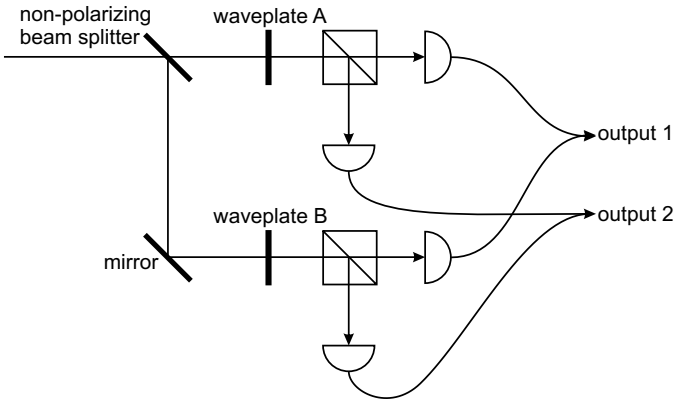


Fig. 5.3 Example of a detector not described by the model of Fig. 5.2. A non-polarizing beam splitter randomly directs a photon to two different, ideal polarization measurement setups. A photon detected in the transmitted channel of either PBS activates the same output state of the detector; a photon in the reflected channel of either PBS activates the other output state.

To finalize the discussion of generalized measurements, let me say that not every physical measurement can be modeled by a projective measurement plus a scram-

bler — an example is shown in Fig. 5.3. However, quite remarkably, *any* detector — that is, any apparatus that provides us with information about a physical system — can be described by a POVM, i.e., a set of non-negative operators whose properties are consistent with Eqs. (5.38), (5.39), and (5.40). We will show how to construct this POVM in the next section, but for now, let us look at an example.

Exercise 5.73. Consider the detector in Fig. 5.3, where the waveplate A is a half-wave plate oriented at 0° (the top polarization detector measures in the canonical basis), and the waveplate B a half-wave plate at 22.5° (the bottom detector measures in the diagonal basis). The non-polarizing beam splitter is symmetric, i.e., it has an equal probability of transmitting and reflecting the photon.

a) Suppose the detector is used to measure an arbitrary state with the density matrix $\hat{\rho} \simeq \begin{pmatrix} \rho_{HH} & \rho_{HV} \\ \rho_{VH} & \rho_{VV} \end{pmatrix}$. Find the probabilities of the two detector outputs in terms of $\rho_{HH}, \rho_{HV}, \rho_{VH}, \rho_{VV}$.
b) Based on Eq. (5.39) and the result of part (a), find the POVM of this detector. Show that the sum of the POVM elements is the identity operator.

A further beautiful result, known as *Neumark's (Naimark's) theorem*, states that, for any set $\{\hat{F}_j\}$ of non-negative Hermitian operators such that $\sum_j \hat{F}_j = \hat{\mathbf{1}}$, one can construct a detector whose POVM is $\{\hat{F}_j\}$. The proof of this statement is beyond the scope of this course, but can be found in textbooks on quantum information theory[6].

Exercise 5.74. A certain detector is described by a POVM $\{\hat{F}_j\}$ such that Eq. (5.39) holds for every physical state $\hat{\rho}$.

a)* Show that each $\hat{F}_j$ is a Hermitian operator.
b) Show that each $\hat{F}_j$ is a non-negative operator.
c) Prove that set $\{\hat{F}_j\}$ complies with Eq. (5.38).

Exercise 5.75. Consider a "detector" that does not provide any information about the state of a quantum system — that is, the probabilities of its outputs are independent of the state of the input quantum system. Show that all elements of the POVM of a such a "detector" are proportional to the identity operator.

5.7 Quantum tomography*

5.7.1 Quantum state tomography

Here we revisit a topic we touched upon briefly in Section 1.4: complete characterization of quantum states by measurements. But we will now utilize the arsenal of

[6] For example, see A. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory*, Springer, 2011.,

tools we have learned in this chapter — the density matrix formalism — to develop tomography of a general quantum state without assuming it to be pure.

As we know, full characterization of a state requires not just multiple measurements on many copies of that state, but also that these measurements be performed in multiple bases. Let us estimate the number of bases necessary for full state tomography in a given Hilbert space.

Exercise 5.76. Consider an arbitrary state $\hat{\rho}$ in a Hilbert space of dimension N.

a) Show that this state can be fully described by $N^2 - 1$ independent real parameters.
b) We perform a projective measurement on multiple copies of $\hat{\rho}$ in a specific basis. Show that the information we learn in this measurement state can be contained within a set of $N - 1$ independent real parameters.

So our goal is to determine $(N^2 - 1)$ numbers, but a measurement in each basis gives us only $N - 1$ numbers. Hence complete state tomography requires acquisition of statistics in a minimum of $(N^2 - 1)/(N - 1) = N + 1$ bases. In practice, the choice of bases is largely dictated by experimental convenience, which means that more bases are sometimes needed. Let us look at two examples.

Exercise 5.77. Redo Ex. 1.15 for density matrices. That is, multiple polarization measurements of photons prepared in the same state $\hat{\rho}$ are done in the canonical, diagonal, and circular bases, and all six corresponding probabilities are determined. Express the four matrix elements of $\hat{\rho}$ through these probabilities.

Exercise 5.78*. Show that full tomography of the polarization state of a photon *pair* can be accomplished by measuring multiple copies of that state in each of the nine bipartite combinations of the canonical, diagonal, and circular bases[7]
Hint: This is a tedious calculation, but it can be simplified by doing it in the right order.

- Start with the bipartite canonical basis: what elements of the density matrix do the measurement statistics in this basis help us determine?
- Let Alice's basis be canonical and Bob's diagonal and subsequently circular. Using the density matrix elements known from the first step, determine four additional elements.
- Now let Bob's basis be canonical and Alice's diagonal and circular. Four more matrix elements can be found.
- The remaining density matrix elements can be evaluated from measurements in the four remaining bipartite bases.

In Ex. 5.77, the Hilbert space dimension is $N = 2$, and the number of bases used is $N + 1 = 3$, consistent with the minimum we found. In Ex. 5.78, on the other hand, $N = 4$, while the number of bases used is 9. This means that we can think of

[7] See the experiment in A. G. White, D. F. V. James, W. J. Munro, and P. G. Kwiat, *Exploring Hilbert space: Accurate characterization of quantum information*, Physical Review A **65**, 012301 (2001).

optimizing our solution by invoking a lower number of bases. We need to be careful, however, that these "optimized" bases are not too difficult to realize in a practical experimental setup.

Another important lesson we learn from Ex. 5.78 is that, although the bipartite Hilbert space contains entangled states, its full tomography does not require measurements in entangled bases. In other words, Alice's and Bob's measurement devices do not need to have any quantum correlation with each other. This is, of course, a great relief for experimentalists.

5.7.2 *Quantum process tomography*

Under a *quantum process* we understand a "black box" that performs some kind of processing on quantum states (Fig. 5.4). For an input state $\hat{\rho}$, the process output state is denoted by $\mathbf{E}(\hat{\rho})$. The goal of *quantum process tomography (QPT)* is to learn enough information about the black box to enable one to predict its effect upon an arbitrary input state. This information is obtained by sending multiple copies of certain *probe states* $\hat{\rho}_j$ into the black box and performing quantum state tomography on the output to find $\mathbf{E}(\hat{\rho}_j)$ for each probe state.

Fig. 5.4 A quantum process.

At the beginning of this course (Sec. 1.10), we learned that quantum evolution processes are represented by unitary linear operators $\hat{U} = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\hat{H}t}$ (where $\hat{H}$ is the Hamiltonian). However, the same is not always true for an arbitrary quantum process, as we shall see in a moment. Still, let us start our discussion of QPT with a black box that is known to be described by some linear operator.

Exercise 5.79. Suppose the process is described by a linear operator $\hat{U}$ and the state $\hat{U}\,|v_i\rangle$ is known for every element of some orthonormal basis $\{|v_i\rangle\}$ of the Hilbert space. Find the density matrix of the process output state $\mathbf{E}(\hat{\rho})$ if the input state density operator $\hat{\rho}$ is given[8].

According to this result, in order to fully characterize a process described by a linear operator, it suffices to probe that process with states from any basis of the Hilbert space.

However, quantum processes are unitary operators only if the system of interest does not interact with the outside world (the "environment"). If such interaction does

[8] Of course, if a quantum process is described by an operator, this operator must be, not just linear, but also unitary (see Sec. 1.10). However, this fact is not relevant to this exercise.

occur, the system and the environment will become entangled. We then have to trace over the environment to determine the final state of the system. This irreversible operation makes the whole process non-unitary.

Consider, for example, decoherence of a spin-$\frac{1}{2}$ particle with the canonical basis being the preferred one. States $|\uparrow\rangle$ and $|\downarrow\rangle$ are not affected by this decoherence: $\mathbf{E}(|\uparrow\rangle\langle\uparrow|) = |\uparrow\rangle\langle\uparrow|$ and $\mathbf{E}(|\downarrow\rangle\langle\downarrow|) = |\downarrow\rangle\langle\downarrow|$. However, any linear combination $|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle$ becomes a statistical mixture: $\mathbf{E}(|\psi\rangle\langle\psi|) = |\alpha|^2|\uparrow\rangle\langle\uparrow| + |\beta|^2|\downarrow\rangle\langle\downarrow|$. If the only information available to us is the effect of the process on the basis states $|\uparrow\rangle$ and $|\downarrow\rangle$, we cannot distinguish this process from the identity process $\mathbf{E}(\hat{\rho}) = \hat{\rho}$.

In view of the above, it may appear that quantum process tomography is an intractable problem. The interaction of systems and environments comes in all shapes and sizes, and since the information about the environment is unavailable, it would seem impossible to determine all properties of the process by just looking at the system. Fortunately, however, this is not the case, as we shall see next.

Exercise 5.80. Show that any process must be linear with respect to density matrices, i.e.,

$$\mathbf{E}(\alpha\hat{\rho}_1 + \beta\hat{\rho}_2) = \alpha\mathbf{E}(\hat{\rho}_1) + \beta\mathbf{E}(\hat{\rho}_2). \tag{5.41}$$

Hint: use the probabilistic nature of the density operator (see Ex. 5.22).

Exercise 5.81. Show that, in the linear space of all linear operators on a Hilbert space of dimension N (see Ex. A.42), one can construct a basis that consists entirely of density operators of physical quantum states.
Hint: consider, for example, the set Q that includes

- N operators $\hat{\rho}_{kk} = |v_k\rangle\langle v_k|$;
- $N(N-1)/2$ operators $\hat{\rho}_{\mathrm{re},kl} = \big|\psi_{\mathrm{re},kl}\big\rangle\big\langle\psi_{\mathrm{re},kl}\big|$ with $\psi_{\mathrm{re},kl} = (|v_k\rangle + |v_l\rangle)/\sqrt{2}$ for each unique pair of indices (k,l);
- $N(N-1)/2$ operators $\hat{\rho}_{\mathrm{im},kl} = \big|\psi_{\mathrm{im},kl}\big\rangle\big\langle\psi_{\mathrm{im},kl}\big|$, with $\psi_{\mathrm{im},kl} = (|v_k\rangle + \mathrm{i}\,|v_l\rangle)/\sqrt{2}$ for each unique pair of indices (k,l),

where $\{|v_k\rangle\}$ is an arbitrary orthonormal basis of the Hilbert space.

Exercise 5.82. Let $\{\hat{\rho}_i\}$ be a basis in the space of *operators* on our Hilbert space with each element corresponding to the density operator of a physical state. Suppose one knows the effect $\mathbf{E}(\hat{\rho}_i)$ of the process on each of these states. Show that the effect of the process on an arbitrary state is given by

$$\mathbf{E}(\hat{\rho}) = \sum_i \lambda_i \mathbf{E}(\hat{\rho}_i), \tag{5.42}$$

where λ_i are the coefficients of the decomposition of the density operator $\hat{\rho}$ into this spanning set:

$$\hat{\rho} = \sum_i \lambda_i \hat{\rho}_i. \tag{5.43}$$

The above exercise provides us with a conceptual framework for quantum process tomography. Any basis[9] $\{\hat{\rho}_i\}$ in the space of operators over the Hilbert space can serve as a set of probe states such that the set of output density matrices $\{\mathbf{E}(\hat{\rho}_i)\}$ comprises complete information about the process. The exercises below provide an example to that effect. These examples specialize to the physics of a spin-$\frac{1}{2}$ particle.

Exercise 5.83. Show that the set of density matrices

$$Q = \big\{\hat{\rho}_\uparrow = |\uparrow\rangle\langle\uparrow|\,,\ \hat{\rho}_\downarrow = |\downarrow\rangle\langle\downarrow|\,,\ \hat{\rho}_+ = |+\rangle\langle+|\,,\ \hat{\rho}_R = |R\rangle\langle R|\,\big\}, \tag{5.44}$$

where $|+\rangle = (|\uparrow\rangle + |\downarrow\rangle)/\sqrt{2}$ and $|R\rangle = (|\uparrow\rangle + i\,|\downarrow\rangle)/\sqrt{2}$ are the eigenstates of $\hat{\sigma}_x$ and $\hat{\sigma}_y$ with the eigenvalue 1, forms a basis in the linear space of all linear operators over the qubit Hilbert space. Express an arbitrary state $\hat{\rho} \simeq \begin{pmatrix} \rho_{\uparrow\uparrow} & \rho_{\uparrow\downarrow} \\ \rho_{\downarrow\uparrow} & \rho_{\downarrow\downarrow} \end{pmatrix}$ as a mixture (5.42) of elements of that basis.

Exercise 5.84. Consider the partial decoherence process studied in Sec. 5.5.1:

$$\mathbf{E}\begin{pmatrix} \rho_{\uparrow\uparrow} & \rho_{\uparrow\downarrow} \\ \rho_{\downarrow\uparrow} & \rho_{\downarrow\downarrow} \end{pmatrix} = \begin{pmatrix} \rho_{\uparrow\uparrow} & \rho_{\uparrow\downarrow}\mathrm{e}^{-t/T_2} \\ \rho_{\downarrow\uparrow}\mathrm{e}^{-t/T_2} & \rho_{\downarrow\downarrow} \end{pmatrix}. \tag{5.45}$$

a) Find the effect $\mathbf{E}(\hat{\rho}_i)$ of this process on all elements of the basis (5.44).
b) Suppose the basis (5.44) is used for quantum process tomography. By expressing an arbitrary state $\hat{\rho} \simeq \begin{pmatrix} \rho_{\uparrow\uparrow} & \rho_{\uparrow\downarrow} \\ \rho_{\downarrow\uparrow} & \rho_{\downarrow\downarrow} \end{pmatrix}$ as a mixture of elements of that basis, check Eq. (5.42) explicitly.

A QPT experiment provides us with a set of density matrices $\{\mathbf{E}(\hat{\rho}_i)\}$. While, as we have shown, this set fully describes the process, it would be good to have a more compact and convenient description — as we had in the case of density operators and POVMs. We thus seek a way to express the information about the process in the form of a *process tensor* — a "supermatrix" $\mathbf{E}^{nm}_{lk}$ that would generate the matrix of the black box output state $\mathbf{E}(\hat{\rho})$ when applied to the matrix input state $\hat{\rho}$:

$$\big[\mathbf{E}(\hat{\rho})\big]_{lk} = \sum_{m=1}^{N}\sum_{n=1}^{N} \mathbf{E}^{nm}_{lk}\rho_{nm}, \tag{5.46}$$

where $\rho_{nm} = \langle v_n|\,\hat{\rho}\,|\,v_m\rangle$ and $[\mathbf{E}(\hat{\rho})]_{lk} = \langle v_l|\,\mathbf{E}(\hat{\rho})\,|\,v_k\rangle$, with $\{|v_j\rangle\}$ an orthonormal basis in $\mathbb{V}$.

Equation (5.46) is reminiscent of matrix multiplication (A.20), except that the summation is over two indices. Both the input and output objects are matrices, they have two indices. And the process tensor $\mathbf{E}^{nm}_{lk}$ that transforms one into the other has four indices — it is a *rank-4 tensor*, an $N\times N\times N\times N$ table of numbers that is easy to handle, store, and communicate.

But does the process tensor exist for every quantum process, and if so, how can it be found? The answer turns out to be relatively simple.

[9] In fact, it suffices for the set $\{\hat{\rho}_i\}$ to be a spanning set; it need not be linearly independent.

