Quirky Quantum Concepts

The Anti-Textbook*

By Eric L. Michelsen

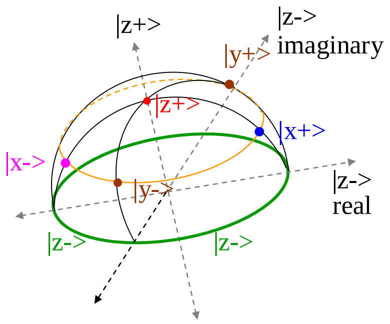
Manuscript Draft Excerpt

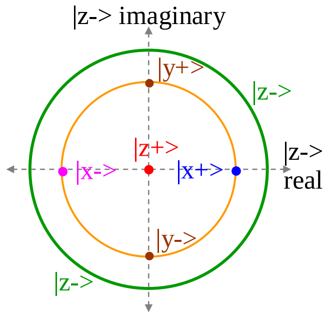
I hope this manuscript draft excerpt will encourage you to buy the final book:

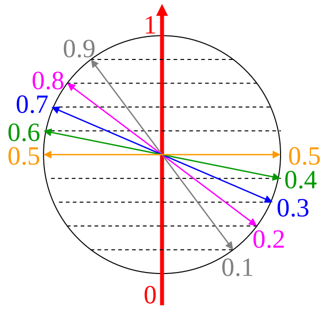
Locally at the UCSD bookstore: Quirky Quantum Concepts

From Springer: Quirky Quantum Concepts

From Amazon: Quirky Quantum Concepts







> "Quantum Mechanics is a silly theory, perhaps the silliest theory to come out of the 20th century. The only reason it has any following at all is that it is completely supported by experiment."
> − *Unknown physicist*

> "We are all agreed that your theory is crazy. The question that divides us is whether it is crazy enough to have a chance of being correct." − *Niels Bohr*

> "Now in the further development of science, we want more than just a formula. First we have an observation, then we have numbers that we measure, then we have a law which summarizes all the numbers. But the real *glory* of science is that we can *find a way of thinking* such that the law is *evident*." − *Richard Feynman*

* Physical, conceptual, geometric, and pictorial physics that didn't fit in your textbook.

Physical constants: 2006 values from NIST. For more, see http://physics.nist.gov/cuu/Constants/ .

Speed of light in vacuum	$c = 299\ 792\ 458$ m s^{-1} (exact)
Gravitational constant	G = 6.674 28(67) x 10–11 m^{3} kg^{-1} s^{-2}
Relative standard uncertainty	$\pm 1.0 \times 10^{-4}$
Boltzmann constant	$k = 1.380\ 6504(24) \times 10^{-23}$ J K^{-1}
Stefan-Boltzmann constant	$\sigma = 5.670\ 400(40) \times 10^{-8}$ W m^{-2} K^{-4}
Relative standard uncertainty	$\pm 7.0 \times 10^{-6}$
Avogadro constant	$N_A, L = 6.022\ 141\ 79(30) \times 10^{23}$ mol^{-1}
Relative standard uncertainty	$\pm 5.0 \times 10^{-8}$
Molar gas constant	$R = 8.314\ 472(15)$ J mol^{-1} K^{-1}
calorie	4.184 J (exact)
Electron mass	$m_e = 9.109\ 382\ 15(45) \times 10^{-31}$ kg
Proton mass	$m_p = 1.672\ 621\ 637(83) \times 10^{-27}$ kg
Proton/electron mass ratio	$m_p/m_e = 1836.152\ 672\ 47(80)$
Elementary charge	$e = 1.602\ 176\ 487(40) \times 10^{-19}$ C
Electron g-factor	$g_e = -2.002\ 319\ 304\ 3622(15)$
Proton g-factor	$g_p = 5.585\ 694\ 713(46)$
Neutron g-factor	$g_N = -3.826\ 085\ 45(90)$
Muon mass	$m_\mu = 1.883\ 531\ 30(11) \times 10^{-28}$ kg
Inverse fine structure constant	$\alpha^{-1} = 137.035\ 999\ 679(94)$
Planck constant	$h = 6.626\ 068\ 96(33) \times 10^{-34}$ J s
Planck constant over 2π	$\hbar = 1.054\ 571\ 628(53) \times 10^{-34}$ J s
Bohr radius	$a_0 = 0.529\ 177\ 208\ 59(36) \times 10^{-10}$ m
Bohr magneton	$\mu_B = 927.400\ 915(23) \times 10^{-26}$ J T^{-1}

Table of Contents

0 Front/Back Matter

0.1 Back Cover

Quirky Quantum Concepts explains the more important and more difficult concepts in theoretical quantum mechanics, especially those which are consistently neglected or confusing in many common expositions. The emphasis is on physical understanding, which is necessary for the development of new, cutting edge science.

Quirky Quantum Concepts is therefore a supplement to almost any existing quantum mechanics text. Students and scientists will appreciate the combination of conversational style, which promotes understanding, with thorough scientific accuracy. The book is not a simplification or popularization: it is real science for real scientists. Physics includes math, and this book does not shy away from it, but neither does it hide behind it. Without conceptual understanding, math is gibberish. In particular, this book explains the basis for many standard quantum methods, which are too often presented without sufficient motivation or interpretation. The discussions here provide the experimental and theoretical reasoning behind some of the great discoveries, so the reader may see how discoveries arise from a rational process of thinking, a process which *Quirky Quantum Concepts* makes accessible to its readers.

0.2 Dedication

To my wife, Laura, for all her patience and support, and to my children, Sarah and Ethan, for understanding of my absences while I was working on the book.

0.3 Preface

Why Quirky?

The purpose of the "Quirky" series is to help develop an accurate physical, conceptual, geometric, and pictorial understanding of important physics topics. We focus on areas that don't seem to be covered well in most texts. The Quirky series attempts to clarify those neglected concepts, and others that seem likely to be challenging and unexpected (quirky?). The Quirky books are intended for serious students of physics; they are not "popularizations" or oversimplifications.

Physics includes math, and we're not shy about it, but we also don't hide behind it.

Without a conceptual understanding, math is gibberish.

We seek to be accurate, but not pedantic. When mathematical or physical words have precise meanings, we adhere to those meanings. Words are the tools of communication; it is impossible to make fine points with dull tools.

Who Is It For?

This work is one of several aimed at graduate and advanced-undergraduate physics students, engineers, scientists, and anyone else who wants a serious understanding of Quantum Mechanics. The material ranges from fairly elementary (though often neglected) to graduate level. Go to http://physics.ucsd.edu/~emichels for the latest versions of the Quirky Series, and for contact information. We're looking for feedback, so please let us know what you think.

How to Use This Book

This book is an informal, topical review. We strive to be accurate, but not tedious.

This work is not a text book.

There are plenty of those, and they cover most of the topics quite well. This work is meant to be used *with* a standard text, to help emphasize those things that are most confusing for new students. When standard presentations don't make sense, come here. In short, our goal is to provide the foundation that enables future learning from other sources.

If you have some background in quantum mechanics, then most of the sections stand alone. The larger sections start by naming the prerequisites needed to understand that section. This work is deliberately somewhat redundant, to make sections more independent, and because learning requires repetition.

You should read all of Chapter 1, Basic Wave Mechanics Concepts, to familiarize yourself with the notation and contents. After the first two chapters, this book is meant to be read in any order that suits you. Each section stands largely alone, though the sections are ordered logically. You may read it from beginning to end, or skip around to whatever topic is most interesting. The "Desultory" chapter is a diverse set of short topics, each meant for quick reading.

We must necessarily sometimes include forward references to material which has not yet been covered in this book. If they are unfamiliar, most such references may be ignored without loss of continuity.

If you don't understand something, read it again *once*, then keep reading.
Don't get stuck on one thing. Often, the following discussion will clarify things.

Scope

What This Text Covers

This text covers most of the unusual or challenging concepts in a first-year graduate course in non-relativistic Quantum Mechanics (QM). Much of it is suitable for undergraduate QM, as well, because it provides a conceptual foundation for all of QM. We expect that you are taking or have taken such a QM course, and have a good text book. This text supplements those other sources.