Exercise 5.85. Consider a certain orthonormal basis $\{|v_n\rangle\}$ of the Hilbert space. Let $\{\hat{\rho}_i\}$ (with $i = 1,\ldots,N^2$) be a set of QPT probe states, i.e., a spanning set in the space of density matrices. Then each operator $|v_m\rangle\langle v_n|$ can be decomposed into this spanning set, viz.,

$$|v_n\rangle\langle v_m| = \sum_{i=1}^{N^2} \lambda_{nmi}\hat{\rho}_i, \tag{5.47}$$

where λ_{nmi} are the decomposition coefficients. Show that Eq. (5.46) is satisfied if the process tensor is given by

$$\mathbf{E}_{lk}^{nm} = \sum_{i=1}^{N^2} \lambda_{nmi} \langle v_l| \mathbf{E}(\hat{\rho}_i)| v_k\rangle . \tag{5.48}$$

Exercise 5.86. Find the coefficients of decomposition (5.47) when $\{|v_n\rangle\}$ is the canonical basis in the qubit space and the basis $\{\hat{\rho}_i\}$ is given by Eq. (5.44).

Exercise 5.87. Use Eq. (5.48) with the result of Ex. 5.84(a) and 5.86 to find the tensor of the partial decoherence process (5.45). Verify that this tensor, when substituted into Eq. (5.46), yields Eq. (5.45).

Answer:

$$\mathbf{E}_{lk}^{nm} = \begin{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} 0 & \mathrm{e}^{-t/T_2} \\ 0 & 0 \end{pmatrix} \\ \begin{pmatrix} 0 & 0 \\ \mathrm{e}^{-t/T_2} & 0 \end{pmatrix} & \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \end{pmatrix}, \tag{5.49}$$

where each pair (n,m) identifies a 2×2 submatrix, while the indices inside each submatrix are (l,k).

The above result illustrates the meaning of the process tensor. The submatrix in the nth row and mth column on the right-hand side of Eq. (5.49) gives the process output $\mathbf{E}(|v_n\rangle\langle v_m|)$ corresponding to the input "state" $|v_n\rangle\langle v_m|$[10]. For example, the input state $|\uparrow\rangle\langle\uparrow| \simeq \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ is not affected by decoherence, so the upper left submatrix is the same as the input: $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. On the other hand, if the input "state" is $|\uparrow\rangle\langle\downarrow| \simeq \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, the decohered output (upper right submatrix) is $\begin{pmatrix} 0 & \mathrm{e}^{-t/T_2} \\ 0 & 0 \end{pmatrix}$, and so on. The mathematics behind this observation can be seen from Eq. (5.46): if we set $\hat{\rho} = |v_n\rangle\langle v_m|$, then $\big[\mathbf{E}(\hat{\rho})\big]_{lk} = \mathbf{E}_{lk}^{nm}$.

As we see, the theoretical treatment of QPT can be complicated and tedious, and its practical realization even more so. To form a basis in the space of operators over the Hilbert space, the set of probe states must contain N^2 elements. For each of these elements, full tomography of the output state $\mathbf{E}(\hat{\rho}_i)$ needs to be performed and a set of N^2-1 parameters defining its density matrix determined. So the total number of

[10] For $m \neq n$, these are only formal mathematical objects that do not correspond to physical states. However, they are handy for developing an intuition.

parameters to be obtained in QPT scales as the fourth power of the Hilbert space dimension, which means that the experimentalist will have to spend many hours in the lab. To make things worse, the required probe basis may contain complex superposition states that are difficult to prepare or may even be beyond the reach of existing quantum state engineering methods.

5.7.3 Quantum detector tomography

Quantum detector tomography can be seen as a simplified case of QPT. Here, instead of a black box with a quantum output, we have a detector — a black box with M possible classical output states. The goal is once again to be able to predict the detector response to an arbitrary state, or determine the detector's POVM, by studying its response to certain probe states.

Exercise 5.88. A certain detector, when measuring the states $\hat{\rho}_{1,2}$, produces the output j with probabilities $\mathrm{pr}_j(\hat{\rho}_{1,2})$, respectively. Show that, when measuring a linear mixture $\alpha\hat{\rho}_1 + \beta\hat{\rho}_2$, the probability of the output j is given by

$$\mathrm{pr}_j(\alpha\hat{\rho}_1 + \beta\hat{\rho}_2) = \alpha\mathrm{pr}_j(\hat{\rho}_1) + \beta\mathrm{pr}_j(\hat{\rho}_2). \tag{5.50}$$

Exercise 5.89. Suppose $\{\hat{\rho}_i\}$ is the basis (or a spanning set) defined in Ex. 5.82. For each of its elements, we have measured the complete statistics of the detector response, i.e., $\mathrm{pr}_j(\hat{\rho}_i)$, where j indexes the detector's output states. From these data, determine $\mathrm{pr}_j(\hat{\rho})$ for an arbitrary input density matrix $\hat{\rho}$ whose decomposition into $\{\hat{\rho}_i\}$ is given by Eq. (5.43).

Exercise 5.90.* Under the conditions of the previous exercise, show that Eq. (5.39) is satisfied if the POVM of the detector is given by

$$\hat{F}_j = \sum_{i=1}^{N^2} \sum_{m,n=1}^{N} \lambda_{nmi}\mathrm{pr}_j(\rho_i)\,|v_m\rangle\langle v_n|\,, \tag{5.51}$$

where λ_{nmi} are the coefficients of the decomposition of the operator $|v_n\rangle\langle v_m|$ into the probe state basis, according to Eq. (5.47).

Exercise 5.91. Consider the detector shown in Fig. 5.3, set up in the same way as in Ex. 5.73.

a) Find the detector output probabilities for the four states from the set $Q = \big\{\hat{\rho}_\uparrow = |H\rangle\langle H|\,,\ \hat{\rho}_\downarrow = |V\rangle\langle V|\,,\ \hat{\rho}_+ = |+\rangle\langle +|\,,\ \hat{\rho}_R = |R\rangle\langle R|\big\}$.
b) Use this information and Eq. (5.51) to find the POVM of the detector. Check that the result is the same as in Ex. 5.73.

As we can see from the last exercise, we now have an algorithm to calculate a detector's POVM, not only from the experimental data obtained by measuring probe states, but also theoretically from a physical model of the detector.

5.8 Problems

Problem 5.1. Find the representation of the density operator of states $|\alpha\rangle + |-\alpha\rangle$ and $|\alpha\rangle\langle\alpha| - |-\alpha\rangle\langle-\alpha|$ of a harmonic oscillator

a) in the Fock basis;
b) in the position basis;
c) in the momentum basis,

where α and $-\alpha$ are coherent states. Discuss the behavior of diagonal and off-diagonal elements in the context of Ex. 5.12. Normalization can be neglected.

Problem 5.2. Consider a photon in the ensemble of states

- $|\psi_1\rangle = (3|H\rangle - 4|V\rangle)/5$ with probability $p_1 = 1/2$;
- $|\psi_2\rangle = (12|H\rangle - 5\mathrm{i}|V\rangle)/13$ with probability $p_2 = 1/4$;
- $|\psi_3\rangle = |-45^\circ\rangle$ with probability $p_3 = 1/4$.

a) Find the density operator.
b) This ensemble is measured in the circular basis. Find the probabilities of each result using the verbal description above and using the density matrix formalism. Check for consistency.

The answers should be in numerical form, up to the third decimal point.

Problem 5.3. The density matrix of a photon state in the canonical basis is

$$\hat{\rho} = \begin{pmatrix} 1/2 & \mathrm{i}/6 \\ -\mathrm{i}/6 & 1/2 \end{pmatrix}.$$

Present this state as a statistical mixture of orthogonal pure states.

Problem 5.4. Alice and Bob share two photons in the state $|\Psi\rangle = (|HV\rangle + |VH\rangle + 2|VV\rangle)/\sqrt{6}$. Alice measures the state in the canonical basis.

a) What state will be prepared at Bob's station in each case?
b) What is the probability of each outcome?
c) Suppose Bob does not know Alice's measurement outcome. Use the results of parts (a) and (b) to write the statistical ensemble describing the state of Bob's photon. Find the corresponding density matrix in the canonical basis.
d) Find the reduced density matrix of Bob's photon using the density matrix formalism. Check that the result is the same as in part (c).
e) Repeat parts (a)–(c) for Alice performing her measurement in the diagonal basis. Check that the reduced density matrix of Bob's photon is the same.

Problem 5.5. Alice and Bob share two photons in a polarization state whose matrix in the canonical basis $\{|HH\rangle, |HV\rangle, |VH\rangle, |VV\rangle\}$ is

$$\hat{\rho}_{AB} = \frac{1}{18}\begin{pmatrix} 3 & 1 & -2 & \mathrm{i} \\ 1 & 1 & 2\mathrm{i} & -3 \\ -2 & -2\mathrm{i} & 4 & 0 \\ -\mathrm{i} & -3 & 0 & 10 \end{pmatrix}$$

.

a) Write the density matrix $\hat{\rho}_b$ of Bob's photon if he has no communication with Alice.
b) Alice measures the polarization of her photon in the canonical basis. What is the probability of each outcome and what state will be prepared at Bob's station in each case?
c) Alice measures her photon using the detector described in Ex. 5.63. What is the probability of each outcome and what state will be prepared at Bob's station in each case?

Problem 5.6. An ensemble of spin-$1/2$ particles initially in the state $|\uparrow\rangle$ undergoes decoherence due to collisions with a buffer gas. Each collision results in complete decoherence of the particle that experienced it. The decoherence preferred basis is $\{|\pm\rangle\} = \{(|\uparrow\rangle \pm |\downarrow\rangle)/\sqrt{2}\}$. The probability of collision per particle per unit time is p. Write the density matrix as a function of time

a) in the decoherence preferred basis;
b) in the canonical basis.

Problem 5.7. Redo Ex. 5.25 for a mixture of the states that corresponds to the spin pointing along the x and y axes with probabilities $\frac{1}{3}$ and $\frac{2}{3}$, respectively. The magnetic field B is along the z axis.

Problem 5.8. Two electrons, whose spins are initially in the state $|\Psi(0)\rangle = |\rightarrow\rangle \otimes |\uparrow\rangle$ (where $|\rightarrow\rangle$ is the eigenstate of $\hat{S}_x$ with eigenvalue $\frac{\hbar}{2}$), associated with fictitious observers Alice and Bob, are interacting with the Hamiltonian $\hat{H} = C\vec{S}_1 \cdot \vec{S}_2$.

a) Find the evolution $|\Psi(t)\rangle$ of the electrons' spin state in the canonical basis.
b) Alice measures the projection of her electron's spin onto the z axis at time t. Find the probabilities of the possible results and the state in which Bob's electron will be prepared in each case. Based on that information, determine the ensemble describing the state of Bob's electron when he does not know the result of Alice's measurement. From that description, obtain the density matrix of Bob's electron in the canonical basis.
c) Repeat part (b) when Alice measures the projection of her electron's spin onto the x axis.
d) Find the density operator $\hat{\rho}_B(t)$ of Bob's electron as a function of time using the partial trace formalism. Check that your result is identical to what was found in parts (b) and (c).
e) Find the Bloch vector trajectory of Bob's electron spin and plot it.
f) Find the state purity of Bob's electron spin as a function of time. Check how it is related to the length of the Bloch vector: $\mathrm{Tr}\hat{\rho}^2 = (|\vec{R}|^2 + 1)/2$.

Problem 5.9. For the two-mode squeezed state (3.186a), calculate the density matrix of Bob's portion.
Hint: to calculate the partial trace in the continuous-variable setting, replace the summation in Eq. (5.18) by integration.

Problem 5.10. Find the process tensor of the homogeneous relaxation process that has both a longitudinal component (T_1) and transverse components (T_2).

Problem 5.11. Analyze the following technique for measuring the longitudinal relaxation time.

- A $\frac{\pi}{2}$ excitation pulse is applied to a thermalized spin ensemble to make the Bloch vector point along the y axis.
- As time elapses, Bloch vectors of different spins will spread over the equator due to inhomogeneous dephasing. At the same time, they will experience longitudinal and transverse relaxation. The longitudinal relaxation will give rise to the z component of the mean Bloch vector.
- After time $t_0 \gg T_2^*$ has elapsed, another $\frac{\pi}{2}$ pulse is applied. The emerging z component of the Bloch vector is now pointing along the y axis and can produce free induction decay.

Calculate the mean magnetic moment of the spin after the second excitation pulse as a function of the time t, the excitation pulse separation t_0, and the longitudinal and transverse time constants of the sample.

Problem 5.12. Calculate the POVM of the non-discriminating detector described in Ex. 5.65 taking dark counts into account. A dark avalanche occurs with probability p_{dark}, independently of other avalanches that may be occurring in the detector at the same time.

Problem 5.13. Consider the polarization detector described in Ex. 5.63, taking quantum efficiency $\eta = 0.8$ into account. In the event that no avalanche is produced in any of the photon detectors in response to an incoming photon, the detector displays "0".

a) Calculate the POVM.
b) Find the probability of each outcome for the input state $\alpha|H\rangle + \beta|V\rangle$.

Problem 5.14. Consider a two-mode optical state

$$|\psi\rangle = \sum_{k,m=0}^{\infty} \psi_{km} |k\rangle_A \otimes |m\rangle_B ,$$

where subscripts A and B denote the modes and the state is written in the Fock basis (e.g., the state $|1\rangle_A \otimes |0\rangle_B$ corresponds to one photon in mode A and the vacuum in mode B).

a) Mode B is discarded. What is the density operator of the state in mode A?

b) Mode B is subjected to a measurement by a non-discriminating single-photon detector with the quantum efficiency η described in Ex. 5.65. What is the density operator of the state in mode A in the event of a "click"?

c) Repeat (b) for the case when the initial state is not pure, but described by the density matrix

$$\hat{\rho} = \sum_{k,l,m,n=0}^{\infty} \rho_{klmn} \, |k\rangle\langle l|_A \otimes |m\rangle\langle n|_B \, .$$

Problem 5.15. A polarization measurement device consisting of a PBS and two perfect photon detectors contains a "gremlin" who, with probability $1/2$, inserts a half-wave plate with its optic axis oriented at $\pi/4$ before the PBS. Find the POVM of that detector.

Problem 5.16. A quantum process $\mathbf{E}$ on a polarization qubit has been subjected to a quantum process tomography experiment. It has revealed the following transformations of the probe states:

$$\begin{aligned}
&|H\rangle \to 1/4\,|H\rangle\langle H| + 3/4\,|V\rangle\langle V|\,;\\
&|V\rangle \to 3/4\,|H\rangle\langle H| + 1/4\,|V\rangle\langle V|\,;\\
&|+\rangle \to |+\rangle\langle +|\,;\\
&|R\rangle \to 1/2\,|H\rangle\langle H| + 1/2\,|V\rangle\langle V| + i/4\,|H\rangle\langle V| - i/4\,|V\rangle\langle H|\,.
\end{aligned}$$

a) Find the process tensor $\mathbf{E}_{lk}^{nm}$ such that

$$[\mathbf{E}(\hat{\rho})]_{lk} = \sum_{nm} \mathbf{E}_{lk}^{nm} \rho_{nm}.$$

b) How will the process transform the states $|-\rangle$, $|L\rangle$, $p\,|H\rangle\langle H| + (1-p)\,|-\rangle\langle -|$?

c) This process can be described as decoherence in a certain preferred basis. What is that basis?

Appendix A
Linear algebra basics

A.1 Linear spaces

Linear spaces consist of elements called *vectors*. Vectors are abstract mathematical objects, but, as the name suggests, they can be visualized as geometric vectors. Like regular numbers, vectors can be added together and subtracted from each other to form new vectors; they can also be multiplied by numbers. However, vectors cannot be multiplied or divided by one another as numbers can.

One important peculiarity of the linear algebra used in quantum mechanics is the so-called *Dirac notation* for vectors. To denote vectors, instead of writing, for example, $\vec{a}$, we write $|a\rangle$. We shall see later how convenient this notation turns out to be.

Definition A.1. A *linear (vector) space* $\mathbb{V}$ over a field[1] $\mathbb{F}$ is a set in which the following operations are defined:

1. Addition: for any two vectors $|a\rangle, |b\rangle \in \mathbb{V}$, there exists a unique vector in $\mathbb{V}$ called their sum, denoted by $|a\rangle + |b\rangle$.
2. Multiplication by a number ("scalar"): For any vector $|a\rangle \in \mathbb{V}$ and any number $\lambda \in \mathbb{F}$, there exists a unique vector in $\mathbb{V}$ called their product, denoted by $\lambda\,|a\rangle \equiv |a\rangle\,\lambda$.

These operations obey the following *axioms*.

1. Commutativity of addition: $|a\rangle + |b\rangle = |b\rangle + |a\rangle$.
2. Associativity of addition: $(|a\rangle + |b\rangle) + |c\rangle = |a\rangle + (|b\rangle + |c\rangle)$.
3. Existence of zero: there exists an element of $\mathbb{V}$ called $|\text{zero}\rangle$ such that, for any vector $|a\rangle$, $|a\rangle + |\text{zero}\rangle = |a\rangle$.[2]

A solutions manual for this appendix is available for download at https://www.springer.com/gp/book/9783662565827

[1] *Field* is a term from algebra which means a complete set of numbers. The sets of rational numbers $\mathbb{Q}$, real numbers $\mathbb{R}$, and complex numbers $\mathbb{C}$ are examples of fields. Quantum mechanics usually deals with vector spaces over the field of complex numbers.

[2] As an alternative notation for $|\text{zero}\rangle$, we some times use "0" but *not* "$|0\rangle$".

A. I. Lvovsky, *Quantum Physics*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-662-56584-1

4. Existence of the opposite element: For any vector $|a\rangle$ there exists another vector, denoted by $-|a\rangle$, such that $|a\rangle + (-|a\rangle) = |\text{zero}\rangle$.
5. Distributivity of vector sums: $\lambda(|a\rangle + |b\rangle) = \lambda|a\rangle + \lambda|b\rangle$.
6. Distributivity of scalar sums: $(\lambda + \mu)|a\rangle = \lambda|a\rangle + \mu|a\rangle$.
7. Associativity of scalar multiplication: $\lambda(\mu|a\rangle) = (\lambda\mu)|a\rangle$.
8. Scalar multiplication identity: For any vector $|a\rangle$ and number $1 \in \mathbb{F}$, $1 \cdot |a\rangle = |a\rangle$.

Definition A.2. *Subtraction* of vectors in a linear space is defined as follows:

$$|a\rangle - |b\rangle \equiv |a\rangle + (-|b\rangle).$$

Exercise A.1. Which of the following are linear spaces (over the field of complex numbers, unless otherwise indicated)?

a) $\mathbb{R}$ over $\mathbb{R}$? $\mathbb{R}$ over $\mathbb{C}$? $\mathbb{C}$ over $\mathbb{R}$? $\mathbb{C}$ over $\mathbb{C}$?
b) Polynomial functions? Polynomial functions of degree $\leq n$? $> n$?
c) All functions such that $f(1) = 0$? $f(1) = 1$?
d) All periodic functions of period T?
e) N-dimensional geometric vectors over $\mathbb{R}$?

Exercise A.2. Prove the following:

a) there is only one zero in a linear space;
b) if $|a\rangle + |x\rangle = |a\rangle$ for some $|a\rangle \in \mathbb{V}$, then $|x\rangle = |\text{zero}\rangle$;
c) for any vector $|a\rangle$ and for number $0 \in \mathbb{F}$, $0|a\rangle = |\text{zero}\rangle$;
d) $-|a\rangle = (-1)|a\rangle$;
e) $-|\text{zero}\rangle = |\text{zero}\rangle$;
f) for any $|a\rangle$, $-|a\rangle$ is unique;
g) $-(-|a\rangle) = |a\rangle$;
h) $|a\rangle = |b\rangle$ if and only if $|a\rangle - |b\rangle = 0$.

Hint: Most of these propositions can be proved by adding the same number to the two sides of an equality.

A.2 Basis and dimension

Definition A.3. A set of vectors $|v_i\rangle$ is said to be *linearly independent* if no nontrivial[2] *linear combination* $\lambda_1|v_1\rangle + \ldots + \lambda_N|v_N\rangle$ equals $|\text{zero}\rangle$.

Exercise A.3. Show that a set of vectors $\{|v_i\rangle\}$ is *not* linearly independent if and only if one of the $|v_i\rangle$ can be represented as a linear combination of others.

Exercise A.4. For linear spaces of geometric vectors, show the following:

a) For the space of vectors in a plane (denoted $\mathbb{R}^3$), any two vectors are linearly independent if and only if they are not parallel. Any set of three vectors is linearly dependent.

[3] That is, in which at least one of the coefficients is nonzero.

b) For the space of vectors in a three-dimensional space (denoted $\mathbb{R}^4$), any three non-coplanar vectors form a linearly independent set.

Hint: Recall that a geometric vector can be defined by its x, y and z components.

Definition A.4. A subset $\{|v_i\rangle\}$ of a vector space $\mathbb{V}$ is said to *span* $\mathbb{V}$ (or to be a *spanning set* for $\mathbb{V}$) if any vector in $\mathbb{V}$ can be expressed as a linear combination of the $|v_i\rangle$.

Exercise A.5. For the linear space of geometric vectors in a plane, show that any set of at least two vectors, of which at least two are non-parallel, forms a spanning set.

Definition A.5. A *basis* of $\mathbb{V}$ is any linearly independent spanning set. A *decomposition* of a vector relative to a basis is its expression as a linear combination of the basis elements.

The basis is a smallest subset of a linear space such that all other vectors can be expressed as a linear combination of the basis elements. The term "basis" may suggest that each linear space has only one basis — just as a building can have only one foundation. Actually, as we shall see, in any nontrivial linear space, there are infinitely many bases.

Definition A.6. The number of elements in a basis is called the *dimension* of $\mathbb{V}$. Notation: $\dim\mathbb{V}$.

Exercise A.6* Prove that in a finite-dimensional space, all bases have the same number of elements.

Exercise A.7. Using the result of Ex. A.6, show that, in a finite-dimensional space,

a) any linearly independent set of $N = \dim\mathbb{V}$ vectors forms a basis;
b) any spanning set of $N = \dim\mathbb{V}$ vectors forms a basis.

Exercise A.8. Show that, for any element of $\mathbb{V}$, there exists only one decomposition into basis vectors.

Definition A.7. For a decomposition of the vector $|a\rangle$ into basis $\{|v_i\rangle\}$, viz.,

$$|a\rangle = \sum_i a_i |v_i\rangle\,, \tag{A.1}$$

we may use the notation

$$|a\rangle \simeq \begin{pmatrix} a_1 \\ \vdots \\ a_N \end{pmatrix}. \tag{A.2}$$

This is called the *matrix form* of a vector, in contrast to the Dirac form (A.1). The scalars a_i are called the *coefficients* or *amplitudes* of the decomposition[3].

[4] We use the symbol $\simeq$ instead of $=$ when expressing vectors and operators in matrix form, e.g., in Eq. (A.2). This is to emphasize the difference: the left-hand side, a vector, is an abstract object

Exercise A.9. Let $|a\rangle$ be one of the elements, $|v_k\rangle$, of the basis $\{|v_i\rangle\}$. Find the matrix form of the decomposition of $|a\rangle$ into this basis.

Exercise A.10. Consider the linear space of two-dimensional geometric vectors. Such vectors are usually defined by two numbers (x,y), which correspond to their x and y components, respectively. Does this notation correspond to a decomposition into any basis? If so, which one?

Exercise A.11. Show the following:

a) For the linear space of geometric vectors in a plane, any two non-parallel vectors form a basis.
b) For the linear space of geometric vectors in a three-dimensional space, any three non-coplanar vectors form a basis.

Exercise A.12. Consider the linear space of two-dimensional geometric vectors. The vectors $\vec{a}, \vec{b}, \vec{c}, \vec{d}$ are oriented with respect to the x axis at angles $0, 45^\circ, 90^\circ, 180^\circ$ and have lengths $2, 1, 3, 1$, respectively. Do the pairs $\{\vec{a},\vec{c}\}$, $\{\vec{b},\vec{d}\}$, $\{\vec{a},\vec{d}\}$ form bases? Find the decompositions of the vector $\vec{b}$ in each of these bases. Express them in the matrix form.

Definition A.8. A subset of a linear space $\mathbb{V}$ that is a linear space on its own is called a *subspace* of $\mathbb{V}$.

Exercise A.13. In an arbitrary basis $\{|v_i\rangle\}$ in the linear space $\mathbb{V}$, a subset of elements is taken. Show that a set of vectors that are spanned by this subset is a subspace of $\mathbb{V}$.

For example, in the space of three-dimensional geometric vectors, any set of vectors within a particular plane or any set of vectors collinear to a given straight line form a subspace.

A.3 Inner Product

Although vectors cannot be multiplied together in the same way that numbers can, one can define a multiplication operation that maps any pair of vectors onto a number. This operation generalizes the scalar product that is familiar from geometry.

Definition A.9. For any two vectors $|a\rangle,\ |b\rangle \in \mathbb{V}$ we define an *inner (scalar) product*[5] — a number $\langle a|\ b\rangle \in \mathbb{C}$ such that:

1. For any three vectors $|a\rangle, |b\rangle, |c\rangle,\ \langle a|\,(|b\rangle + |c\rangle) = \langle a|\ b\rangle + \langle a|\ c\rangle$.
2. For any two vectors $|a\rangle, |b\rangle$ and number $\lambda,\ \langle a|\,(\lambda\,|b\rangle) = \lambda\,\langle a|\ b\rangle$.

---

and is basis-independent, while the right-hand side is a set of numbers and depends on the choice of basis $\{|v_i\rangle\}$. However, in the literature, the equality sign is generally used for simplicity.

[5] The inner product of two vectors is sometimes called the *overlap* in the context of quantum physics.

3. For any two vectors $|a\rangle, |b\rangle$, $\langle a|\, b\rangle = \langle b|\, a\rangle^*$.
4. For any $|a\rangle$, $\langle a|\, a\rangle$ is a nonnegative real number, and $\langle a|\, a\rangle = 0$ if and only if $|a\rangle = 0$.

Exercise A.14. In geometry, the scalar product of two vectors $\vec{a} = (x_a, y_a)$ and $\vec{b} = (x_b, y_b)$ (where all components are real) is defined as $\vec{a} \cdot \vec{b} = x_a x_b + y_a y_b$. Show that this definition has all the properties listed above.

Exercise A.15. Suppose a vector $|x\rangle$ is written as a linear combination of some vectors $|a_i\rangle$: $|x\rangle = \sum_i \lambda_i |a_i\rangle$. For any other vector $|b\rangle$, show that $\langle b|\, x\rangle = \sum_i \lambda_i \langle b|\, a_i\rangle$ and $\langle x|\, b\rangle = \sum_i \lambda_i^* \langle a_i|\, b\rangle$.

Exercise A.16. For any vector $|a\rangle$, show that $\langle \text{zero}|\, a\rangle = \langle a|\, \text{zero}\rangle = 0$.

Definition A.10. $|a\rangle$ and $|b\rangle$ are said to be *orthogonal* if $\langle a|\, b\rangle = 0$.

Exercise A.17. Prove that a set of nonzero mutually orthogonal vectors is linearly independent.

Definition A.11. $\| |a\rangle \| = \sqrt{\langle a|\, a\rangle}$ is called the *norm (length)* of a vector. Vectors of norm 1 are said to be *normalized*. For a given vector $|a\rangle$, the quantity $\mathcal{N} = 1/\| |a\rangle \|$ (such that the vector $\mathcal{N} |a\rangle$ is normalized) is called the *normalization factor*.

Exercise A.18. Show that multiplying a vector by a *phase factor* $\mathrm{e}^{\mathrm{i}\phi}$, where ϕ is a real number, does not change its norm.

Definition A.12. A linear space in which an inner product is defined is called a *Hilbert space*.

A.4 Orthonormal Basis

Definition A.13. An *orthonormal basis* $\{|v_i\rangle\}$ is a basis whose elements are mutually orthogonal and have norm 1, i.e.,

$$\langle v_i|\, v_j\rangle = \delta_{ij}, \tag{A.3}$$

where δ_{ij} is the Kronecker symbol.

Exercise A.19. Show that any orthonormal set of N (where $N = \dim \mathbb{V}$) vectors forms a basis.

Exercise A.20. Show that, if $\begin{pmatrix} a_1 \\ \vdots \\ a_N \end{pmatrix}$ and $\begin{pmatrix} b_1 \\ \vdots \\ b_N \end{pmatrix}$ are the decompositions of vectors $|a\rangle$ and $|b\rangle$ in an orthonormal basis, their inner product can be written in the form

$$\langle a|\, b\rangle = a_1^* b_1 + \ldots + a_N^* b_N. \tag{A.4}$$

Equation (A.4) can be expressed in matrix form using the "row-times-column" rule:

$$\langle a|\, b\rangle = \begin{pmatrix} a_1^* & \dots & a_N^* \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_N \end{pmatrix}. \tag{A.5}$$

One context where we can use the above equations for calculating the inner product is ordinary spatial geometry. As we found in Ex. A.10, the coordinates of geometric vectors correspond to their decomposition into orthogonal basis $\{\hat{i}, \hat{j}\}$, so not surprisingly, their scalar products are given by Eq. (A.4).

Suppose we calculate the inner product of the same pair of vectors using Eq. (A.5) in two different bases. Then the right-hand side of that equation will contain different numbers, so it may seem that the inner product will also depend on the basis chosen. This is not the case, however: according to Defn. A.9, the inner product is defined for a pair of vectors, and is basis-independent.

Exercise A.21. Show that the amplitudes of the decomposition $\begin{pmatrix} a_1 \\ \vdots \\ a_N \end{pmatrix}$ of a vector $|a\rangle$ into an orthonormal basis can be found as follows:

$$a_i = \langle v_i|\, a\rangle. \tag{A.6}$$

In other words [see Eq. (A.1)],

$$|a\rangle = \sum_i \langle v_i|\, a\rangle\, |v_i\rangle. \tag{A.7}$$

Exercise A.22. Consider two vectors in a two-dimensional Hilbert space, $|\psi\rangle = 4|v_1\rangle + 5|v_2\rangle$ and $|\phi\rangle = -2|v_1\rangle + 3\mathrm{i}|v_2\rangle$, where $\{|v_1\rangle, |v_2\rangle\}$ is an orthonormal basis.

a) Show that the set $\{|w_1\rangle = (|v_1\rangle + \mathrm{i}|v_2\rangle)/\sqrt{2}, |w_2\rangle = (|v_1\rangle - \mathrm{i}|v_2\rangle)/\sqrt{2}\}$ is also an orthonormal basis.
b) Find the matrices of vectors $|\psi\rangle$ and $|\phi\rangle$ in both bases.
c) Calculate the inner product of these vectors in both bases using Eq. (A.5). Show that they are the same.

Exercise A.23. Show that, if $|a\rangle$ is a normalized vector and $\{a_i = \langle v_i|\, a\rangle\}$ is its decomposition in an orthonormal basis $\{|v_i\rangle\}$, then

$$\sum_i |a_i|^2 = 1. \tag{A.8}$$

Exercise A.24. Suppose $\{|w_i\rangle\}$ is some basis in $\mathbb{V}$. It can be used to find an orthonormal basis $\{|v_i\rangle\}$ by applying the following equation in sequence to each basis element:

$$|v_{k+1}\rangle = \mathcal{N}\left[|w_{k+1}\rangle - \sum_{i=1}^{k} \langle v_i|\, w_{k+1}\rangle\, |v_i\rangle\right], \tag{A.9}$$

where $\mathcal{N}$ is the normalization factor. This is called the *Gram-Schmidt procedure*.

Exercise A.25*. For a normalized vector $|\psi\rangle$ in an N-dimensional Hilbert space, and any natural number $m \leq N$, show that it is possible to find a basis $\{|v_i\rangle\}$ such that $|\psi\rangle = 1/\sqrt{m}\sum_{i=1}^{m}|v_i\rangle$.

Exercise A.26*. Prove the *Cauchy-Schwarz inequality* for any two vectors $|a\rangle$ and $|b\rangle$:

$$|\langle a|\, b\rangle| \leq \||a\rangle\| \times \||b\rangle\|. \tag{A.10}$$

Show that the inequality is saturated (i.e., becomes an equality) if and only if the vectors $|a\rangle$ and $|b\rangle$ are collinear (i.e., $|a\rangle = \lambda\,|b\rangle$).
Hint: Use the fact that $\||a\rangle - \lambda\,|b\rangle\|^2 \geq 0$ for any complex number λ.

Exercise A.27. Prove the *triangle inequality* for any two vectors $|a\rangle$ and $|b\rangle$:

$$\|(|a\rangle + |b\rangle)\| \leq \||a\rangle\| + \||b\rangle\|. \tag{A.11}$$

A.5 Adjoint Space

The scalar product $\langle a|\, b\rangle$ can be calculated as a matrix product (A.5) of a row and a column. While the column $\begin{pmatrix} b_1 \\ \vdots \\ b_N \end{pmatrix}$ corresponds directly to the vector $|b\rangle$, the row $\begin{pmatrix} a_1^* & \dots & a_N^* \end{pmatrix}$ is obtained from the column corresponding to vector $|a\rangle$ by transposition and complex conjugation. Let us introduce a convention associating this row with the vector $\langle a|$, which we call the *adjoint* of $|a\rangle$.