What This Text Doesn't Cover

This text is not a QM course in itself, nor a review of such a course. We do not cover all basic QM concepts; only those that are unusual or especially challenging (quirky?). There is almost no relativistic QM here.

What You Already Know

This text assumes you understand basic integral and differential calculus, partial differential equations, have seen complex numbers, and have some familiarity with probability. You must have a working knowledge of basic physics: mass, force, momentum, energy, etc. Further, it assumes you have a Quantum Mechanics text for the bulk of your studies, and are using *Quirky Quantum Concepts* to supplement it. You must have been introduced to the idea of particles as waves, and photons as particles of light. Beyond that, different sections require different levels of preparation; some are much more advanced than others. Each section lists any particular prerequisites at the beginning. Some sections require some familiarity with classical Lagrangian and Hamiltonian mechanics, including canonical momentum.

Notation

Important points are highlighted in solid-border boxes.

Common misconceptions are sometimes written in dashed-line boxes.

References: As is common, we include references to published works in square brackets, where the abbreviations in the brackets are defined in the "References" section of this document. Where page numbers are given, they may be augmented by "t", "m", or "b", for "top", "middle", and "bottom" of the page.

Unit vectors: We use $\mathbf{e}_x$, $\mathbf{e}_y$, etc. for unit spatial vectors. In other disciplines, these are more likely written as $\hat{\mathbf{x}}, \hat{\mathbf{y}}$, etc., but we reserve "hats" to indicate quantum operators.

Keywords are listed in **bold** near their definitions. All keywords also appear in the glossary.

We use the following symbols regularly:

Symbol	Name	Common scientific meanings
∇	del	gradient of a scalar field, divergence or curl of a vector field.
$\equiv$	identity	(1) is defined as; (2) identically (always) equal to.
$\forall$	for all	for all.
$\in$	element	is an element of (sometimes written as epsilon, ε).
$\approx$	approximately equals	approximately equals.
~	tilde	scales like; is proportional to in some limit. For example, $1/(r + 1) \sim 1/r$ for large r because $\lim_{r\to\infty} \frac{1}{r+1} = \frac{1}{r}$. Note that we want to preserve the scaling property, so we don't take such a limit all the way to $r \to \infty$ (limit of zero), which would hide any r dependence.
!!	double factorial	$n!! \equiv n(n-2)(n-4) \ldots (2 \text{ or } 1)$.

Integrals: In many cases, we present a general integral, whose exact limits depend on what problem or coordinate system the integral is applied to. To provide a general formula, independent of such particulars, we give the limits of integration as a single "∞", meaning integrate over the entire domain appropriate to the problem:

$$\int_\infty f(x)\, dx \equiv \text{integrate over entire domain relevant to the given problem} .$$

Open and closed intervals: An **open interval** between c and d is written "(c, d)", and means the range of numbers from c to d *exclusive* of c and d. A **closed interval** between c and d is written "$[c, d]$", and means the range of numbers from c to d *including* c and d. A half-open interval "$[c, d)$" has the expected meaning of c to d including c but not d, and "$(c, d]$" means c to d excluding c but including d.

Operators: I write most operators with a "hat" over them, e.g. $\hat{x}$. Rarely, the hat notation is cumbersome, so I sometimes use the subscript $_{op}$ to denote quantum operators, as in [P&C]. Thus the symbol x is a real variable, $\hat{x}$ is the position operator, and $\left(p^2\right)_{op}$ is the operator for p^2.

Conjugates and Adjoints: We use "*" for complex conjugation, and "†" for adjoint:

$$z^* \equiv \text{complex conjugate of the number 'z'}, \quad \hat{a}^\dagger \equiv \text{adjoint operator of } \hat{a} .$$

[Note that some math texts use a bar for conjugate: $\bar{a} \equiv$ complex conjugate of 'a', and a "*" for adjoint. This is confusing to physicists, but c'est la vie.]

> [Interesting paragraphs that may be skipped are "asides," shown in square brackets, smaller font, and narrowed margins.]

[Short asides may be also be written in-line in square brackets.]

Vector variables: In some cases, to emphasize that a variable is a vector, it is written in bold; e.g., $V(\mathbf{r})$ is a scalar function of the vector, $\mathbf{r}$. $\mathbf{E}(\mathbf{r})$ is a vector function of the vector, $\mathbf{r}$. We write a zero vector as $\mathbf{0}_\mathbf{v}$ (this is different than the number zero).

Matrices: Matrices are in bold, **B**. A particular element of a single matrix may be specified with subscripts, e.g. B_{ij}. A particular element of a matrix expression uses brackets, e.g. $[\mathbf{AB}]_{ij} \equiv$ the ij^{th} element of the matrix product **AB**.

Tensor products: Sometimes, we write a tensor product explicitly with the ⊗ symbol.

In-line derivatives sometimes use the notation d/dx and $\partial/\partial x$. There is not always a clear mathematical distinction between d/dx and $\partial/\partial x$. When the function arguments are independent, they are both the same thing. I use d/dx when a function is clearly a total derivative, and $\partial/\partial x$ when it is clearly a partial derivative. However, in some cases, it's not clear what arguments a function has, and it's not important. In that case, I tend to use $\partial/\partial x$ for generality, but don't worry about it.

Also, for the record, derivatives *are* fractions, despite what you might have been told in calculus. They are a special case of fraction: the limiting case of fractions of differentially small changes. But they are still fractions, with all the rights and privileges thereof. All of physics treats them like fractions, multiplies and divides them like fractions, etc., because they *are* fractions.

Greek Letters

The Greek alphabet is probably the next-best well known alphabet (after our Latin alphabet). But Greek letters are often a stumbling block for readers unfamiliar with them. So here are all the letters, their pronunciations, and some common meanings from all over physics. Note that every section defines its own meanings for letters, so look for those definitions.

The Greek alphabet has 24 letters, and each has both upper-case (capital) and lower-case forms. Not all can be used as scientific symbols, though, because some look identical to Latin letters. When both upper- and lower-case are useable, the lower-case form is listed first. Lower case Greek variables are italicized, but by convention, upper case Greek letters are not. Don't worry if you don't understand all the common meanings; we'll define as we go everything you need to know for this book.

Letter	Name (pronunciation)	Common scientific meanings
α	alpha (al'fuh)	coefficient of linear thermal expansion. (Capital: A, not used.)
β	beta (bae'tuh)	velocity as a fraction of the speed of light ($\beta \equiv v/c$). (Capital: B, not used).
γ	gamma (gam'uh)	the relativistic factor $(1-\beta^2)^{-1/2}$, aka time-dilation/length-contraction factor.
Γ	capital gamma	Christoffel symbols (General Relativity); generalized factorial function.
δ	delta (del'tuh)	the Dirac delta (impulse) function; the Kronecker delta; an inexact differential (calculus).
∂	old-style delta	partial derivative (calculus).
Δ	capital delta	a small change.
ε	epsilon (ep'si-lon)	a small value. (Capital: E, not used.)
ζ	zeta (zae'tuh)	damping ratio. (Capital: Z, not used.)
η	eta (ae'tuh)	efficiency; flat-space metric tensor. (Capital: H, not used.)
θ	theta (thae'tuh)	angle.
Θ	capital theta	not commonly used. Sometimes angle.
ι	iota (ie-o'tuh)	not commonly used. (Capital: I, not used.)
κ	kappa (kap'uh)	decay constant. (Capital: K, not used.)
λ	lambda (lam'duh)	wavelength.

Λ	capital lambda	cosmological constant.
μ	mu (mew)	micro (10^{-6}); reduced mass. (Capital: M, not used.)
ν	nu (noo)	frequency. Not to be confused with an italic v: v vs. nu: ν. (Capital: N, not used.)
ξ	xi (zie, sometimes ksee)	dimensionless distance measure.
Ξ	capital xi	not commonly used.
o	omicron (oe'mi-kron)	not used. (Capital: O, not used.)
π	pi (pie)	ratio of a circle's circumference to its diameter, ≈ 3.14159... .
Π	capital pi	product (multiplication).
ρ	rho (roe)	mass density; charge density; correlation coefficient. (Capital: P, not used.)
σ	sigma (sig'muh)	standard deviation; surface charge density.
Σ	capital sigma	sum (addition).
τ	tau (rhyme: cow, or sometimes saw)	time; torque. (Capital: T, not used.)
υ	upsilon (oops'i-lon)	not commonly used. (Capital: Y, not used.)
ϕ	phi (fee or fie)	angle.
φ	old-style phi	angle
Φ	capital phi	electric potential; general potential.
χ	chi (kie)	degrees of freedom. (Capital: X, not used.)
ψ	psi (sie)	wave-function amplitude.
Ψ	capital psi	not commonly used.
ω	omega (oe-mae'guh)	angular velocity; angular frequency.
Ω	capital omega	angle; solid angle; ohm (unit of electrical resistance).