Definition A.14. For the Hilbert space $\mathbb{V}$, we define the *adjoint space* $\mathbb{V}^\dagger$ (read "V-dagger"), which is in one-to-one correspondence with $\mathbb{V}$, in the following way: for each vector $|a\rangle \in \mathbb{V}$, there is one and only one *adjoint* vector $\langle a| \in \mathbb{V}^\dagger$ with the property

$$\text{Adjoint}(\lambda\,|a\rangle + \mu\,|b\rangle) = \lambda^*\,\langle a| + \mu^*\,\langle b|\,. \tag{A.12}$$

Exercise A.28. Show that $\mathbb{V}^\dagger$ is a linear space.

Exercise A.29. Show that if $\{|v_i\rangle\}$ is a basis in $\mathbb{V}$, $\{\langle v_i|\}$ is a basis in $\mathbb{V}^\dagger$, and if the vector $|a\rangle$ is decomposed into $\{|v_i\rangle\}$ as $|a\rangle = \sum a_i\,|v_i\rangle$, the decomposition of its adjoint is

$$\langle a| = \sum a_i^*\,\langle v_i|\,. \tag{A.13}$$

Exercise A.30. Find the matrix form of the vector adjoint to $|v_1\rangle + \mathrm{i}\,|v_2\rangle$ in the basis $\{\langle v_1|, \langle v_2|\}$.

"Direct" and adjoint vectors are sometimes called *ket* and *bra* vectors, respectively. The rationale behind this terminology, introduced by P. Dirac together with

the symbols $\langle|$ and $|\rangle$, is that the bra-ket combination of the form $\langle a|\, b\rangle$, a "bracket", gives the inner product of the two vectors.

Note that $\mathbb{V}$ and $\mathbb{V}^\dagger$ are different linear spaces. We cannot add a bra-vector and a ket-vector.

A.6 Linear Operators

A.6.1 Operations with linear operators

Definition A.15. A *linear operator* $\hat{A}$ on a linear space $\mathbb{V}$ is a map[6] of linear space $\mathbb{V}$ onto itself such that, for any vectors $|a\rangle$, $|b\rangle$ and any scalar λ

$$\hat{A}(|a\rangle + |b\rangle) = \hat{A}\,|a\rangle + \hat{A}\,|b\rangle\,; \tag{A.14a}$$

$$\hat{A}(\lambda\,|a\rangle) = \lambda\hat{A}\,|a\rangle\,. \tag{A.14b}$$

Exercise A.31. Decide whether the following maps are linear operators[7]:

a) $\hat{A}\,|a\rangle \equiv 0$.
b) $\hat{A}\,|a\rangle = |a\rangle$.
c) $\mathbb{C}^2 \to \mathbb{C}^2 : \hat{A}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -y \end{pmatrix}$.
d) $\mathbb{C}^2 \to \mathbb{C}^2 : \hat{A}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ xy \end{pmatrix}$.
e) $\mathbb{C}^2 \to \mathbb{C}^2 : \hat{A}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+1 \\ y+1 \end{pmatrix}$.
f) Rotation by angle ϕ in the linear space of two-dimensional geometric vectors (over $\mathbb{R}$).

Definition A.16. For any two operators $\hat{A}$ and $\hat{B}$, their sum $\hat{A}+\hat{B}$ is an operator that maps vectors according to

$$(\hat{A}+\hat{B})\,|a\rangle \equiv \hat{A}\,|a\rangle + \hat{B}\,|a\rangle\,. \tag{A.15}$$

For any operator $\hat{A}$ and any scalar λ, their product $\lambda\hat{A}$ is an operator that maps vectors according to

$$(\lambda\hat{A})\,|a\rangle \equiv \lambda\,(\hat{A}\,|a\rangle). \tag{A.16}$$

Exercise A.32. Show that the set of all linear operators over a Hilbert space of dimension N is itself a linear space, with the addition and multiplication by a scalar given by Eqs. (A.15) and (A.16), respectively.

---

[6] A *map* is a function that establishes, for every element *ket*a in $\mathbb{V}$, a unique "image" $\hat{A}\,|a\rangle$.

[7] $\mathbb{C}^2$ is the linear space of columns $\begin{pmatrix} x \\ y \end{pmatrix}$ consisting of two complex numbers.

a) Show that the operators $\hat{A}+\hat{B}$ and $\lambda\hat{A}$ are linear in the sense of Defn. A.15.
b) In the space of liner operators, what is the zero element and the opposite element $-\hat{A}$ for a given $\hat{A}$?
c)§ Show that the space of linear operators complies with all the axioms introduced in Definition A.1.

Definition A.17. The operator $\hat{\mathbf{1}}$ that maps every vector in $\mathbb{V}$ onto itself is called the *identity operator*.

When writing products of a scalar with identity operators, we sometimes omit the symbol $\hat{\mathbf{1}}$, provided that the context allows no ambiguity. For example, instead of writing $\hat{A}-\lambda\hat{\mathbf{1}}$, we may simply write $\hat{A}-\lambda$.

Definition A.18. For operators $\hat{A}$ and $\hat{B}$, their *product* $\hat{A}\hat{B}$ is an operator that maps every vector $|a\rangle$ onto $\hat{A}\hat{B}|a\rangle \equiv \hat{A}(\hat{B}|a\rangle)$. That is, in order to find the action of the operator $\hat{A}\hat{B}$ on a vector, we must first apply $\hat{B}$ to that vector, and then apply $\hat{A}$ to the result.

Exercise A.33. Show that a product of two linear operators is a linear operator.

It does matter in which order the two operators are multiplied, i.e., generally $\hat{A}\hat{B}\neq\hat{B}\hat{A}$. Operators for which $\hat{A}\hat{B}=\hat{B}\hat{A}$ are said to *commute*. Commutation relations between operators play an important role in quantum mechanics, and will be discussed in detail in Sec. A.9.

Exercise A.34. Show that the operators of counterclockwise rotation by angle $\pi/2$ and reflection about the horizontal axis in the linear space of two-dimensional geometric vectors do not commute.

Exercise A.35. Show that multiplication of operators has the property of associativity, i.e., for any three operators, one has

$$\hat{A}(\hat{B}\hat{C})=(\hat{A}\hat{B})\hat{C}. \tag{A.17}$$

A.6.2 Matrices

It may appear that, in order to fully describe a linear operator, we must say what it does to every vector. However, this is not the case. In fact, it is enough to say how the operator maps the elements of some basis $\{|v_1\rangle,\ldots,|v_N\rangle\}$ in $\mathbb{V}$, i.e., it is enough to know the set $\{\hat{A}|v_1\rangle,\ldots,\hat{A}|v_N\rangle\}$. Then, for any other vector $|a\rangle$, which can be decomposed as

$$|a\rangle = a_1|v_1\rangle+\ldots+a_N|v_N\rangle,$$

we have, thanks to linearity,

$$\hat{A}|a\rangle = a_1\hat{A}|v_1\rangle+\ldots+a_N\hat{A}|v_N\rangle. \tag{A.18}$$

How many numerical parameters does one need to completely characterize a linear operator? Each image $\hat{A}\left|v_j\right\rangle$ of a basis element can be decomposed into the same basis:

$$\hat{A}\left|v_j\right\rangle = \sum_i A_{ij}\left|v_i\right\rangle. \tag{A.19}$$

For every j, the set of N parameters $A_{1j},\ldots,A_{Nj}$ fully describes $\hat{A}\left|v_j\right\rangle$. Accordingly, the set of N^2 parameters A_{ij}, with both i and j varying from 1 to N, contains full information about a linear operator.

Definition A.19. The *matrix of an operator* in the basis $\{|v_i\rangle\}$ is an $N\times N$ square table whose elements are given by Eq. (A.21). The first index of A_{ij} is the number of the row, the second is the number of the column.

Suppose, for example, that you are required to prove that two given operators are equal: $\hat{A}=\hat{B}$. You can do so by showing the identity for the matrices A_{ij} and B_{ij} of the operators in any basis. Because the matrix contains full information about an operator, this is sufficient. Of course, you should choose your basis judiciously, so that the matrices A_{ij} and B_{ij} are as easy as possible to calculate.

Exercise A.36. Find the matrix of $\hat{\mathbf{1}}$. Show that this matrix does not depend on the choice of basis.

Exercise A.37. Find the matrix representation of the vector $\hat{A}\left|v_j\right\rangle$ in the basis $\{|v_i\rangle\}$, where $\left|v_j\right\rangle$ is an element of this basis, j is given, and the matrix of $\hat{A}$ is known.

Exercise A.38. Show that, if $|a\rangle \simeq \begin{pmatrix} a_1 \\ \vdots \\ a_N \end{pmatrix}$ in some basis, then the vector $\hat{A}|a\rangle$ is given by the matrix product

$$\hat{A}|a\rangle \simeq \begin{pmatrix} A_{11} & \ldots & A_{1N} \\ \vdots & & \vdots \\ A_{N1} & \ldots & A_{NN} \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_N \end{pmatrix} = \begin{pmatrix} \sum_j A_{1j}a_j \\ \vdots \\ \sum_j A_{Nj}a_j \end{pmatrix}. \tag{A.20}$$

Exercise A.39. Given the matrices A_{ij} and B_{ij} of the operators $\hat{A}$ and $\hat{B}$, find the matrices of the operators

a) $\hat{A}+\hat{B}$;
b) $\lambda\hat{A}$;
c) $\hat{A}\hat{B}$.

The last two exercises show that operations with operators and vectors are readily represented in terms of matrices and columns. However, there is an important caveat: matrices of vectors and operators depend on the basis chosen, in contrast to "physical" operators and vectors that are defined irrespectively of any specific basis.

This point should be taken into account when deciding whether to perform a calculation in the Dirac or matrix notation. If you choose the matrix notation to save

ink, you should be careful to keep track of the basis you are working with, and write all the matrices in that same basis.

Exercise A.40. Show that the matrix elements of the operator $\hat{A}$ in an *orthonormal* basis $\{|v_i\rangle\}$ are given by

$$A_{ij} = \langle v_i | \left(\hat{A} \left| v_j \right\rangle\right) \equiv \langle v_i | \hat{A} \left| v_j \right\rangle . \tag{A.21}$$

Exercise A.41. Find the matrices of operators $\hat{R}_\phi$ and $\hat{R}_\theta$ that correspond to the rotation of the two-dimensional geometric space through angles ϕ and θ, respectively [Ex. A.31(f)]. Find the matrix of $\hat{R}_\phi \hat{R}_\theta$ using the result of Ex. A.39 and check that it is equivalent to a rotation through $(\phi + \theta)$.

Exercise A.42. Give an example of a basis and determine the dimension of the linear space of linear operators over a Hilbert space of dimension N (see Ex. A.32).

A.6.3 Outer products

Definition A.20. *Outer products* $|a\rangle\langle b|$ are understood as operators acting as follows:

$$(|a\rangle\langle b|)\,|c\rangle \equiv |a\rangle\,(\langle b|\,c\rangle) = (\langle b|\,c\rangle)\,|a\rangle . \tag{A.22}$$

(The second equality comes from the fact that $\langle b|\,c\rangle$ is a number and commutes with everything.)

Exercise A.43. Show that $|a\rangle\langle b|$ as defined above is a linear operator.

Exercise A.44. Show that $(\langle a|\,b\rangle)\,(\langle c|\,d\rangle) = \langle a|\,(|b\rangle\langle c|)\,|d\rangle$.

Exercise A.45. Show that the matrix of the operator $|a\rangle\langle b|$ is given by

$$|a\rangle\langle b| \simeq \begin{pmatrix} a_1 \\ \vdots \\ a_N \end{pmatrix} \begin{pmatrix} b_1^* \ \ldots \ b_N^* \end{pmatrix} = \begin{pmatrix} a_1 b_1^* & \ldots & a_1 b_N^* \\ \vdots & & \vdots \\ a_N b_1^* & \ldots & a_N b_N^* \end{pmatrix} . \tag{A.23}$$

This result explains the intuition behind the notion of the outer product. As discussed in the previous section, a ket-vector corresponds to a column and a bra-vector to a row. According to the rules of matrix multiplication, the product of the two is a square matrix, and the outer product is simply the operator corresponding to this matrix.

Exercise A.46. Let A_{ij} be the matrix of the operator $\hat{A}$ in an orthonormal basis $\{|v_i\rangle\}$. Show that

$$\hat{A} = \sum_{i,j} A_{ij} \left| v_i \right\rangle \left\langle v_j \right| . \tag{A.24}$$

Exercise A.47. Let $\hat{A}$ be an operator and $\{|v_i\rangle\}$ an orthonormal basis in the Hilbert space. It is known that $\hat{A}|v_1\rangle = |w_1\rangle, \ldots, \hat{A}|v_N\rangle = |w_N\rangle$, where $|w_1\rangle, \ldots, |w_N\rangle$ are some vectors (not necessarily orthonormal). Show that

$$\hat{A} = \sum_i |w_i\rangle\langle v_i| . \tag{A.25}$$

These exercises reveal the significance of outer products. First, they provide a way to convert the operator matrix into the Dirac notation as per Eq. (A.24). This result complements Eq. (A.21), which serves the reverse purpose, converting the operator from the Dirac form into the matrix notation. Second, Eq. (A.25) allows us to construct the expression for an operator based on our knowledge of how it maps elements of an arbitrary orthonormal basis. We find it to be of great practical utility when we try to associate an operator with a physical process.

Below are two practice exercises using these results, followed by one very important additional application of the outer product.

Exercise A.48. The matrix of the operator $\hat{A}$ in the basis $\{|v_1\rangle, |v_2\rangle\}$ is $\begin{pmatrix} 1 & -3\mathrm{i} \\ 3\mathrm{i} & 4 \end{pmatrix}$. Express this operator in the Dirac notation.

Exercise A.49. Let $\{|v_1\rangle, |v_2\rangle\}$ be an orthonormal basis in a two-dimensional Hilbert space. Suppose the operator $\hat{A}$ maps $|u_1\rangle = (|v_1\rangle + |v_2\rangle)/\sqrt{2}$ onto $|w_1\rangle = \sqrt{2}|v_1\rangle$ and $|u_2\rangle = (|v_1\rangle - |v_2\rangle)/\sqrt{2}$ onto $|w_2\rangle = \sqrt{2}(|v_1\rangle + 3\mathrm{i}|v_2\rangle)$. Find the matrix of $\hat{A}$ in the basis $\{|v_1\rangle, |v_2\rangle\}$.
Hint: Notice that $\{|u_1\rangle, |u_2\rangle\}$ is an orthonormal basis.

Exercise A.50. Show that for any orthonormal basis $\{|v_i\rangle\}$,

$$\sum_i |v_i\rangle\langle v_i| = \hat{\mathbf{1}}. \tag{A.26}$$

This result is known as *the resolution of the identity*. It is useful for the following application. Suppose the matrix of $\hat{A}$ is known in some orthonormal basis $\{|v_i\rangle\}$ and we wish to find its matrix in another orthonormal basis, $\{|w_i\rangle\}$. This can be done as follows:

$$\begin{aligned}
(\hat{A}_{ij})_{w\text{-basis}} &= \langle w_i | \hat{A} | w_j \rangle \\
&= \langle w_i | \hat{\mathbf{1}}\hat{A}\hat{\mathbf{1}} | w_j \rangle \\
&= \langle w_i | \left(\sum_k |v_k\rangle\langle v_k| \right) \hat{A} \left(\sum_m |v_m\rangle\langle v_m| \right) | w_j \rangle \\
&= \sum_k \sum_m \langle w_i | v_k \rangle \langle v_k | \hat{A} | v_m \rangle \langle v_m | w_j \rangle .
\end{aligned} \tag{A.27}$$

The central object in the last line is the matrix element of $\hat{A}$ in the "old" basis $\{|v_i\rangle\}$. Because we know the inner products between each pair of elements in the old and

new bases, we can use the above expression to find each matrix element of $\hat{A}$ in the new basis. We shall use this trick throughout the course.

The calculation can be simplified if we interpret the last line of Eq. (A.27) as a product of three matrices. An example to that effect is given in the solution to the exercise below.

Exercise A.51. Find the matrix of the operator $\hat{A}$ from Ex. A.48 in the basis $\{|w_1\rangle, |w_2\rangle\}$ such that

$$|w_1\rangle = (|v_1\rangle + \mathrm{i}\,|v_2\rangle)/\sqrt{2}, \qquad (A.28)$$
$$|w_2\rangle = (|v_1\rangle - \mathrm{i}\,|v_2\rangle)/\sqrt{2}.$$

a) using the Dirac notation, starting with the result of Ex. A.48 and then expressing each bra and ket in the new basis;
b) according to Eq. (A.27).

Check that the results are the same.

A.7 Adjoint and self-adjoint operators

The action of an operator $\hat{A}$ on a ket-vector $|c\rangle$ corresponds to multiplying the square matrix of $\hat{A}$ by the column associated with $|c\rangle$. The result of this operation is another column, $\hat{A}\,|c\rangle$.