0.4 Acknowledgements

Thank you, Herbert Shore, San Diego State University, Physics Department, for giving me a chance, and for being a good teacher. I am indebted to Daniel Arovas for his detailed comments. I also owe a big thank you to many professors at both SDSU and UCSD, for their generosity even before I was a real student: Peter Salamon, Arlette Baljon, Andrew Cooksy, George Fuller, Tom O'Neil, Terry Hwa, and many others. Thanks to Peter Betts, who first set the equations in MathType, and to Eric Williams, Yaniv Rosen, and Jason Leonard for their thorough comments, and to Kevin Duggento and Chad Kishimoto for their thorough discussions. Kim Griest gets a special mention for his unfailing moral support, which means so much. Also, all the students I have worked with over the years have helped, both knowingly and unknowingly, to make this work possible. As always, since I have collected all the input and put it into this form, any errors in the text are solely my own.

Please send any comments or corrections to me, at emichels@physics.ucsd.edu.

1 Basic Wave Mechanics Concepts

The goal of this chapter is to convey a conceptual and pictorial understanding of the workings of quantum mechanics. We introduce the experimental results that motivate quantum mechanics, turning them into a reliable, quantitative model that will underpin all of our future developments. This chapter is conceptually the most difficult, but it is essential.

1.1 The Meaning of Science

Quantum Theory is true. It is *not* a speculation, nor are the major parts of it in any doubt at all. In science, unlike ordinary conversation, a **theory** is the highest level of scientific achievement: a quantitative, predictive, testable model which unifies and relates a body of facts. A theory becomes accepted science only after being supported by overwhelming evidence. A theory is not a speculation, e.g. Maxwell's electromagnetic theory. Note that every generally accepted theory was, at one time, *not* generally accepted. A **fact** is a small piece of information backed by solid evidence (in hard science, usually repeatable evidence). If someone disputes a fact, it is still a fact. ("If a thousand people say a foolish thing, it is still a foolish thing.") A **speculation** is a guess, perhaps motivated by facts, perhaps by intuition.

We must be careful when we talk of what is "really" happening in quantum mechanics. Because QM is so far removed from everyday experience, there is no "really."

All we can hope for are mathematical models which predict the outcomes of experiments.
Any such model is valid (or "real").

Throughout this work, we adopt the simplest valid model we can. Other valid models exist. They may disagree on the wording or interpretation of the theory, but by definition, all valid models agree on known experimental results. When two models disagree on predictions that are not yet tested, then in principle, performing an experiment can falsify one model, and be consistent with the other. In general, it is not possible to prove that a theory is correct, only that it is consistent with all known facts (definitive experiments).

1.2 Not Your Grandfather's Physics: Quantum Weirdness

The following realistic (but here idealized) experiment demonstrates that quantum mechanics is not your everyday physics.

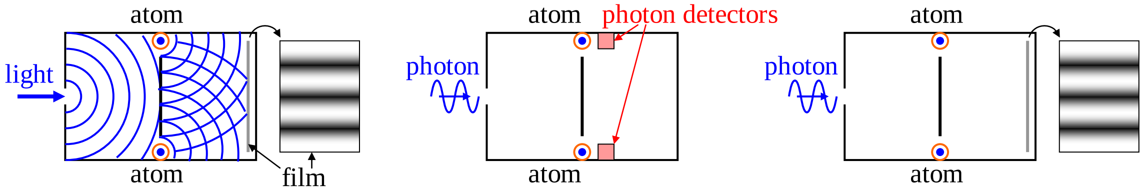


Figure 1.1 (Left) A bright light fluoresces from both atoms and exposes an interference pattern. (Middle) A single photon at a time fluoresces off one atom or the other, but never both. (Right) A large number of individual photons still produces interference on the film.

Consider a box (Figure 1.1, left), that admits light on the left, and has two identical atoms midway across, one each at the top and bottom. A black screen blocks light from passing in-between the atoms. On the right is photographic film which records the cumulative energy of light striking it at each point. The atoms fluoresce, i.e. they absorb photons, and quickly re-radiate new photons of fixed energy. If we shine a bright light into the hole, the atoms fluoresce, and the film records a standard interference pattern (dark indicates higher exposure). Classical physics predicts this.

If we now place photon detectors in front of each atom (Figure 1.1, middle), and send in a series of individual photons (one at a time), our detectors show that either one atom or the other fluoresces, but never both. Recall that all photons of a given frequency have the same energy. When an atom does fluoresce, it radiates a single photon with the full photon energy, just as when under a bright light. The energy is never split between the two atoms. This is quantum: an atom either absorbs and radiates, or it doesn't. The event is quantized: there is no half-way. That is perhaps unexpected, but not yet truly weird.

Next, we remove the detectors and replace the film (Figure 1.1, right). We again send in a series of individual photons (one at a time), and accumulate enough photons to expose the film. By extension of our previous single-photon experiment, we expect only one atom to radiate from each entering photon, and therefore we do not expect interference. However, the actual experimental result is an image of interference on the film.

It seems that the atoms behave differently when we are "looking at" them with detectors, than when we are not. Some physicists have called this "dual" behavior "quantum weirdness." This result says that quantum mechanics isn't just a modified set of laws from classical mechanics, but instead, the very nature of causality, time evolution, and observation are somehow different on the quantum scale. For example, this experiment goes far beyond saying "electrons are waves." Water travels in waves, too, but does not (macroscopically) exhibit this "weirdness."

Quantum mechanics is more than just a new "wavy model" of particles.
QM also implies that the fundamental nature of causality, time evolution, and observation is different at the quantum scale than at the everyday macroscopic scale.

This experiment is essentially the famous double-slit experiment (each atom is a slit), but with the added twist that the atoms seems to change their behavior when the observational setup changes, even though that setup does not directly impinge on the process of photon absorption and re-radiation: the observational setup only changes how we observe the results. A short time after a photon enters the box, we might think the two possible states of the system are "top atom is excited" or "bottom atom is excited." However, the recording of an interference pattern on the film indicates that the actual state is some kind of superposition of both.

This experiment is also directly analogous to the Schrödinger's cat experiment, wherein a radioactive element sits in a box with a cat. Over time, the element either decays or not, causing a mechanism to either kill the cat, or not. After a time interval, the two states of the system are "the cat is alive" and "the cat is dead." In principle, we could add film to the box, and a "life-detector:" If the cat is alive, it radiates a photon from the top; if the cat is dead, it radiates a photon from the bottom. After killing a large number of cats, the idealized experiment would produce an interference pattern on the film: each cat has indeed been in a superposition of "alive" and "dead."

The reason we cannot successfully perform the cat experiment has nothing to do with life, death, or consciousness. It has to do with the cat being macroscopic. We describe all these aspects of quantum mechanics in much more detail throughout this book.

Although the photon-atom-film experimental result demands a new mechanics, quantum mechanics, I think it is not truly the essence of QM. QM is more about quantization of energy, angular momentum, and other properties of microscopic systems, which have a long history of experimental verification. In many ways, these effects are more important, though perhaps less exciting, than the mind-bending nature of causality and observation. Either way, it is the quantization effects that consume the vast majority of the study of QM, and to which we devote the vast majority of this book.

1.3 The Job of Quantum Mechanics

What, then, does Quantum Mechanics do? Like any physical theory, it predicts the results of experiments. All other physical theories, given enough information, predict the exact result of an experiment. The unique feature of QM theory is that predicting one single result is not generally possible. The universe itself is probabilistic. The *exact* same experiment can produce *different* results when repeated.

Quantum mechanics tells what results are possible, and with what probabilities.

In QM talk: QM defines the "spectrum" of results, and their probabilities of occurrence. Therefore, by combining this information, you can also predict the average value of a measurement *A*. We'll have more on this later, but for now, recall some statistical facts about averages:

$$\langle A \rangle = \sum_{i=1}^{N} \Pr(a_i)\, a_i , \qquad \text{or} \qquad \langle A \rangle = \int_{-\infty}^{\infty} \text{pdf}(a)\, a\, da ,$$

where the *a* are possible values, Pr(a_i) is the probability of a discrete measurement, and pdf(a) is the probability distribution function of a continuous measurement.

1.3.1 The Premises of Quantum Mechanics

We introduce here some of the underlying principles of quantum mechanics, and then explore those principles more deeply in the "Axioms" section below.

Quantum mechanics says that, in general, systems exist in a linear superposition
of oscillating states, i.e. they exist simultaneously in many different states.

These superpositions have two important characteristics:

First: Each state in the superposition has its own mathematical weight. For example, a system might be 1/3 state A, 1/6 state B, and 1/2 state C. We say the system's "component states" are A, B, and C. Before a measurement, interactions occur with each component state acting independently, and as if it were the *whole* system. The resulting final state is a superposition of all the resulting component states.

For brevity, it is often said that taking an (ideal) measurement "collapses" the system to a state (possibly a superposition) consistent with the measurement. Often, the system then evolves into a broader superposition of states again, over time. We will see, however, that such a simple model cannot explain the full set of experimental observations, and leads to conceptual difficulties. A more complete model of measurements includes "decoherence," discussed later.

Second: The states oscillate sinusoidally (in some way). Therefore, each state has not only weight (or "magnitude"), but **phase**. The phase of a state describes where in its periodic oscillation the state is, at a given moment. States oscillate sinusoidally, and since sin( ) has period 2π, the phase is often given as an angle, between 0 and 2π.

Note There is a big difference between a "superposition" state, and a "mixed" state. We will describe the difference in more detail later, but for now: superpositions include phase information in the set of component states, and describe a *single* system (e.g., particle). "Mixed states" have no phase information between the constituent states, and describe one of a set (ensemble) of multiple systems (e.g. particles), each in a different state. This is essentially a classical concept, but with very non-classical (quantum) consequences. See "Density Matrices" later for more discussion.

1.3.2 The Success of Quantum Mechanics

There are many fundamental successes of quantum mechanics. The following rules of calculation follow from the axioms of linearity and superposition (described more fully below):

1. Dynamic quantities are described by complex numbers, called **amplitudes** (or "complex amplitudes"), which have a magnitude and a phase. (This is different than some other applications where "amplitudes" are real numbers that quantify only the size of a sinusoidal oscillation.)
2. The probability of an observation is proportional to the squared-magnitude of the amplitude for that observation.
3. When there are two or more paths from an initial state to a final state, the complex amplitudes for each path add as complex numbers, called adding **coherently**.
4. When there is a path from initial state A, through an intermediate state B, to a final state C, the amplitude for that path is the complex product of the amplitudes for transitions from A→B and B→C.

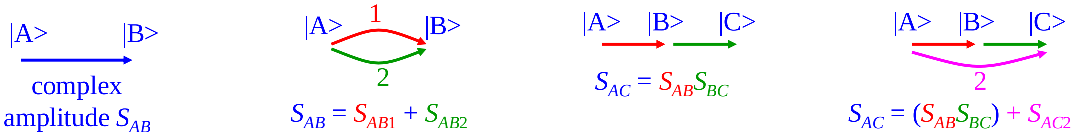


Figure 1.2 The Laws of Quantum Mechanics for combining amplitudes.

These rules, shown diagrammatically in Figure 1.2, are the basis for essentially all of quantum mechanics. [Quantum Field Theory uses them heavily, and Feynman diagrams build on the scheme depicted above.]

1.3.3 The Failure of Quantum Mechanics

The original Quantum Mechanics was *non-relativistic*, and it did not describe the creation and annihilation of particles. For relativistic quantum physics, which requires particle creation and annihilation, one must use Quantum Field Theory. QFT is the most successful physical theory of all time, in that it predicts the magnetic dipole moment (g-factor) of the electron to 13 digits. The reasons for the failures of QM are that it assumes that the number of particles is known and definite, and the hamiltonian is non-relativistic. Whether we use 1-particle wave functions, or multi-particle wave functions, QM assumes a fixed set of particles. For non-relativistic physics, a fixed set of massive particles is a good approximation, because the low energies have negligible probabilities of creating a particle/antiparticle pair, or other massive particles. However, at relativistic energies, the probability of massive particle creation can be significant. Particle creation (including photon creation) adds new “paths” for quantized interactions to take place, thus changing the probabilities of the outcomes. For example, the g-factor of the electron is measurably affected by these phenomena. Quantum Field Theory is the extension of quantum mechanics to relativistic dynamics and particle creation.

1.4 Axioms to Grind: The Foundation of Quantum Mechanics

One can develop most of quantum mechanics from a small set of principles. Well-chosen axioms provide great power to work with quantum mechanics, and learn new systems (unlike mathematical axioms, which are often cryptic and unenlightening). Each of our axioms is directly motivated by experimental results. Different references choose different axioms, and even different numbers of axioms. Their developments are generally not mathematically rigorous, and neither is ours. Instead, our development aims to illuminate the conceptual workings of quantum mechanics. We reiterate that at the quantum level, we cannot talk of what is “really” happening; the best we can hope for is a simple, valid model. Here, we choose a model that is grounded in simple experiments, and therefore intuitive and instructive to the behavior of quantum systems.

This section assumes that you have seen some quantum mechanics and wave-functions, and know the de Broglie relation between momentum and wavelength. It also refers to phasors, and their representation of oscillations, which are described in more detail later. Briefly, a phasor is a complex number that represents the amplitude and phase of a real sinusoid. The axioms below are written for a single-particle continuous system (wave-functions), but can be adapted easily to discrete systems (e.g., angular momentum). We also later extend them to multiple particles.

Quantum mechanics can be formulated from the following observations, taken as axioms:

1. Quantum systems oscillate sinusoidally *in space* with a frequency proportional to momentum: $\mathbf{k} = \mathbf{p}/\hbar$. In other words, quantum systems have a wavelength, given by de Broglie. **k** is the spatial frequency in (say) radians per meter, and points in the direction of momentum. This momentum-wavelength relation was known true for photons from simple interferometry [which means it’s true relativistically, as well, since photons travel at light speed]. This was demonstrated for electrons serendipitously by Davisson and Germer in 1927, after their equipment failed. No one knows *what* exactly is oscillating, as it cannot be directly observed, however its interference patterns can be observed. We model the (unknown) thing that oscillates as the quantum state of the system. Initially, we consider quantum states which are wave-functions. Later, we’ll extend that to include discrete states.

2. Quantum systems oscillate sinusoidally *in time* with a frequency proportional to energy: $\omega = E/\hbar$. This was known true for photons from the photo-electric effect [which again means it's true relativistically]. It is also experimentally confirmed for both electrostatic and gravitational potential energy [Sak p125-9]. Any temporal oscillation has a frequency, phase, and real amplitude. The real amplitude and phase can be described mathematically by a phasor: a complex number that represents a sinusoid. The frequency is not given by the phasor, and must come from other information, in this case, the energy. In quantum mechanics, a phasor may be called a **complex amplitude**, or confusingly, often just "amplitude." Because quantum systems oscillate, and we represent those oscillations with complex-valued phasors, the wave-function is a complex-valued function of space (or a phasor-valued function of space). Since energy can be positive or negative, the frequency of oscillation can be positive or negative. We describe phasors and negative frequency later. [In a relativistic treatment, axioms (1) and (2) combine into a Lorentz covariant form that unifies the time and space pieces.]