Let us by analogy consider an operation in which a row corresponding to a bra-vector $\langle b|$ is multiplied on the right by the square matrix of $\hat{A}$. The result of this operation will be another row corresponding to a bra-vector. We can associate such multiplication with the action of the operator $\hat{A}$ on $\langle b|$ *from the right*, denoted in the Dirac notation as $\langle b|\hat{A}$. The formal definition of this operation is as follows:

$$\langle b|\hat{A} \equiv \sum_{ij} b_i^* A_{ij} \langle v_j|\,, \qquad (A.29)$$

where A_{ij} and b_i are, respectively, the matrix elements of $\hat{A}$ and $|b\rangle$ in the orthonormal basis $\{|v_i\rangle\}$.

Exercise A.52. Derive the following properties of the operation defined by Eq. (A.29):

a) $\hat{A}$ acting from the right is a linear operator in the adjoint space;
b) $\langle a|\,b\rangle\langle c| = \langle a|\,(|b\rangle\langle c|)$;
c) for vectors $|a\rangle$ and $|c\rangle$,

$$\left(\langle a|\hat{A}\right)|c\rangle = \langle a|\left(\hat{A}\,|c\rangle\right); \qquad (A.30)$$

d) the vector $\langle a|\hat{A}$ as defined by Eq. (A.29) does not depend on the basis in which the matrix (A_{ij}) is calculated.

Let us now consider the following problem. Suppose we have an operator $\hat{A}$ that maps a ket-vector $|a\rangle$ onto ket-vector $|b\rangle$: $\hat{A}|a\rangle = |b\rangle$. What is the operator $\hat{A}^\dagger$ which, when acting from the right, maps *bra*-vector $\langle a|$ onto bra-vector $\langle b|$: $\langle a|\hat{A}^\dagger = \langle b|$? It turns out that this operator is not the same as $\hat{A}$, but is related relatively simply to it.

Definition A.21. An operator $\hat{A}^\dagger$ ("A-dagger") is called the *adjoint* (*Hermitian conjugate*) of $\hat{A}$ if for any vector $|a\rangle$,

$$\langle a|\hat{A}^\dagger = \text{Adjoint}\left(\hat{A}|a\rangle\right). \tag{A.31}$$

If $\hat{A} = \hat{A}^\dagger$, the operator is said to be *Hermitian* or *self-adjoint*.

Unlike bra- and ket-vectors, operators and their adjoints live in the same Hilbert space. More precisely, they live in both the bra- and ket- spaces: they act on bra-vectors from the right, and on ket-vectors from the left. Note that an operator cannot act on a bra-vector from the left or on a ket-vector from the right.

Exercise A.53. Show that the matrix of $\hat{A}^\dagger$ is related to the matrix of $\hat{A}$ through transposition and complex conjugation.

Exercise A.54. Show that, for any operator, $(\hat{A}^\dagger)^\dagger = \hat{A}$.

Exercise A.55. Show that the Pauli operators (1.7) are Hermitian.

Exercise A.56. By way of counterexample, show that two operators being Hermitian does not guarantee that their product is also Hermitian.

Exercise A.57. Show that

$$(|c\rangle\langle b|)^\dagger = |b\rangle\langle c|. \tag{A.32}$$

It may appear from this exercise that the adjoint of an operator is somehow related to its inverse: if the "direct" operator maps $|b\rangle$ onto $|c\rangle$, its adjoint does the opposite. This is not always the case: as we know from the Definition A.20 of the outer product, the operator $|c\rangle\langle b|$, when acting from the left, maps *everything* (not only $|c\rangle$) onto $|b\rangle$, while $|c\rangle\langle b|$ maps everything onto $|c\rangle$. However, there is an important class of operators, the so-called unitary operators, for which the inverse is the same as the adjoint. We discuss these operators in detail in Sec. A.10.

Exercise A.58. Show that

a)
$$(\hat{A}+\hat{B})^\dagger = \hat{A}^\dagger + \hat{B}^\dagger; \tag{A.33}$$

b)
$$(\lambda\hat{A})^\dagger = \lambda^*\hat{A}^\dagger; \tag{A.34}$$

c)
$$(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger. \tag{A.35}$$

We can say that every object in linear algebra has an adjoint. For a number, its adjoint is its complex conjugate; for a ket-vector it is a bra-vector (and vice versa);

for an operator it is the adjoint operator. The matrices of an object and its adjoint are related by transposition and complex conjugation.

Suppose we are given a complex expression consisting of vectors and operators, and are required to find its adjoint. Summarizing Eqs. (A.12), (A.32) and (A.35), we arrive at the following algorithm:

a) invert the order of all products;
b) conjugate all numbers;
c) replace all kets by bras and vice versa;
d) replace all operators by their adjoints.

Here is an example.

$$\text{Adjoint}\left(\lambda\hat{A}\hat{B}\,|a\rangle\,\langle b|\,\hat{C}\right) = \lambda^*\hat{C}^\dagger\,|b\rangle\,\langle a|\,\hat{B}^\dagger\hat{A}^\dagger \tag{A.36}$$

This rule can be used to obtain the following relation.

Exercise A.59. Show that

$$\langle\phi|\,\hat{A}\,|\psi\rangle = \langle\psi|\hat{A}^\dagger\,|\phi\rangle^*\,. \tag{A.37}$$

A.8 Spectral decomposition

We will now prove an important theorem for Hermitian operators. I will be assuming you are familiar with the notions of determinant, eigenvalue, and eigenvector of a matrix and the methods for finding them. If this is not the case, please refer to any introductory linear algebra text.

Exercise A.60.* Prove the *spectral theorem*: for any Hermitian operator $\hat{V}$, there exists an orthonormal basis $\{|v_i\rangle\}$ (which we shall call the *eigenbasis*) such that

$$\hat{V} = \sum_i v_i\,|v_i\rangle\,\langle v_i|\,, \tag{A.38}$$

with all the v_i being real.

The representation of an operator in the form (A.38) is called the *spectral decomposition* or *diagonalization* of the operator. The basis $\{|v_i\rangle\}$ is called an *eigenbasis* of the operator.

Exercise A.61. Write the matrix of the operator (A.38) in its eigenbasis.

Exercise A.62. Show that the elements of the eigenbasis of $\hat{V}$ are the eigenvectors of $\hat{V}$ and the corresponding values v_i are its eigenvalues, i.e., for any i,

$$\hat{V}\,|v_i\rangle = v_i\,|v_i\rangle\,.$$

Exercise A.63.[*§] Show that a spectral decomposition (not necessarily with real eigenvalues) exists for any operator $\hat{V}$ such that $\hat{V}\hat{V}^\dagger = \hat{V}^\dagger\hat{V}$ (such operators are said to be *normal*).

Exercise A.64. Find the eigenvalues and eigenbasis of the operator associated with the rotation of the plane of two-dimensional geometric vectors through angle ϕ (see Ex. A.41), but over the field of *complex* numbers.

Exercise A.65.[§] In a three-dimensional Hilbert space, three operators have the following matrices in an orthonormal basis $\{|v_1\rangle, |v_2\rangle, |v_3\rangle\}$:

a) $\hat{L}_x \simeq \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$,

b) $\hat{L}_y \simeq \begin{pmatrix} 0 & -\mathrm{i} & 0 \\ \mathrm{i} & 0 & -\mathrm{i} \\ 0 & \mathrm{i} & 0 \end{pmatrix}$;

c) $\hat{L}_z \simeq \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

Show that these operators are Hermitian. Find their eigenvalues and eigenvectors.

So we have found that every Hermitian operator has a spectral decomposition. But is the spectral decomposition of a given operator unique? The answer is affirmative as long as the operator has no *degenerate eigenvalues*, i.e., eigenvalues associated with two or more eigenvectors.

Exercise A.66. The Hermitian operator $\hat{V}$ diagonalizes in an orthonormal basis $\{|v_i\rangle\}$. Suppose there exists a vector $|\psi\rangle$ that is an eigenvector of $\hat{V}$ with eigenvalue v, but is not proportional to any $|v_i\rangle$. Show that this is possible only if v is a degenerate eigenvalue of $\hat{V}$ and $|\psi\rangle$ is a linear combination of elements of $\{|v_i\rangle\}$ corresponding to that eigenvalue.

Exercise A.67. Show that, for a Hermitian operator $\hat{V}$ whose eigenvalues are non-degenerate,

a) the eigenbasis is unique up to phase factors;
b) any set that contains all linearly independent normalized eigenvectors of $\hat{V}$ is identical to the eigenbasis of $\hat{V}$ up to phase factors.

The latter result is of primary importance, and we shall make abundant use of it throughout this course. It generalizes to Hilbert spaces of infinite dimension and even to those associated with continuous observables. Let us now look into the case of operators with degenerate eigenvalues.

Exercise A.68. Find the eigenvalues of the identity operator in the qubit Hilbert space and show that they are degenerate. Give two different examples of this operator's eigenbasis.

Exercise A.69. Show that eigenvectors of a Hermitian operator $\hat{V}$ that are associated with different eigenvalues are orthogonal. Do not assume non-degeneracy of the eigenvalues.

Exercise A.70. Suppose an eigenvalue v of an operator $\hat{V}$ is degenerate. Show that a set of corresponding eigenvectors forms a linear subspace (see Defn. A.8).

Exercise A.71*.

a) Show that if $\langle\psi|\,\hat{A}\,|\psi\rangle = \langle\psi|\,\hat{B}\,|\psi\rangle$ for all $|\psi\rangle$, then $\hat{A} = \hat{B}$.
b) Show that if $\langle\psi|\,\hat{A}\,|\psi\rangle$ is a real number for all $|\psi\rangle$, then $\hat{A}$ is Hermitian.

Definition A.22. A Hermitian operator $\hat{A}$ is said to be *positive (non-negative)* if $\langle\psi|\,\hat{A}\,|\psi\rangle > 0$ ($\langle\psi|\,\hat{A}\,|\psi\rangle \geq 0$) for any non-zero vector $|\psi\rangle$.

Exercise A.72. Show that a Hermitian operator $\hat{A}$ is positive (non-negative) if and only if all its eigenvalues are positive (non-negative).

Exercise A.73. Show that a sum $\hat{A} + \hat{B}$ of two positive (non-negative) operators is positive (non-negative).

A.9 Commutators

As already discussed, not all operators commute. The degree of non-commutativity turns out to play an important role in quantum mechanics and is quantified by the operator known as the commutator.

Definition A.23. For any two operators $\hat{A}$ and $\hat{B}$, their *commutator* and *anticommutator* are defined respectively by

$$[\hat{A},\hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}; \tag{A.39a}$$
$$\{\hat{A},\hat{B}\} = \hat{A}\hat{B} + \hat{B}\hat{A}. \tag{A.39b}$$

Exercise A.74. Show that:

a)
$$\hat{A}\hat{B} = \frac{1}{2}([\hat{A},\hat{B}] + \{\hat{A},\hat{B}\}); \tag{A.40}$$

b)
$$[\hat{A},\hat{B}] = -[\hat{B},\hat{A}]; \tag{A.41}$$

c)
$$[\hat{A},\hat{B}]^\dagger = [\hat{B}^\dagger,\hat{A}^\dagger]; \tag{A.42}$$

d)
$$[\hat{A},\hat{B}+\hat{C}] = [\hat{A},\hat{B}] + [\hat{A},\hat{C}]; \tag{A.43a}$$
$$[\hat{A}+\hat{B},\hat{C}] = [\hat{A},\hat{C}] + [\hat{B},\hat{C}]; \tag{A.43b}$$

e)

$$[\hat{A},\hat{B}\hat{C}] = [\hat{A},\hat{B}]\hat{C} + \hat{B}[\hat{A},\hat{C}]; \tag{A.44a}$$

$$[\hat{A}\hat{B},\hat{C}] = [\hat{A},\hat{C}]\hat{B} + \hat{A}[\hat{B},\hat{C}]; \tag{A.44b}$$

f)

$$\begin{aligned}[\hat{A}\hat{B},\hat{C}\hat{D}] &= \hat{C}\hat{A}[\hat{B},\hat{D}] + \hat{C}[\hat{A},\hat{D}]\hat{B} + \hat{A}[\hat{B},\hat{C}]\hat{D} + [\hat{A},\hat{C}]\hat{B}\hat{D} \\ &= \hat{A}\hat{C}[\hat{B},\hat{D}] + \hat{C}[\hat{A},\hat{D}]\hat{B} + \hat{A}[\hat{B},\hat{C}]\hat{D} + [\hat{A},\hat{C}]\hat{D}\hat{B}.\end{aligned} \tag{A.45}$$

When calculating commutators for complex expressions, it is advisable to use the relations derived in this exercise rather than the definition (A.39a) of the commutator. There are many examples to this effect throughout this book.

Exercise A.75. Express the commutators

a) $[\hat{A}\hat{B}\hat{C},\hat{D}]$;
b) $[\hat{A}^2 + \hat{B}^2, \hat{A} + \mathrm{i}\hat{B}]$

in terms of the pairwise commutators of the individual operators $\hat{A},\hat{B},\hat{C},\hat{D}$.

Exercise A.76. For two operators $\hat{A}$ and $\hat{B}$, suppose that $[\hat{A},\hat{B}] = ic\hat{\mathbf{1}}$, where c is a complex number. Show that

$$[\hat{A},\hat{B}^n] = nc\hat{B}^{n-1}. \tag{A.46}$$

Exercise A.77. Show that, if $\hat{A}$ and $\hat{B}$ are Hermitian, so are

a) $i[\hat{A},\hat{B}]$;
b) $\{\hat{A},\hat{B}\}$.

Exercise A.78. Find the commutation relations of the Pauli operators (1.7).
Answer:

$$[\hat{\sigma}_m,\hat{\sigma}_j] = 2\mathrm{i}\varepsilon_{mjk}\sigma_k, \tag{A.47}$$

where ε is the Levi-Civita symbol given by

$$\varepsilon_{mjk} \equiv \begin{cases} +1 \text{ for } mjk = xyz,\ yzx \text{ or } zxy \\ -1 \text{ for } mjk = xzy,\ yxz \text{ or } zyx \\ \qquad 0 \text{ otherwise.} \end{cases} \tag{A.48}$$

A.10 Unitary operators

Definition A.24. Linear operators that map all vectors of norm 1 onto vectors of norm 1 are said to be *unitary*.

Exercise A.79. Show that unitary operators preserve the norm of any vector, i.e., if $|a'\rangle = \hat{U}|a\rangle$, then $\langle a|\,a\rangle = \langle a'|\,a'\rangle$.

Exercise A.80. Show that an operator $\hat{U}$ is unitary if and only if it preserves the inner product of any two vectors, i.e., if $|a'\rangle = \hat{U}\,|a\rangle$ and $|b'\rangle = \hat{U}\,|b\rangle$, then $\langle a|\ b\rangle = \langle a'|\ b'\rangle$.

Exercise A.81. Show that:

a) a unitary operator maps any orthonormal basis $\{|w_i\rangle\}$ onto an orthonormal basis;
b) conversely, for any two orthonormal bases $\{|v_i\rangle\}, \{|w_i\rangle\}$, the operator $\hat{U} = \sum_i |v_i\rangle\langle w_i|$ is unitary (in other words, *any* operator that maps an orthonormal basis onto an orthonormal basis is unitary).

Exercise A.82. Show that an operator $\hat{U}$ is unitary if and only if $\hat{U}^\dagger\hat{U} = \hat{U}\hat{U}^\dagger = \hat{\mathbf{1}}$ (i.e., its adjoint is equal to its inverse).

Exercise A.83. Show the following:

a) Any unitary operator can be diagonalized and all its eigenvalues have absolute value 1, i.e., they can be written in the form $e^{i\theta},\ \theta \in \mathbb{R}$.
Hint: use Ex. A.63.
b) A diagonalizable operator (i.e., an operator whose matrix becomes diagonal in some basis) with eigenvalues of absolute value 1 is unitary.

Exercise A.84. Show that the following operators are unitary:

a) the Pauli operators (1.7);
b) rotation through angle ϕ in the linear space of two-dimensional geometric vectors over $\mathbb{R}$.

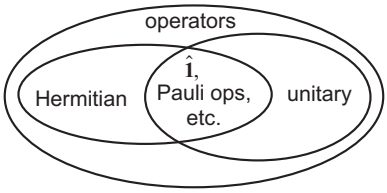


Fig. A.1 Relations among types of operators

The families of Hermitian and unitary operators overlap, but they are not identical (Fig. A.1). An operator that is both Hermitian and unitary must be self-inverse, as per Ex. A.82. Such operators are relatively rare.

A.11 Functions of operators

The concept of function of an operator has many applications in linear algebra and differential equations. It is also handy in quantum mechanics, as operator functions permit easy calculation of evolution operators.