3. Systems exist in a *linear* **superposition** of states. In other words, even a single, indivisible system behaves as if it is separated into pieces (components), each in a different state, and each of which behaves like the *whole* system. This is suggested by the observed interference of matter waves: electrons, atoms, molecules, etc. When measuring such a superposition, one will get a measurement consistent with *one* of the components of the system, but no one can predict exactly which component will be seen in the measurement. The "weights" of the pieces, though, determine the probability of measuring the system according to that component: larger weights have larger probabilities of being observed. For example, a spin-1/2 particle could be 1/3 spin up, and 2/3 spin down. Furthermore, because systems oscillate *sinusoidally*, each piece has not only a magnitude, but a phase of its oscillation. This means we can represent each piece by a complex number (phasor), giving its weight and phase of oscillation. Again, the frequency of oscillation is given by the energy of each piece (Axiom 2). The concept of a superposition leads to the need for linear *operators*, which quantum mechanics uses heavily.

4. Quantum mechanics is *consistent*, in the sense that if you measure something twice, you get the same answer (provided you measure quickly, before the system time evolves into a different state). This implies the collapse of the quantum state (aka, loosely, "collapse of the wave-function"). This was observed easily in many experiments, such as Stern-Gerlach filtering.

5. The weight of a given component of a superposition is proportional to the square of the amplitude of its oscillations, much like the intensity of an EM wave is proportional to the square of the E-field (or B-field, or A-field). A "point particle" actually exists in a superposition of locations. Thus each point in space has a phasor (complex number) representing the fractional *density* (per unit volume) of the particle being at that point. This function of space is the particle's **wave-function**. The volume density is therefore given by the square of the wave function, which we normalize such that

$$\rho(\mathbf{r}) = \left|\psi(\mathbf{r})\right|^2 \qquad \Rightarrow \qquad \int_\infty \left|\psi(\mathbf{r})\right|^2 \, d^3\mathbf{r} = 1 .$$

6. Note that the axioms of time and space oscillation, and superposition, imply the Schrödinger equation for classically allowed regions where the particle energy is above the potential energy. We could show this by considering particles of definite momentum, and noting that any particle can be written as a superposition of such momentum states. However, we must *postulate* that the Schrödinger equation is also valid in the classically forbidden regions, where the particle energy is *less than* the potential energy. In these regions, the quantum kinetic energy is negative, and the "momentum" is imaginary. This axiom is consistent with the empirically confirmed phenomenon of tunneling.

These axioms are the foundation of quantum mechanics, and we rely on them implicitly and explicitly throughout our discussion. We also develop many of the implications of these axioms, which yields the incredibly large and diverse science of quantum mechanics.

Note that the quantum formulas for energy use the classical nonrelativistic hamiltonian, e.g. $E = p^2/2m + V(x)$, though x and p become operators. You might be surprised that QM is so weird, yet still uses a classical hamiltonian. However, classical mechanics is the high-energy and/or large-system limit of QM, so perhaps it is to be expected that the classical hamiltonian matches the underlying quantum hamiltonian.

What is a wave function? A wave-function is some unknown "thing" that oscillates. That thing can be thought of as a "quantum field," similar to an electric or magnetic field, but with some important differences. [This is the basis of the more-complete theory of QFT.] Note that an EM field is a classical field on large scales, and also a quantum field on small scales, so the two types of field are clearly related. In QM, though, a whole molecule (or other complex system) can be represented by a single "wave-function." This follows from the linear nature of superpositions. Such a representation is an approximation, but is quite good if the internal details of the system are not important to its observable behavior.

Collapse of the wave-function: This is the most difficult concept of quantum mechanics. It stems from the consistency postulate. There are several models of wave-function collapse, all of which predict the same results; therefore, all of them are valid. In this book, we choose the model we believe is simplest and most consistent over all sizes of systems, spanning a continuum from microscopic to macroscopic. We compare some other "collapse models" later. We introduce the ideas here, with more details to come.

Briefly, a wave-function is said to "collapse" when you make a measurement on it. If the quantum state were a superposition of multiple possibilities, your measurement changes the state to one consistent with your measurement. A more detailed analysis reveals that you can think of "collapse" as happening in two steps: first, the system "decoheres" by being measured (entangled) by *any* macroscopic measuring system (dead or alive). Note that there is no concern for sentience, just "bigness." Once a system has decohered, it follows classical probabilities, and there is no further possibility of quantum interference. In other words, decoherence eliminates any "quantum weirdness" (i.e., nonclassical behavior). However, in our model, time evolution is governed by the Schrödinger equation, and even a decohered system is in a superposition of states. Only when you observe a measurement do the other possibilities disappear (the quantum state collapses). This decoherence model is equivalent to one in which the quantum state *collapses upon decoherence*, but you simply don't know the result until you look at it. However, the decoherence model fully explains partial coherence with no additional assumptions, and is therefore preferable.

Note that since all sentient observers are macroscopic, they necessarily decohere a system before "seeing" the result of a measurement. Also, both models above have a collapse somewhere in them.

Before decoherence was widely recognized, scientists wondered about whether consciousness somehow affected quantum results. This was the puzzle of Schrödinger's cat: is it "sentient enough" to collapse a quantum state? One interpretation was that, regardless of whether a cat can collapse a wave function or not, an outside observer experiences the (idealized) "cat-in-the-box" as being in a superposition of states. This is still true even if the cat is replaced by an idealized physicist (whom we shall call "Alice"). However, realistic macroscopic systems (cats, physicists, or dial gauges) cannot exist in "coherent superpositions" for any noticeable time, because uncontrollable and unrepeatable interactions with the environment randomize the phase relationship between components of the superposition. This fact renders the riddle of Schrodinger's Cat moot: decoherence leads to classical probabilities, and prevents any observable quantum effect. It is impossible to measure whether a cat collapses a wave-function, so the question is scientifically meaningless.

Uncertainty is not an axiom: Note that there is no axiom about uncertainty, or wave-particle duality (complementarity). These concepts follow from the given axioms. The wave nature of the Schrödinger equation, combined with the collapse of the wave-function, allows us to summarize "duality" roughly as "particles propagate like waves, but measure like points." We examine the nature of measurements in some detail later.

Note on superpositions: *Every* quantum state can be written as a superposition in some bases, and also as a (non-superposed) basis state in other bases. Thus, there is nothing fundamental about a state being either a "superposition" or a "basis" state. (We discuss bases in more detail later.) However, some bases are special, with special properties, such as energy, position, and momentum. Basis states in those bases have properties of particular interest, which we will discuss throughout this work.

1.5 Energy and Momentum Are Special

Energy and momentum are very special physical quantities: they are linearly related to a quantum particle's frequency and wave-number, by the constant $\hbar$ ("h-bar"):

$$E = \hbar\omega \qquad \text{and} \qquad p = \hbar k \,.$$

Energy and momentum relate directly to the wave nature of particles. This is not true of other physical quantities such as velocity, acceleration, etc. Therefore, energy and momentum get special treatment; we use them as much as possible, because they fit most easily into the wave equations. For example, we quantify kinetic energy as

$$T = \frac{p^2}{2m} \qquad \text{rather than} \qquad T = \frac{1}{2}mv^2 \,,$$

because p (momentum) is directly related to wave functions, and v (velocity) is not.

All other physical quantities have to be expressed in terms of energy and momentum. For example, to find the velocity of a particle (in the absence of a magnetic field), we find its momentum, which is fundamental, and divide by its mass:

$$v = \frac{p}{m} \qquad \text{(absent magnetic fields)} \,.$$

Not coincidentally, energy and momentum have a special place in relativity, as well: they compose the energy-momentum 4-vector, which is Lorentz invariant. In other words, relativity ties together energy and momentum in a special way, so they must *both* be special.

There is also a complication in that the quantum momentum p is *canonical* momentum, not kinetic momentum. As we will see, in systems with magnetic fields, this distinction becomes important.

1.6 Complex Numbers

1.6.1 Fundamental Properties of Complex Numbers

OK, so we admit complex numbers aren't really all that quirky, but here's a summary of the characteristics of complex numbers that are most relevant to QM.

Why do we use complex numbers? We *could* write everything with only real numbers; after all, everything measurable uses real numbers to measure it. But using only real numbers makes the formulas very complicated.

We use complex numbers because they make mathematical formulas simpler.

Since we spend more time working with formulas than with the actual numbers, it's much more efficient to use complex numbers and simpler formulas, rather than simple (real) numbers and complicated formulas. We start by defining:

$$i \equiv \sqrt{-1}\,, \qquad \text{an imaginary number.}$$

i is a constant, whose square is –1. There is no *real* number with this property, so i is said to be **imaginary**. [Electrical engineers write 'j' instead of 'i', because 'i' stands for electric current.] From the definition:

$$i^2 = (-i)^2 = -1 \qquad \text{and} \qquad -i^2 = 1 \,.$$

Since all the rules of arithmetic, algebra, and calculus apply to imaginary numbers, you'll commonly see these types of identities:

$$i = \frac{-1}{i}, \qquad -i\hbar = \frac{\hbar}{i}, \qquad \text{etc.}$$

i can be multiplied by any real number, say *y*, to give *iy*. Any such real multiple of *i* is also imaginary, and its square is a negative number: $-y^2$. The sum of a real number and an imaginary number is a **complex number**: $x + iy$. The label 'z' is often used for complex variables:

$$z = x + iy .$$

The graphical interpretation of complex numbers is critical. Recall that we can plot any real number as a point on a 1-dimensional number line. Similarly, we can plot any complex number as a point, or a vector, on a 2-dimensional plane (the **complex plane**); see Figure 1.3.

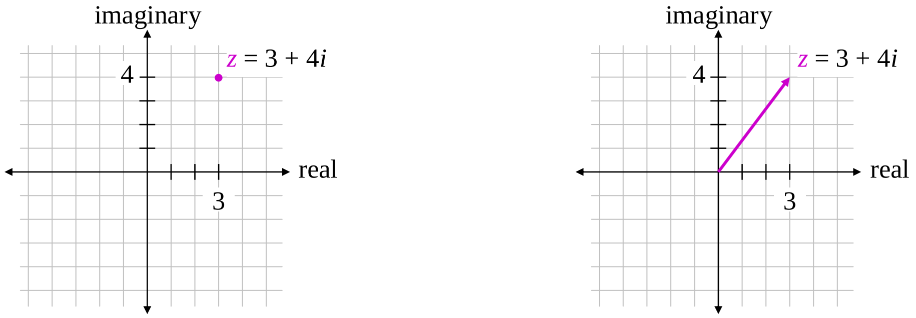


Figure 1.3 Rectangular picture of *z* in the complex plane. (Left) As a point. (Right) As a vector.

The $z = x + iy$ form is called the **rectangular form** of a complex number. Adding (or subtracting) complex numbers is straightforward; the rectangular *components* add (or subtract), because complex numbers are vectors:

$$(a+ib)+(c+id)=(a+c)+i(b+d) \qquad (1+2i)+(3+4i)=4+6i .$$

Besides rectangular form, complex numbers can also be expressed in **polar form**. Polar form is intimately tied to Euler's identity (pronounced "oilers"), which relates an angle to real and imaginary components (Figure 1.4, left).

$$e^{i\theta}=\cos\theta+i\sin\theta, \qquad \theta \text{ in radians} .$$

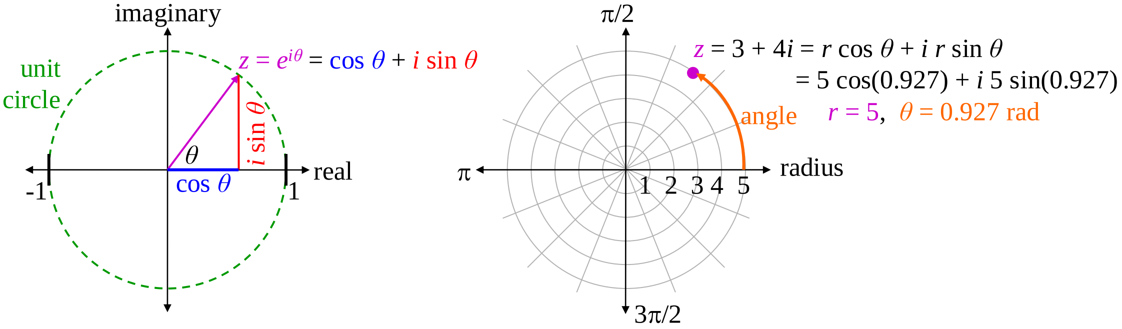


Figure 1.4 (Left) Euler's identity in the complex plane. (Right) Polar picture of *z* in the complex plane.

[Aside: Euler's identity can be derived from the Maclaurin expansion $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$, or from solving the 2nd order differential equation, $\frac{d^2}{dx^2} f(x) = -f(x)$.]

Multiplying $e^{i\theta}$ by a magnitude, *r*, allows us to write any complex number in **polar form**: $z = (r, \theta)$, by choosing *r* and θ appropriately (Figure 1.4, right). Comparing the two graphs in Figure 1.4, we see that we can relate the polar and rectangular forms by a mathematical expression:

$$z = (r,\theta) = re^{i\theta} = r(\cos\theta\ +\ i\sin\theta) = r\cos\theta\ +\ ir\sin\theta = x + iy\ .$$

$$x = r\cos\theta, \qquad y = r\sin\theta, \qquad \text{and} \qquad r = \sqrt{x^2+y^2}, \qquad \theta = \tan_{4q}^{-1}(x,y),$$

where $\tan_{4q}^{-1}(x, y)$ is a full, 4-quadrant arctangent of (y/x). E.g., the usual 2-quadrant arctangent gives:

$$\tan^{-1}\left(\frac{-1}{-\sqrt{3}}\right) = \tan^{-1}\left(\frac{1}{\sqrt{3}}\right) = \frac{\pi}{6}, \qquad \text{but} \qquad \tan_{4q}^{-1}\left(-1,-\sqrt{3}\right) = \frac{-5\pi}{6}\ .$$

Rectangular and polar forms are two different, but equivalent,
ways of expressing complex numbers.

Note that θ may be positive or negative. Polar form is not unique, because:

$$e^{i\theta} = e^{i(\theta+2n\pi)}, \qquad n \text{ any integer}\ .$$

The angle of purely real numbers is either 0 or π, and the angle of purely imaginary numbers is $\pm\,\pi/2$.

All the rules of algebra and calculus work with complex numbers: commutativity, associativity, distributivity, exponentials, derivatives, integrals:

$$e^z e^w = e^{z+w}, \qquad \frac{d}{dz}e^z = e^z, \qquad \int e^z\ dz = e^z + C\ , \text{ etc.}$$

Therefore, multiplying (or dividing) is easier in polar form, using Euler's identity (Figure 1.5):

$$\left(re^{i\theta}\right)\left(se^{i\phi}\right) = rse^{i(\theta+\phi)}, \qquad \text{and} \qquad \frac{re^{i\theta}}{se^{i\varphi}} = \frac{r}{s}e^{i(\theta-\varphi)}\ .$$

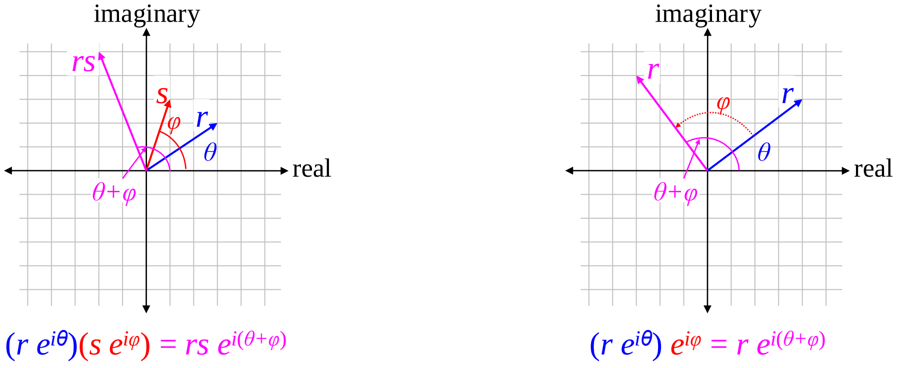


Figure 1.5 (Left) Picture of complex multiply. (Right) Special case: multiplication by $e^{i\varphi}$

When multiplying (or dividing) complex numbers, the radii multiply (or divide), and the angles add (or subtract). This fact suggests the concept of a "magnitude" for complex numbers. We define the **magnitude** (aka absolute value) of a complex number as its radius in polar form:

$$|z| \equiv r = \sqrt{x^2+y^2} \qquad \text{(magnitude is a real number } \ge 0)\ .$$

Thus when multiplying (or dividing) complex numbers, the magnitudes (radii) multiply (or divide):

$$|z_1 z_2| = |z_1|\ |z_2| \qquad \text{and} \qquad \left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|}\ .$$

Note that when z happens to be real, $|z|$ reduces to the familiar real-valued magnitude (aka absolute value).