Definition A.25. Consider a complex function $f(x)$ defined on $\mathbb{C}$. The *function of operator* $f(\hat{A})$ of a diagonalizable operator $\hat{A}$ is the following operator:

$$f(\hat{A}) = \sum_i f(a_i) |a_i\rangle \langle a_i| , \tag{A.49}$$

where $\{|a_i\rangle\}$ is an orthonormal basis in which $\hat{A}$ diagonalizes:

$$\hat{A} = \sum_i a_i |a_i\rangle \langle a_i| . \tag{A.50}$$

Exercise A.85. Show that, if the vector $|a\rangle$ is an eigenvector of a Hermitian operator $\hat{A}$ with eigenvalue a, then $f(\hat{A})|a\rangle = f(a)|a\rangle$.

Exercise A.86. Suppose that the operator $\hat{A}$ is Hermitian and the function $f(x)$, when applied to a real argument x, takes a real value. Show that $f(\hat{A})$ is a Hermitian operator, too.

Exercise A.87. Suppose that the operator $\hat{A}$ is Hermitian and function $f(x)$, when applied to any real argument x, takes a real non-negative value. Show that $f(\hat{A})$ is a non-negative operator (see Defn. A.22).

Exercise A.88. Find the matrices of $\sqrt{\hat{A}}$ and $\ln\hat{A}$ in the orthonormal basis in which

$$\hat{A} \simeq \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$$

Exercise A.89. Find the matrix of $e^{i\theta\hat{A}}$, where $\hat{A} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$.

Hint: One of the eigenvalues of $\hat{A}$ is 0, which means that the corresponding eigenvector does not appear in the spectral decomposition (A.50) of $\hat{A}$. However, the exponential of the corresponding eigenvalue is not zero, and the corresponding eigenvectors do show up in the operator function (A.49).

Exercise A.90. Show that, for any operator $\hat{A}$ and function f, $[\hat{A}, f(\hat{A})] = 0$.

Exercise A.91. Suppose $f(x)$ has a Taylor decomposition $f(x) = f_0 + f_1 x + f_2 x^2 + \ldots$. Show that $f(\hat{A}) = f_0\hat{\mathbf{1}} + f_1\hat{A} + f_2\hat{A}^2 + \ldots$

Exercise A.92. Show that, if the operator $\hat{A}$ is Hermitian, the operator $e^{i\hat{A}}$ is unitary and $e^{i\hat{A}} = \left(e^{-i\hat{A}}\right)^{-1}$.

Exercise A.93*. Let $\vec{s} = (s_x, s_y, s_z)$ be a unit vector (i.e. a vector of length 1). Show that:

$$e^{i\theta\vec{s}\cdot\hat{\vec{\sigma}}} = \cos\theta\,\hat{\mathbf{1}} + i\sin\theta\,\vec{s}\cdot\hat{\vec{\sigma}}, \tag{A.51}$$

where $\hat{\vec{\sigma}} = (\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z)$, $\vec{s}\cdot\hat{\vec{\sigma}} = s_x\hat{\sigma}_x + s_y\hat{\sigma}_y + s_z\hat{\sigma}_z$.
Hint: There is no need find the explicit solutions for the eigenvectors of the operator $\vec{s}\cdot\hat{\vec{\sigma}}$.

Exercise A.94§ Find the matrices of the operators $e^{i\theta\hat{\sigma}_x}$, $e^{i\theta\hat{\sigma}_y}$, $e^{i\theta\hat{\sigma}_z}$ in the canonical basis.
Answer:

$$e^{i\theta\hat{\sigma}_x} = \begin{pmatrix} \cos\theta & i\sin\theta \\ i\sin\theta & \cos\theta \end{pmatrix};$$
$$e^{i\theta\hat{\sigma}_y} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix};$$
$$e^{i\theta\hat{\sigma}_z} = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}.$$

Definition A.26. Suppose the vector $|\psi(t)\rangle$ depends on a certain parameter t. The *derivative* of $|\psi(t)\rangle$ with respect to t is defined as the vector

$$\frac{d|\psi\rangle}{dt} = \lim_{\Delta t\to 0}\frac{|\psi(t+\Delta t)\rangle - |\psi(t)\rangle}{\Delta t}. \tag{A.52}$$

Similarly, the derivative of the operator $\hat{Y}(t)$ with respect to t is the operator

$$\frac{d\hat{Y}}{dt} = \lim_{\Delta t\to 0}\frac{\hat{Y}(t+\Delta t) - \hat{Y}(t)}{\Delta t}. \tag{A.53}$$

Exercise A.95. Suppose that the matrix form of the vector $|\psi(t)\rangle$ is

$$|\psi(t)\rangle = \begin{pmatrix} \psi_1(t) \\ \vdots \\ \psi_N(t) \end{pmatrix}$$

in some basis. Show that

$$\frac{d|\psi\rangle}{dt} = \begin{pmatrix} d\psi_1(t)/dt \\ \vdots \\ d\psi_N(t)/dt \end{pmatrix}.$$

Write an expression for the matrix form of an operator derivative.

Exercise A.96. Suppose the operator $\hat{A}$ is diagonalizable in an orthonormal basis and independent of t, where t is a real parameter. Show that $\frac{d}{dt}e^{i\hat{A}t} = i\hat{A}e^{i\hat{A}t} = ie^{i\hat{A}t}\hat{A}$.

Exercise A.97:* For two operators $\hat{A}$ and $\hat{B}$, suppose that $[\hat{A},\hat{B}] = ic\hat{\mathbf{1}}$, where c is a complex number. Prove the *Baker-Hausdorff-Campbell formula*[8]

$$e^{\hat{A}+\hat{B}} = e^{\hat{A}}e^{\hat{B}}e^{-ic/2} = e^{\hat{B}}e^{\hat{A}}e^{ic/2} \tag{A.54}$$

using the following steps.

a) Show that

$$[\hat{A}, e^{\hat{B}}] = ce^{\hat{B}}. \tag{A.55}$$

Hint: use the Taylor series expansion for the exponential and Eq. (A.46).

b) For an arbitrary number λ and operator $\hat{G}(\lambda) = e^{\lambda\hat{A}}e^{\lambda\hat{B}}$, show that

$$\frac{d\hat{G}(\lambda)}{d\lambda} = \hat{G}(\lambda)(\hat{A}+\hat{B}+\lambda c) \tag{A.56}$$

c) Solve the differential equation (A.56) to show that

$$\hat{G}(\lambda) = e^{\lambda\hat{A}+\lambda\hat{B}+\lambda^2 c/2}. \tag{A.57}$$

d) Prove the Baker-Hausdorff-Campbell formula using Eq. (A.57).

[8] This is a simplified form of the Baker-Hausdorff-Campbell formula. The full form of this formula is more complicated and holds for the case when $[\hat{A},\hat{B}]$ does not commute with $\hat{A}$ or $\hat{B}$.

Appendix B
Probabilities and distributions

B.1 Expectation value and variance

Definition B.1. Suppose a (not necessarily quantum) experiment to measure a quantity Q can yield any one of N possible outcomes $\{Q_i\}$ $(1 \leq i \leq N)\}$, with respective probabilities pr_i. Then Q is called a *random variable* and the set of values $\{\mathrm{pr}_i\}$ for all values of i is called the *probability distribution*. The *expectation (mean) value* of Q is

$$\langle Q \rangle = \sum_{i=1}^{N} \mathrm{pr}_i Q_i. \tag{B.1}$$

Exercise B.1. Find the expectation of the value displayed on the top face of a fair die.

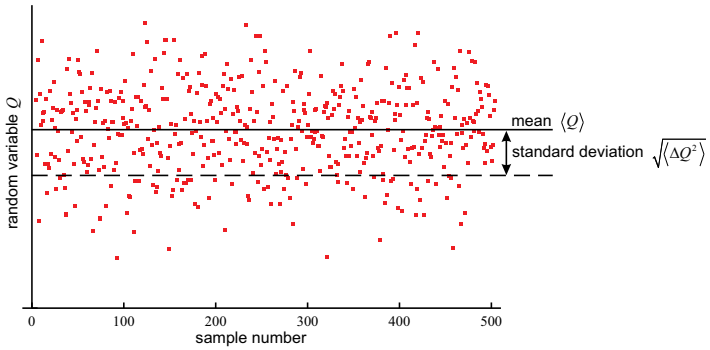


Fig. B.1 Mean and rms standard deviation of a random variable.

Definition B.2. The *mean square variance* of random variable Q is

A. I. Lvovsky, *Quantum Physics*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-662-56584-1

$$\left\langle \Delta Q^2 \right\rangle = \left\langle (Q - \langle Q \rangle)^2 \right\rangle = \sum_i \mathrm{pr}_i \left(Q_i - \langle Q \rangle\right)^2 . \tag{B.2}$$

The *root mean square (rms) standard deviation*, or *uncertainty* of Q is then $\sqrt{\langle \Delta Q^2 \rangle}$.

While the expectation value, $\langle Q \rangle = \sum_{i=1}^{N} \mathrm{pr}_i Q_i$, shows the mean measurement output, the statistical uncertainty shows by how much, on average, a particular measurement result will deviate from the mean (Fig. B.1).

Exercise B.2. Show that, for any random variable Q,

$$\left\langle \Delta Q^2 \right\rangle = \left\langle Q^2 \right\rangle - \langle Q \rangle^2 . \tag{B.3}$$

Exercise B.3. Calculate the mean square variance of the value displayed on the top face of a fair die. Show by direct calculation that Eqs. (B.2) and (B.3) yield the same.

Exercise B.4. Two random variables Q and R are independent, i.e., the realization of one does not affect the probability distribution of the other (for example, a die and a coin being tossed next to each other). Show that $\langle QR \rangle = \langle Q \rangle \langle R \rangle$. Is this statement valid if Q and R are not independent?
Hint: Independence means that events Q_i and R_j occur at the same time with probability $\mathrm{pr}_i^Q \mathrm{pr}_j^R$ for each pair (i, j), where pr_i^Q is the probability of the ith value of variable Q and pr_j^R is the probability of the jth value of R.

Exercise B.5. Suppose a random variable Q is measured (for example, a die is thrown) N times. Consider the random variable $\tilde{Q}$ that is the sum of the N outcomes. Show that the expectation and variance of $\tilde{Q}$ equal

$$\left\langle \tilde{Q} \right\rangle = N \langle Q \rangle$$

and

$$\left\langle \Delta \tilde{Q}^2 \right\rangle = N \left\langle \Delta Q^2 \right\rangle ,$$

respectively.

B.2 Conditional probabilities

The *conditional probability* $\mathrm{pr}_{A|B}$ is the probability of some event A given that another event, B, is known to have occurred. Examples are:

- the probability that the value on a die is odd given that it is greater than 3;
- the probability that Alice's HIV test result will be positive given that she is actually not infected;
- the probability that Bob plays basketball given that he is a man and 185 cm tall;

- the probability that it will rain tomorrow given that it has rained today.

Let us calculate the conditional probability using the third example. Event A is "Bob plays basketball". Event B is "Bob is a 185-cm tall man". The conditional probability is equal to the number $N(A \text{ and } B)$ of 185-cm tall men who play basketball divided by the number $N(B)$ of 185-cm tall men [Fig. B.2(a)]:

$$\mathrm{pr}_{A|B} = \frac{N(A \text{ and } B)}{N(B)}. \tag{B.4}$$

Let us divide both the numerator and the denominator of the above fraction by N, the total number of people in town. Then we have in the numerator $N(A \text{ and } B)/N = \mathrm{pr}_{A \text{ and } B}$ — the probability that a randomly chosen person is a 185-cm tall man who plays basketball, and in the denominator, $N(B)/N = \mathrm{pr}_B$ — the probability that a random person is a 185-cm tall man. Hence

$$\mathrm{pr}_{A|B} = \frac{\mathrm{pr}_{A \text{ and } B}}{\mathrm{pr}_B}. \tag{B.5}$$

This is a general formula for calculating conditional probabilities.

Exercise B.6. Suppose events $B_1, \ldots, B_n$ are mutually exclusive and collectively exhaustive, i.e., one of them must occur, but no two occur at the same time [Fig. B.2(b)]. Show that, for any other event A,

$$\mathrm{pr}_A = \sum_{i=1}^{n} \mathrm{pr}_{A|B_i} \mathrm{pr}_{B_i}. \tag{B.6}$$

This result is known as the *theorem of total probability*.

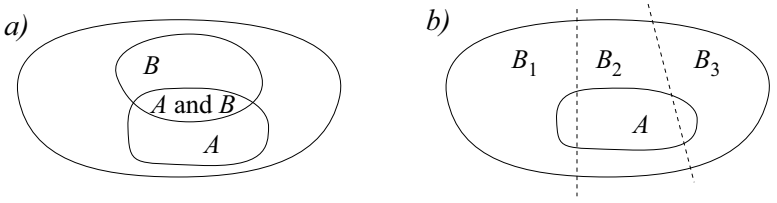


Fig. B.2 Conditional probabilities. a) Relation between the conditional and combined probabilities, Eq. (B.5). b) Theorem of total probability (Ex. B.6).

Exercise B.7. The probability that a certain HIV test gives a false positive result is

$$\mathrm{pr}_{\text{positive}|\text{not infected}} = 0.05.$$

The probability of a false negative result is zero. It is known that, of all people taking the test, the probability of actually being infected is $\mathrm{pr}_{\text{infected}} = 0.001$.

a) What is the probability $\mathrm{pr}_{\text{positive and not infected}}$ that a random person taking the test is not infected *and* shows a false positive result?
b) What is the probability $\mathrm{pr}_{\text{positive}}$ that a random person taking the test shows a positive result?
c) A random person, Alice, has been selected and the test has been performed on her. Her result turned out to be positive. What is the probability that she is not infected?

Hint: To visualize this problem, imagine a city of one million. How many of them are infected? How many are not? How many positive test results will there be all together?

B.3 Binomial and Poisson distributions

Exercise B.8. A coin is tossed n times. Find the probability that heads will appear k times, and tails $n-k$ times:

a) for a fair coin, i.e., the probability of getting heads or tails in a single toss is $1/2$;
b) for a *biased* coin, with the probabilities for the heads and tails being p and $1-p$, respectively.
Answer:

$$\mathrm{pr}_k = \binom{n}{k} p^k (1-p)^{n-k}. \tag{B.7}$$

The probability distribution defined by Eq. (B.7) is called the *binomial* distribution. We encounter this distribution in everyday life, often without realizing it. Here are a few examples.

Exercise B.9[§]

a) On a given day in a certain city 20 babies were born. What is the probability that exactly nine of them are girls?
b) A student answers $3/4$ of questions on average. What is the probability that (s)he scores perfectly on a 10-question test?
c) A certain politician has 60% electoral support. What is the probability that (s)he receives more than 50% of the votes in a district with 100 voters?

Exercise B.10. Find the expectation value and the uncertainty of the binomial distribution (B.7).
Answer:

$$\langle k \rangle = np; \quad \left\langle \Delta k^2 \right\rangle = np(1-p). \tag{B.8}$$

Exercise B.11. In a certain big city, 10 babies are born per day on average. What is the probability that on a given day, exactly 12 babies are born?

a) The city population is 100000.

b) The city population is 1000000.

Hint: Perhaps there is a way to avoid calculating 1000000!.

We see from the above exercise that in the limit $p \to 0$ and $n \to \infty$, but $\lambda = pn =$ const, the probabilities in the binomial distribution become dependent on λ, rather than p and n individually. This important extension of the binomial distribution is known as the *Poisson (Poissonian)* distribution.

Exercise B.12. Show that in the limit $p \to 0$ and $n \to \infty$, but $\lambda = pn =$ const, the binomial distribution (B.7) becomes

$$\mathrm{pr}_k = \mathrm{e}^{-\lambda} \frac{\lambda^k}{k!} \tag{B.9}$$

using the following steps.

a) Show that $\lim_{n\to\infty} \frac{1}{n^k} \binom{n}{k} = \frac{1}{k!}$.
b) Show that $\lim_{n\to\infty} (1-p)^{n-k} = \mathrm{e}^{-\lambda}$.
c) Obtain Eq. (B.9).

Exercise B.13. Find the answer to Ex. B.11 in the limit of an infinitely large city.

Here are some more examples of the Poisson distribution.

Exercise B.14[§]

a) A patrol policeman posted on a highway late at night has discovered that, on average, 60 cars pass every hour. What is the probability that, within a given minute, exactly one car will pass that policeman?
b) A cosmic ray detector registers 500 events per second on average. What is the probability that this number equals exactly 500 within a given second?
c) The average number of lions seen on a one-day safari is 3. What is the probability that, if you go on that safari, you will not see a single lion?

Exercise B.15. Show that both the mean and variance of the Poisson distribution (B.9) equal λ.