An important special case: When we multiply z by some unit-magnitude number ($e^{i\varphi}$), the magnitude of z doesn't change, but z gets rotated in the complex plane by the angle φ (Figure 1.5, right). This is important for phasors, described shortly.

Using the polar form, we can easily raise a complex number to any real power:

$$z^a = \left(re^{i\theta}\right)^a = r^a e^{ia\theta} \qquad \text{magnitude is raised to power } a\text{; angle is multiplied by } a.$$

Adding complex numbers follows the "head to tail" rule for adding vectors. The two examples in Figure 1.6 illustrate two different results possible from adding complex numbers of the same magnitude, but different angles:

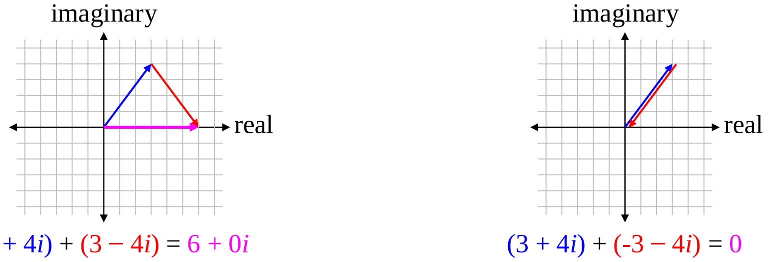


Figure 1.6 A variety of results of adding complex numbers of equal magnitude.

A crucial fact is:

When *adding* or *subtracting* complex numbers, there is no simple relationship between the magnitudes of the addends and the magnitude of the sum.

Their angles, or "phases," are important to the result. This is the essence of interference.

You can also define logarithms and exponents of complex numbers, but we don't need them at this level.

The angle of a complex number may be written as arg(z), and is sometimes called its **phase** (or its **argument**). As shown above, when multiplying (or dividing) complex numbers, the angles add (or subtract):

$$\arg(z_1 z_2) = \arg(z_1) + \arg(z_2), \qquad \text{and} \qquad \arg\left(\frac{z_1}{z_2}\right) = \arg(z_1) - \arg(z_2)\,.$$

The **complex conjugate** of a number z is written z^*, and is defined as follows:

If $z = a + ib, \quad z^* \equiv a - ib,$ and in polar form,

if $z = re^{i\theta}$, then $z^* = re^{-i\theta}$.

Note that the conjugate of a *real* number is itself: $x^* = x$.

An important identity is

$$|z|^2 = z^* z \qquad \text{because } |z|^2 = r^2 = x^2 + y^2 = (x - iy)(x + iy) = z^* z,$$

$$\text{or in polar form: } |z|^2 = r^2 = \left(re^{i\theta}\right)\left(re^{-i\theta}\right) = z^* z.$$

Any complex function can be expressed by its real and imaginary parts, e.g. let $\psi(x)$ be a complex valued function of the *real* value, x:

$$\psi(x) = a(x) + i\,b(x)\,.$$

Linearity then provides identities for derivatives and integrals, which are also complex valued functions:

$$\frac{d}{dx}\psi(x)=\frac{d}{dx}a(x)+i\frac{d}{dx}b(x), \qquad \int \psi(x)\,dx=\int a(x)\,dx+i\int b(x)\,dx\,.$$

The conjugate of a complex function $[\psi(x)]^*$ is written as $\psi^*(x)$. Important conjugation identities (derived simply from the rules above):

$$(z+w)^*=z^*+w^* \qquad (zw)^*=z^*w^*$$

$$\frac{\partial}{\partial x}\psi^*(x)=\left[\frac{\partial}{\partial x}\psi(x)\right]^* \qquad \int \psi^*(x)\,dx=\left[\int \psi(x)\,dx\right]^*.$$

Let $v(z)$ be a complex valued function of the complex value, z. In general, $v(z^*) \neq v^*(z)$. A simple counter example is $v(z) = z + i$; then $v^*(z) = z^* - i \neq v(z^*)$. However, there are some special cases relating functions and conjugates. Most importantly:

For y real: $\left(e^{iy}\right)^*=e^{-iy}$.

For $z=x+iy$: $e^{(z^*)}=\left(e^{z}\right)^*$

$$\text{because } e^{(z^*)}=e^{(x-iy)}=e^{x}\left(e^{+iy}\right)^*=\left(e^{x}e^{+iy}\right)^*=\left(e^{x+iy}\right)^*=\left(e^{z}\right)^*.$$

The real part of a complex number, z, is written as Re(z). The imaginary part of 'z' is written Im(z). Note that Im(z) is a *real* number: it is the coefficient of i in the complex number z. Thus,

$$z = \mathrm{Re}(z) + i\,\mathrm{Im}(z)\,.$$

Derivatives and integrals of functions of *complex* values [say, $w(z)$] are uncommon in elementary QM, because dz has both a real and imaginary differential component. Though contour integrals of complex differentials do appear in more advanced studies, we won't need them here.

1.6.2 Phasors, Rotation in the Complex Plane, and Negative Frequency

Phasors are critical to nearly all aspects of oscillation in engineering and physics. The simple picture of a phasor starts with a clockwise rotating stick, and its projection (shadow) onto the horizontal axis (Figure 1.7). The stick rotates at a constant angular velocity, ω. From the definition of cosine, the shadow traces out a sinusoid, of frequency ω, with amplitude equal to the length of the stick. We can add a phase to the sinusoid by having the stick start at some angle, θ, when $t = 0$:

$$a(t)=(length)\cos(\omega t-\theta)\,.$$

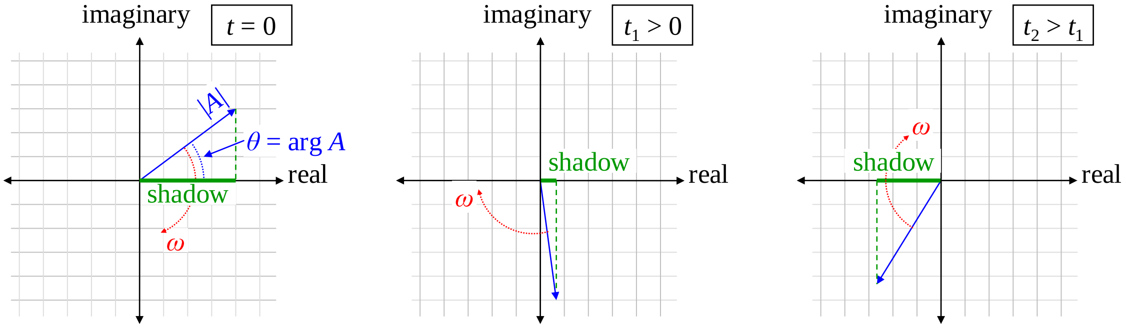


Figure 1.7 For physicists, the stick usually rotates clockwise (as shown), per $e^{-i\omega t}$. Note that $\theta > 0$ in this example.

Rotation in the complex plane: We've seen how multiplication by a unit-magnitude complex number simply rotates another complex number in the complex plane. Now imagine the unit-magnitude angle is not fixed, but changes linearly with time, i.e. rotate not by a fixed angle, but by an increasing angle

ωt. 'ω' is the angular frequency, in rad/s. In physics, time-dependence usually goes as $e^{-i\omega t}$. If we multiply some complex number $r\,e^{i\theta}$ by $e^{-i\omega t}$, we get

$$\left(r\,e^{i\theta}\right)e^{-i\omega t}=r\,e^{i(\theta-\omega t)},$$

a complex function of time that rotates continuously in the complex plane. The magnitude of the result is fixed at $|r|$, because $|e^{-i\omega t}| = 1$ at all times. But the angle of the result decreases with time, at the rate ω (Fig Figure 1.8, left). (NB: Engineers, and some physics applications, use a different time dependence, $e^{+i\omega t}$, in which the complex vector rotates counter-clockwise for $\omega > 0$.)