For example, in a certain city, 25 babies are born per day on average, so $\lambda = 25$. The root mean square uncertainty in this number $\sqrt{\lambda} = 5$, i.e., on a typical day we are much more likely to see 20 or 30 babies rather than 10 or 40 (Fig. B.3).

Although the absolute uncertainty of n increases with $\langle n \rangle$, the *relative* uncertainty $\sqrt{\lambda}/\lambda$ decreases. In our example above, the relative uncertainty is $5/25 = 20\%$. But in a smaller city, where $\langle n \rangle = 4$, the relative uncertainty is as high as $2/4 = 50\%$.

B.4 Probability densities

So far, we have studied random variables that can take values from a discrete set, with the probability of each value being finite. But what if we are dealing with

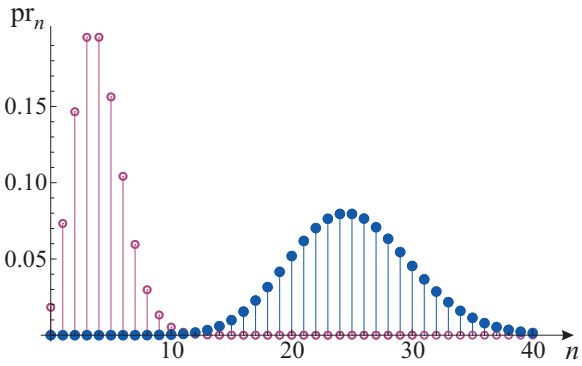


Fig. B.3 the Poisson distribution with $\langle n \rangle = 4$ (empty circles) and $\langle n \rangle = 25$ (filled circles).

a *continuous* random variable — for example, the wind speed, the decay time of a radioactive atom, or the range of a projectile? In this case, there is now way to assign a finite probability value to each specific value of Q. The probability that the atom decays after *precisely* two milliseconds, or the wind speed is *precisely* five meters per second, is infinitely small.

However, the probability of detecting Q within some range — for example, that the atom decays between times 2 ms and 2.01 ms — is finite. We can therefore *discretize* the continuous variable: divide the range of values that Q can take into equal bins of width δQ. Then we define a discrete random variable $\tilde{Q}$ with possible values $\tilde{Q}_i$ equal to the central point of each bin, and the associated finite probability $\mathrm{pr}_{\tilde{Q}_i}$ that Q falls within that bin [Fig. B.4(a,b)]. As for any probability distribution, $\sum_i \mathrm{pr}_{\tilde{Q}_i} = 1$. Of course, the narrower the bin width we choose, the more precisely we describe the behavior of the continuous random variable.

The probability values associated with neighboring bins can be expected to be close to each other if the bin width is chosen sufficiently small. For atomic decay, for example, we can write $\mathrm{pr}_{[2.00\ \mathrm{ms},2.01\ \mathrm{ms}]} \approx \mathrm{pr}_{[2.01\ \mathrm{ms},2.02\ \mathrm{ms}]} \approx \frac{1}{2}\mathrm{pr}_{[2.00\ \mathrm{ms},2.02\ \mathrm{ms}]}$. In other words, for small bin widths, the quantity $\mathrm{pr}_{\tilde{Q}_i}/\delta Q$ is independent of δQ. Hence we can introduce the notion of the *probability density* or *continuous probability distribution*[1]:

$$\mathrm{pr}(Q) = \lim_{\delta Q \to 0} \frac{\mathrm{pr}_{\tilde{Q}_i(Q)}}{\delta Q}, \tag{B.10}$$

where $i(Q)$ is the number of the bin within which the value of Q is located and the limit is taken over a set of discretized probability distributions for Q. This probability density is the primary characteristic of a continuous random variable.

Note also that, because the discrete probability $\mathrm{pr}_{\tilde{Q}_i(Q)}$ is a dimensionless quantity, the dimension of a continuous probability density $\mathrm{pr}(Q)$ is always the reciprocal dimension of the corresponding random variable Q.

[1] Throughout this book, I use subscripts for discrete probabilities, such as in pr_i or $\mathrm{pr}_{\tilde{Q}_i}$, and parentheses for continuous probability densities, e.g., $\mathrm{pr}(Q)$.

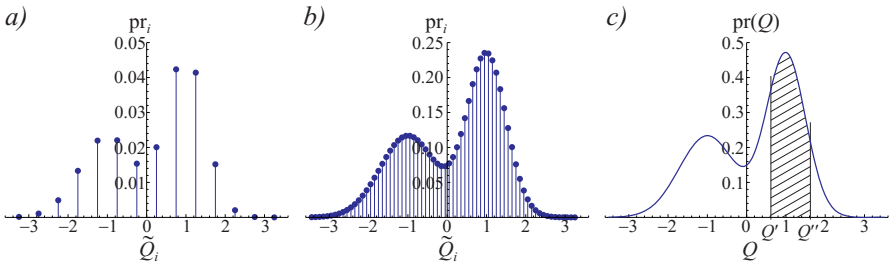


Fig. B.4 Continuous probability distribution. a), b) Discretization of the continuous random variable with bin widths $\delta Q = 0.5$ and 0.1, respectively. c) Continuous probability density. The probability of observing the variable in a range between Q' and Q'' is $\int_{Q'}^{Q''} \mathrm{pr}(Q)\mathrm{d}Q$. Note the variation of the vertical scales in the three plots.

Exercise B.16. For a continuous random variable with probability density $\mathrm{pr}(Q)$, show that:

a) the probability of observin the variable in the range between Q' and Q'' is

$$\mathrm{pr}_{[Q',Q'']} = \int_{Q'}^{Q''} \mathrm{pr}(Q)\mathrm{d}Q; \tag{B.11}$$

b) the probability density function is *normalized*:

$$\int_{-\infty}^{+\infty} \mathrm{pr}(Q)\mathrm{d}Q = 1; \tag{B.12}$$

c) the expectation value of Q is given by

$$\langle Q \rangle = \int_{-\infty}^{+\infty} Q\mathrm{pr}(Q)\mathrm{d}Q; \tag{B.13}$$

d)[§] the variance of Q is given by

$$\left\langle \Delta Q^2 \right\rangle = \int_{-\infty}^{+\infty} (Q - \langle Q \rangle)^2 \mathrm{pr}(Q)\mathrm{d}Q = \left\langle Q^2 \right\rangle - \langle Q \rangle^2 . \tag{B.14}$$

Exercise B.17. Find the probability density, expectation and root mean square uncertainty for the decay time t of a radioactive nucleus with half-life $\tau = 1$ ms.

A probability density that frequently occurs in nature is the *Gaussian*, or *normal* distribution:

$$G_b(x) = \frac{1}{b\sqrt{\pi}} \mathrm{e}^{-x^2/b^2}, \tag{B.15}$$

where b is the *width* of the Gaussian distribution (Fig. B.5). Typically, the Gaussian distribution governs physical quantities that are affected by multiple small random effects that add up[2]. Examples include:

- the position of a particle subjected to Brownian motion;
- the time shown by a clock affected by random fluctuations of the temperature in the room;
- the component of the velocity of a gas molecule along a particular axis.

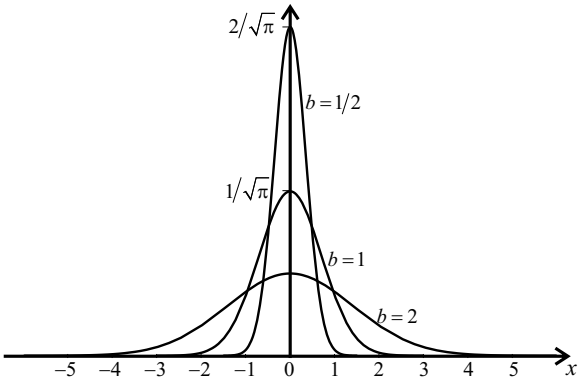


Fig. B.5 Normalized Gaussian functions of different widths.

Exercise B.18. For a Gaussian distribution $G_b(x-a)$, show the following:

a) Normalization holds, i.e.,

$$\int_{-\infty}^{+\infty} G_b(x)dx = 1. \tag{B.16}$$

Note that Eq. (B.17) also holds for complex b, as long as $\mathrm{Re}(b) > 0$.

b) The mean equals $\langle x \rangle = a$.

c) The variance is $\langle \Delta x^2 \rangle = b^2/2$.

Hint: use

[2] The rigorous formulation of this statement is called the *central limit theorem*.

$$\int_{-\infty}^{+\infty} \mathrm{e}^{-x^2/b^2}\,dx = b\sqrt{\pi}; \tag{B.17}$$

$$\int_{-\infty}^{+\infty} x^2 \mathrm{e}^{-x^2/b^2}\,dx = \frac{b^3\sqrt{\pi}}{2}. \tag{B.18}$$

Appendix C
Optical polarization tutorial

C.1 Polarization of light

Consider a classical electromagnetic plane wave propagating along the (horizontal) z-axis with angular frequency ω and wavenumber $k = \omega/c$, where c is the speed of light. The electromagnetic wave is transverse, so its electric field vector lies in the x-y plane:

$$\vec{E}(z,t) = A_H \hat{\imath} \cos(kz - \omega t + \varphi_H) + A_V \hat{\jmath} \cos(kz - \omega t + \varphi_V), \tag{C.1}$$

or in the complex form

$$\vec{E}(z,t) = \mathrm{Re}[(A_H e^{i\varphi_H} \hat{\imath} + A_V e^{i\varphi_V} \hat{\jmath}) e^{ikz - i\omega t}]. \tag{C.2}$$

Here $\hat{\imath}$ and $\hat{\jmath}$ are unit vectors along the x and y axes, respectively; A_H and A_V are the real amplitudes of the x and y components (which we will refer to as *horizontal* and *vertical*), and φ_H and φ_V are their phases.

Exercise C.1.§ Show that Eqs. (C.1) and (C.2) are equivalent.

The intensity of light in each polarization is proportional to:

$$I_H \propto A_H^2; \tag{C.3a}$$

$$I_V \propto A_V^2. \tag{C.3b}$$

The total intensity of the wave is the sum of its two components: $I_{\text{total}} \propto A_H^2 + A_V^2$.

Let us study the behavior of the electric field vector at some specific point in space, say $z = 0$. If the two components of the field have different phases, $\vec{E}(z,t)$ will change its direction as a function of time, as illustrated in Fig. C.1. To understand this interesting phenomenon better, try the following exercise.

Exercise C.2. Plot, as a function of time, the horizontal and vertical components of $\vec{E}(0,t)$ for $0 \le \omega t \le 2\pi$, in the following cases:

a) $A_H = 1$ V/m, $A_V = 0$, $\varphi_H = \varphi_V = 0$;

A. I. Lvovsky, *Quantum Physics*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-662-56584-1

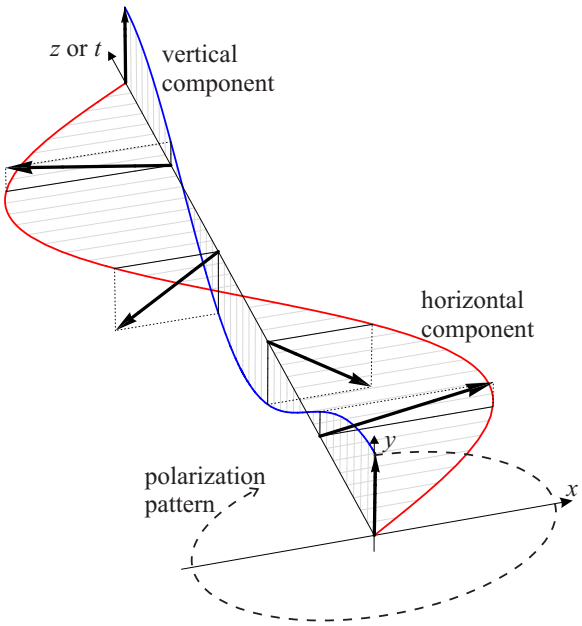


Fig. C.1 Polarization pattern of a plane wave. When the vertical and horizontal components of the electric field vector oscillate with different phases, the direction of that vector (shown with thick arrows) does not remain constant in phase and time. The trajectory of the tip of that vector determines the polarization pattern.

b) $A_H = 5$ V/m, $A_V = -3$ V/m, $\varphi_H = \varphi_V = 0$;
c) $A_H = 5$ V/m, $A_V = -3$ V/m, $\varphi_H = \pi/2$, $\varphi_V = 0$;
d) $A_H = 5$ V/m, $A_V = -3$ V/m, $\varphi_H = \pi/4$, $\varphi_V = -\pi/4$;
e) $A_H = 5$ V/m, $A_V = -3$ V/m, $\varphi_H = 0$, $\varphi_V = \pi/6$.

For each of the above cases, plot the trajectory of the point (E_x, E_y) for a constant z as a function of time.

The field vector trajectory defines the so-called *polarization state (pattern)* of light. The polarization state is one of the primary parameters of an electromagnetic wave; it determines how this field interferes with other waves or interacts with matter. Importantly, the polarization pattern is conserved as the wave is propagating through space and time, with the exception of certain materials which we will study a bit later.

Exercise C.3. Show that the polarization pattern of a plane wave is the same for all values of z.

This can be restated more generally: adding an arbitrary shift to both phases φ_H and φ_V will not change the polarization pattern. One can say that the pattern depends not on the individual phases of its two components, but on their difference $\varphi_H - \varphi_V$

[see Ex. C.2(c,d) for an example]. This property of classical polarization patterns has a direct counterpart in the quantum world: applying an overall phase shift to a quantum state does not change its physical properties (see Sec. 1.3 for a more detailed discussion).

In general, the polarization pattern is elliptical; however, as we have seen above, there exist special cases when the ellipse collapses into a straight line or blows out into a circle. Let us look at these cases more carefully.

Exercise C.4. Show the following:

a) The polarization pattern is linear if and only if $\varphi_H = \varphi_V + m\pi$, where m is an integer, or $A_H = 0$ or $A_V = 0$. The angle θ of the field vector with respect to the x axis is given by $\tan\theta = A_V/A_H$.
b) The polarization pattern is circular if and only if $\varphi_H = \varphi_V \frac{\pi}{2} + m\pi$, where m is an integer, and $A_H = \pm A_V$.

Important specific cases of linear polarization are horizontal ($A_V = 0$), vertical ($A_H = 0$), and $\pm 45^\circ$ ($A_V = \pm A_H$). For circular polarization, one can distinguish two cases according to the helicity of the wave: right and left circular.

- For right circular polarization, $A_V = A_H$ and $\varphi_V = \varphi_H + \frac{\pi}{2} + 2\pi m$ or $A_V = -A_H$ and $\varphi_V = \varphi_H - \frac{\pi}{2} + 2\pi m$, where m is an integer.
- For left circular polarization, $A_V = A_H$ and $\varphi_V = \varphi_H - \frac{\pi}{2} + 2\pi m$ or $A_V = -A_H$ and $\varphi_V = \varphi_H + \frac{\pi}{2} + 2\pi m$, where m is an integer[1].

Exercise C.5.* Show that, when none of the conditions of Ex. C.4 are satisfied, the tip of the electric field vector follows an elliptical pattern.

C.2 Polarizing beam splitter

The *polarizing beam splitter (PBS)* (Fig. C.2) is an important optical device for analyzing polarization. It is a transparent cube consisting of two triangular prisms glued to each other, constructed to transmit horizontally polarized light, but reflect vertically polarized. If a classical wave (C.2) is incident on such a beam splitter, the intensities of the transmitted and reflected light will be proportional to A_H^2 and A_V^2, respectively.

[1] Defining what circular polarization pattern should be called "left" or "right" is a matter of convention. Here we follow the convention that is common in the quantum optics community. In the right-circular pattern, the trajectory of the electric field vector is clockwise when viewed from the "back" of the wave (from the source). However, it is counterclockwise when viewed from the "front", or in the x-y plane with the traditional axis orientation. In space, this trajectory has the shape of a *left-handed* screw.

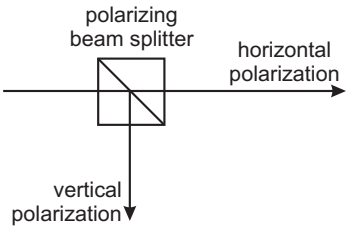


Fig. C.2 Polarizing beam splitter.

C.3 Waveplates

It is sometimes necessary to change the polarization state of light without splitting the vertical and horizontal components spatially. This is normally achieved using an optical instrument called a *waveplate*. The waveplate relies on *birefringence*, or *double refraction* — an optical property displayed by certain materials, primarily crystals, for example quartz or calcite. Birefringent crystals have an anisotropic structure, such that a light wave propagating through them will not conserve its polarization pattern unless it is linearly polarized along one of the two directions: either along or perpendicular to the crystal's *optic axis*. Traditionally, these directions are referred to as *extraordinary* and *ordinary*, respectively.

A birefringent material exhibits different indices of refraction for these two polarizations. Therefore, after propagation through the crystal, the ordinary and extraordinary waves will acquire different phases: $\Delta\varphi_e$ and $\Delta\varphi_o$, respectively. Because an overall phase shift has no effect on the polarization pattern, the quantity of interest is the difference $\delta\varphi = \Delta\varphi_e - \Delta\varphi_o$.

Exercise C.6. The indices of refraction for light polarized along and perpendicular to the optic axis are n_e and n_o, respectively, the length of the crystal is L, and the wavelength in vacuum is λ. Find $\delta\varphi$.

A waveplate is a birefringent crystal of a certain length, so $\Delta\varphi$ is precisely known. Two kinds of waveplates are manufactured commercially: $\lambda/2$*-plate (half-wave plate)* with $\delta\varphi = \pi$ and $\lambda/4$*-plate (quarter-wave plate)* with $\delta\varphi = \pi/2$.

If the polarization pattern is not strictly ordinary or extraordinary, propagation through a birefringent crystal will transform it. In order to determine this transformation, we decompose the wave into the extraordinary and ordinary components. The phase shift of each component is known. Knowing the new phases of both components, we can combine them to find the new polarization pattern.

Exercise C.7. For each of the polarization states of Ex. C.2, plot the polarization patterns that the waves will acquire when they propagate through (a) a half-wave plate, (b) a quarter-wave plate with the optical axes oriented vertically.

Solving the above exercise, you may have noticed that the half-wave plate "flips" the polarization pattern around the vertical (or horizontal) axis akin to a mirror.

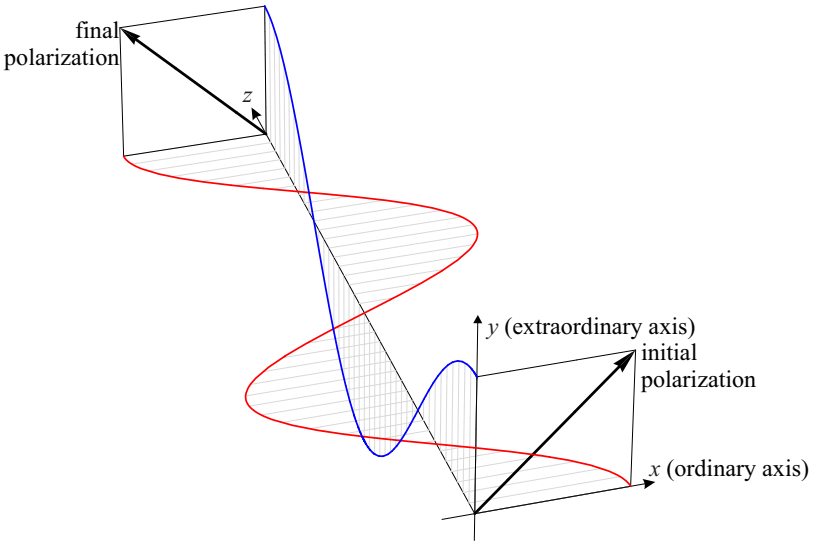


Fig. C.3 Action of a $\lambda/2$-plate with optic axis oriented vertically. Different refractive indices for the ordinary and extraordinary polarizations result in different optical path lengths, thereby rotating the polarization axis.

This is not surprising: the phase shift of π in the vertical component is equivalent to multiplication of A_V by -1. Of course, this mirroring property applies, not only when the optic axis is oriented vertically, but for *any* orientation, making the half-wave plate a universal tool for rotating the polarization of an electromagnetic field. Specifically, a light wave that is linearly polarized at angle θ to the horizontal, after propagating through a half-wave plate with its optic axis oriented at angle α to the horizontal, will transform into a linearly polarized wave at angle $2\alpha - \theta$ (Fig. C.4).

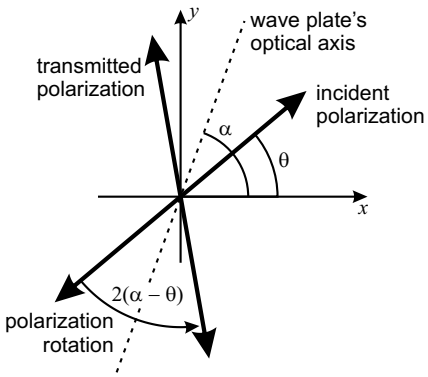


Fig. C.4 Polarization rotation by a $\lambda/2$ plate.

Exercise C.8.[§] Show that a $\lambda/2$-plate with the optic axis oriented at 22.5° to the horizontal interconverts between the horizontal and 45° polarizations, as well as between the vertical and -45° polarizations.

However, rotations alone do not provide a full set of possible transformations. For example, half-wave plates cannot transform between linear and circular/elliptical patterns. To accomplish this, we would need a quarter-wave plate.

Exercise C.9. Show that a $\lambda/4$-plate with the optic axis oriented horizontally or vertically interconverts between the circular and $\pm 45^\circ$ polarizations.

Exercise C.10. Linearly polarized light at angle θ to the horizontal propagates through a $\lambda/4$-plate with the optic axis oriented vertically. For the resulting elliptical pattern, find the angle between the major axis and the horizontal and the ratio of the minor to major axes.

Exercise C.11.[*] Suppose you have a source of horizontally polarized light. Show that, by using one half-wave plate and one quarter-wave plate, you can obtain light with an arbitrary polarization pattern.
Hint: It is easier to tackle this problem using geometric arguments, particularly the result of Ex. C.5, rather than formal algebra.

Exercise C.12.[*] Linearly polarized light propagates through a half-wave plate, then through a quarter-wave plate at angle 45° to the horizontal, then through a polarizing beam splitter. Show that the transmitted intensity does not depend on the angle of the half-wave plate.

Appendix D
Dirac delta function and the Fourier transformation

D.1 Dirac delta function

The delta function can be visualized as a Gaussian function (B.15) of infinitely narrow width b (Fig. B.5):

$$G_b(x) = \frac{1}{b\sqrt{\pi}} \mathrm{e}^{-x^2/b^2} \to \delta(x) \qquad \text{for} \qquad b \to 0. \tag{D.1}$$

The delta function is used in mathematics and physics to describe density distributions of infinitely small (*singular*) objects. For example, the position-dependent density of a one-dimensional particle of mass m located at $x = a$, can be written as $m\delta(x-a)$. Similarly, the probability density of a continuous "random variable" that takes on a certain value $x = a$ is $\delta(x-a)$. In quantum mechanics, we use $\delta(x)$, for example, to write the wave function of a particle that has a well-defined position.

The notion of function in mathematics refers to a map that relates a number, x, to another number, $f(x)$. The delta function is hence not a function in the traditional sense: it maps all $x \neq 0$ to zero, but $x = 0$ to infinity, which is not a number. It belongs to the class of so-called *generalized functions*. A rigorous mathematical theory of generalized functions can be found in most mathematical physics textbooks. Here, we discuss only those properties of the delta function that are useful for physicists.

Exercise D.1. Show that, for any smooth[1], bounded function $f(x)$,

$$\lim_{b\to 0} \int_{-\infty}^{+\infty} G_b(x) f(x) \mathrm{d}x = f(0). \tag{D.2}$$

From Eqs. (D.1) and (D.2) and for any smooth function $f(x)$, we obtain

[1] A *smooth* function is one that has derivatives of all finite orders.

A. I. Lvovsky, *Quantum Physics*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-662-56584-1

$$\int_{-\infty}^{+\infty} \delta(x) f(x) \mathrm{d}x = f(0) \tag{D.3}$$

This property is extremely important because it allows one to perform meaningful calculations with the delta function in spite of its singular nature. Although the delta function does not have a numerical value for all values of its argument, the integral of the delta function multiplied by another function does. We may write a delta function outside of an integral, but we always keep in mind that it will eventually become a part of an integral, and only then will it produce a quantitative value — for example, a prediction of an experimental result.

In fact, Eq. (D.3) can be viewed as a rigorous mathematical definition of the delta function. Using this definition, we can obtain its other primary properties.

Exercise D.2. Show that

a)
$$\int_{-\infty}^{+\infty} \delta(x) \mathrm{d}x = 1; \tag{D.4}$$

b) for any function $f(x)$,

$$\int_{-\infty}^{+\infty} \delta(x-a) f(x) \mathrm{d}x = f(a); \tag{D.5}$$

c) for any real number a,
$$\delta(ax) = \delta(x)/|a|. \tag{D.6}$$

Exercise D.3. For the Heaviside step function

$$\theta(x) = \begin{cases} 0 \text{ if } x < 0 \\ 1 \text{ if } x \geq 0 \end{cases}, \tag{D.7}$$

show that

$$\frac{\mathrm{d}}{\mathrm{d}x}\theta(x) = \delta(x). \tag{D.8}$$

Hint: use Eq. (D.3).

Exercise D.4. Show that, for any $c < 0$ and $d > 0$,

$$\int_{c}^{d} \delta(x) \mathrm{d}x = 1 \tag{D.9}$$

D.2 Fourier transformation

Definition D.1. The *Fourier transform* $\tilde{f}(k) \equiv \mathscr{F}[f](k)$ of a function $f(x)$ is a function of the parameter k defined as follows:[2]

$$\tilde{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{e}^{-\mathrm{i}kx} f(x)\mathrm{d}x. \tag{D.10}$$

This is an important integral transformation used in all branches of physics. Suppose, for example, that you have a light wave of the form $f(\omega)\mathrm{e}^{-\mathrm{i}\omega t}$, where ω is the frequency and $f(\omega)$ is the complex amplitude, or the *frequency spectrum* of the signal. Then the time-dependent signal from all sources is $\int_{-\infty}^{+\infty} f(\omega)\mathrm{e}^{-\mathrm{i}\omega t}\mathrm{d}\omega$ — that is, the Fourier transform of the spectrum. The power density of the spectrum, i.e., the function $|f(\omega)|^2$, can be measured experimentally by means of a dispersive optical element, such as a prism.

Exercise D.5. Show that, if $\tilde{f}(k) = \mathscr{F}[f(x)]$ exists, then

a)
$$\tilde{f}(0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x)\mathrm{d}x; \tag{D.11}$$

b) for a real $f(x)$, $\tilde{f}(-k) = \tilde{f}^*(k)$;

c) for $a \neq 0$,
$$\mathscr{F}[f(ax)] = \frac{1}{|a|}\tilde{f}(k/a); \tag{D.12}$$

d)
$$\mathscr{F}[f(x-a)] = \mathrm{e}^{-\mathrm{i}ka}\tilde{f}(k); \tag{D.13}$$

e)
$$\mathscr{F}[\mathrm{e}^{\mathrm{i}\xi x} f(x)] = \tilde{f}(k-\xi); \tag{D.14}$$

f) assuming that $f(x)$ is a smooth function approaching zero at $\pm\infty$,
$$\mathscr{F}[\mathrm{d}f(x)/\mathrm{d}x] = \mathrm{i}k\tilde{f}(k). \tag{D.15}$$

Exercise D.6. Show that the Fourier transform of a Gaussian function is also a Gaussian function:
$$\mathscr{F}[\mathrm{e}^{-x^2/b^2}] = \frac{b}{\sqrt{2}}\mathrm{e}^{-k^2b^2/4}. \tag{D.16}$$

We see from Eq. (D.12) that scaling the argument x of a function results in inverse scaling of the argument k of its Fourier transform. In particular (Ex. D.6), a

[2] There is no common convention as to whether to place the negative sign in the complex exponent of Eqs. (D.10) or (D.21), nor how to distribute the factor of $1/2\pi$ between them. Here I have chosen the convention arbitrarily.

signal with a Gaussian spectrum of width b is a Gaussian pulse of width $2/b$, so the product of the two widths is a constant. This is a manifestation of the *time-frequency uncertainty* that applies to a wide range of wave phenomena in classical physics. In fact, as we see in Sec. 3.3.2, in its application to the position and momentum observables, the Heisenberg uncertainty principle can also be interpreted in this fashion.

Let us now consider two extreme cases of the Fourier transform of Gaussian functions.

Exercise D.7. Show that:

a) in the limit $b \to 0$, Eq. (D.16) takes the form

$$\mathscr{F}[\delta(x)] = \frac{1}{\sqrt{2\pi}}; \tag{D.17}$$

b) in the opposite limit, $b \to \infty$, one obtains

$$\mathscr{F}[1] = \sqrt{2\pi}\,\delta(k). \tag{D.18}$$

If the spectrum contains only the zero frequency, the signal, not surprisingly, is time-independent. If, on the other hand, the spectrum is constant, the signal is an instant "flash" occurring at $t = 0$. Here is an interesting consequence of this observation.

Exercise D.8. Show that, for $a \neq 0$,

$$\int_{-\infty}^{+\infty} \mathrm{e}^{\mathrm{i}akx}\mathrm{d}x = 2\pi\delta(k)/|a|. \tag{D.19}$$

This result is of paramount importance for many calculations involving the Fourier transform. We will see its utility shortly.

Exercise D.9. Assuming a and b to be real and positive, find the Fourier transforms of the following:

a) $\delta(x+a)+\delta(x-a)$
b) $\cos(ax+b)$;
c) $\mathrm{e}^{-ax^2}\cos bx$;
d) $\mathrm{e}^{-a(x+b)^2}+\mathrm{e}^{-a(x-b)^2}$;
e) $\theta(x)\mathrm{e}^{-ax}$, where $\theta(x)$ is the Heaviside function;
f) a "top-hat function" $\begin{cases} 0 \text{ if } x < -a \text{ or } x > a; \\ A \text{ if } -a \le x \le a; \end{cases}$.

The Fourier transform can be inverted: for any given time-dependent pulse one can calculate its frequency spectrum such that the pulse is the Fourier transform of that spectrum. Remarkably, the Fourier transform is very similar to its inverse. This similarity can be observed, for example, by comparing Eqs. (D.13) and (D.14). Displacing the argument of $f(x)$ leads to the multiplication of $\tilde{f}(k)$ by a complex

Box D.1 Interpreting Eq. (D.8)

The result (D.8) seems to tell us that the integral $\int_{-\infty}^{+\infty} e^{ikx} dx$ equals zero for $k \neq 0$. This does not reconcile with traditional calculus, according to which the integral of a finite oscillating function e^{ikx} must diverge for any k. To address this apparent inconsistency, we need to remember that Eq. (D.19) is valid only as a generalized function — that is, as a part of the integral (D.3). Indeed, if we substitute Eq. (D.19) into Eq. (D.3), we obtain a convergent integral

$$\int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} e^{ikx} dx \right] f(k) dk = \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} e^{ikx} f(k) dk \right] dx = \sqrt{2\pi} \int_{-\infty}^{+\infty} \mathscr{F}[f](-k) dk \overset{(D.11)}{=} 2\pi f(0). \tag{D.20}$$

Therefore, while the numerical value of the integral (D.19) for any specific k does not exist, it is meaningfully defined as a generalized function of k.

phase. On the other hand, if we multiply $f(x)$ by a complex phase, the argument of $\tilde{f}(k)$ gets shifted.

Definition D.2. The *inverse Fourier transform* $\mathscr{F}^{-1}[g](x)$ of the function $g(k)$ is a function of the parameter x such that

$$\mathscr{F}^{-1}[g](x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{ikx} g(k) dk. \tag{D.21}$$

Exercise D.10. Show that

$$\mathscr{F}^{-1}[\mathscr{F}[f]](x) = f(x). \tag{D.22}$$

Exercise D.11. Show that

$$\mathscr{F}^{-1}[f(x)](k) = \mathscr{F}[f(x)](-k) = \mathscr{F}[f(-x)](k). \tag{D.23}$$

Exercise D.12.§ Derive the analogues of the rules found in Ex. D.5 for the inverse Fourier transform.
Answer: Denoting $\breve{g}(x) = \mathscr{F}^{-1}[g(k)]$,

a)

$$\breve{g}(0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} g(k) dk; \tag{D.24}$$

b) for a real $g(k)$, $\breve{g}(x) = \breve{g}^*(-x)$;
c)

$$\mathscr{F}^{-1}[g(ak)](x) = \frac{1}{|a|} \breve{g}(k/a); \tag{D.25}$$

d)

$$\mathscr{F}^{-1}[g(k-a)](x) = \mathrm{e}^{\mathrm{i}xa}\breve{g}(k); \tag{D.26}$$

e)

$$\mathscr{F}^{-1}[\mathrm{e}^{\mathrm{i}\xi k}g(k)](x) = \breve{g}(x+\xi); \tag{D.27}$$

f)

$$\mathscr{F}^{-1}[dg(k)/dk] = -ix\breve{g}(x). \tag{D.28}$$

Index

A. I. Lvovsky, *Quantum Physics*, Undergraduate Lecture Notes in Physics, https://doi.org/10.1007/978-3-662-56584-1