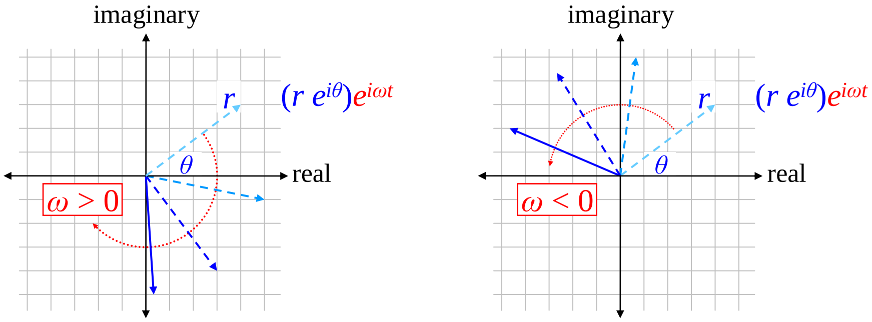


Figure 1.8 (Left) Clockwise rotation in time at positive frequency $\omega > 0$.
(Right) Counter-clockwise rotation in time at negative frequency $\omega < 0$.

We are now prepared to define a phasor: a **phasor** is a complex number which represents the amplitude and phase of a real-valued sinusoid. The phasor does not define the frequency of the sinusoid, which must come from other information. Specifically, the real-valued sinusoid $a(t)$ represented by a phasor A is:

$$a(t)=\operatorname{Re}\left\{Ae^{-i\omega t}\right\}=\cos\left(\omega t-\arg(A)\right)$$

The Re{} operator returns the "shadow" of the complex number on the real axis. Thus graphically, as before, $a(t)$ is the projection of a rotating vector onto the real axis; the vector's magnitude is $|A|$, and its starting angle ($t = 0$) is $\arg(A)$. A phasor is also called a "complex amplitude."

A key fact, proven in *Quirky Electromagnetic Concepts*, is that the sum of any two sinusoids of frequency ω is another sinusoid of frequency ω. Furthermore:

The phasor for the sum of two sinusoids is the sum of the phasors of the two sinusoids.

Mathematically, for a given ω:

$$\text{If}\qquad c(t)=a(t)+b(t),\qquad \text{then}\qquad C=A+B$$

$$\textit{where}\qquad A,B,C\equiv\text{phasors for } a(t),b(t),c(t).$$

Negative frequency: In QM, system frequency is proportional to energy, and energy can be negative. Therefore, instead of just oscillating, quantum systems can be thought of as *rotating* (in some unknown space). This requires both positive and negative frequencies (rotations in opposite directions). Again, no one knows *what* is rotating, but the interference consequences can be observed. Therefore, angular frequency is not constrained to be positive; it can just as well be negative. Rotation by a negative frequency rotates in the counter-clockwise direction, rather than clockwise (Figure 1.8, right). Hence, both positive and negative frequencies occur in complex rotations, and in Quantum Mechanics.

In QM, only frequency (and energy) *differences* are significant. Therefore we could, in principle, eliminate negative frequencies in all our calculations by adding a large, positive constant frequency to everything. That would be tedious, and problematic in future studies of QFT and anti-particles. Therefore, negative frequency gives us a convenient way to accommodate arbitrary frequency differences, while still allowing an arbitrary zero-point for frequencies. Since frequency is proportional to the system's energy (E

= $\hbar\omega$), the zero point of frequency corresponds to the zero point of energy, and negative energies have negative frequencies. In other words, negative frequencies preserve our freedom to choose the zero-point of energy.

1.7 Probability, Density, and Amplitude

Quantum Mechanics relies heavily on probability, probability density, and probability amplitude. Probability and probability density (henceforth often shortened to "density") are concepts common to many fields; "probability amplitude" is unique to QM. We briefly review here the concepts of probability, and precisely define our terms.

1.7.1 Probability of Discrete Events

Probability is a quantification of how likely a **random** event is to occur. A random event is one whose occurrence cannot be predicted exactly. The probability of an event is the fraction of the time that the event will occur in a large number of repetitions (**trials**) of an experiment. Each trial, of course, is set up exactly like all the others. It is in the nature of "random" events that even though all trials are set up identically, different trials may produce different events. So it is with QM measurements: most measurements of observables are random events. For experiments with **discrete** outcomes, such as rolling a die or measuring the energy of a particle in a superposition of discrete energy states, each possible outcome has a finite probability. For experiments with continuous outcomes, such as the position or momentum of a particle, each outcome has a finite probability density (more on this below).

Discrete example: measuring the spin of a spin-½ particle can produce only two possible values: parallel to the measuring equipment (call it "up"), or anti-parallel to it ("down"). The two values, "up" and "down," are mutually exclusive. If in a given experiment the probability of "up" is 0.7, then the probability of "down" must be 0.3, because the sum of all possibilities must be 1 (i.e., the probability of measuring up *or* down is 1).

Note that an event is random if its outcome cannot be exactly predicted. All the possibilities need not be equally likely. Some references use the term "completely random," to mean all possibilities are equally likely. We find that term ambiguous. "Random" events are not exactly predictable, and the possible outcomes may or may not be equally likely.

A more complicated discrete example: the energy of a particle in a bound state has discrete values. However, though discrete, there may be an infinite set of possible values. For example, an electron bound to a proton (hydrogen) has an infinite set of possible energies, characterized by an integer, *n*, such that:

$$E_n = \mathrm{Ry}/n^2 \qquad \textit{where} \qquad \mathrm{Ry} \equiv \text{Rydberg constant} .$$

Clearly, not all of them can be equally likely, since the infinite sum of their probabilities is 1.

The probability of an event, *e*, may be written as Pr(*e*), e.g., the probability of measuring E_n may be written "Pr(E_n)" or "Pr($E = E_n$)".

1.7.2 Probability Density of Continuous Values

We've seen that probability describes discrete random variables, such as rolling a die, or measuring (discrete) bound-state energy. But what about continuous random variables? Suppose I spin the pointer on a board game (Figure 1.9, left), and consider the angle at which the pointer stops. That angle could be any of an infinite set of values, but a *continuous* infinite set, not a *discrete* infinite set. Angle, in this case, is a **continuous** random variable: between any two values are an infinite number of other possible values. There are uncountably many angle values that could occur. Probability density (aka "density") addresses such continuous random variables.

In the case of a simple pointer that rotates in a plane, all angle values are equally likely; the pointer has no preference for one position over any other. The range of possible measurements is the half-open interval $[0, 2\pi)$, i.e., $0 \le \theta < 2\pi$. The probability of measuring an angle < 0 or $\ge 2\pi$ is zero:

$$\Pr(\theta < 0) = 0, \qquad \text{and} \qquad \Pr(\theta \ge 2\pi) = 0.$$

However, the probability of measuring the pointer between a range of allowed angles is finite. Suppose we divide the angles into two halves: upper $[0, \pi)$, and lower $[\pi, 2\pi)$. The probability of measuring in the upper half is 0.5, and the probability of measuring in the lower half is also 0.5. We could divide the circle into quadrants, and the probability of measuring in any quadrant is 0.25. In fact, for any finite angular range, there is a finite probability of measuring an angle in that range. In the simple pointer example, the probability of measuring in a range is proportional to the size of the range, i.e. all angles are equally likely. Noting that $\Pr(0 \le \theta < 2\pi) = 1$, and that all angles are equally likely, it must be that (given $0 \le \theta_1 \le \theta_2 < 2\pi$):

$$\Pr(\theta_1 < \theta < \theta_2) = \frac{\theta_2 - \theta_1}{2\pi} .$$

We can find and graph a function of angle, pdf(θ), such that (Figure 1.9, right):

$$\Pr(\theta_1 < \theta < \theta_2) = \int_{\theta_1}^{\theta_2} \text{pdf}(\theta)\, d\theta \qquad \text{for all } \theta, -\infty < \theta < \infty .$$

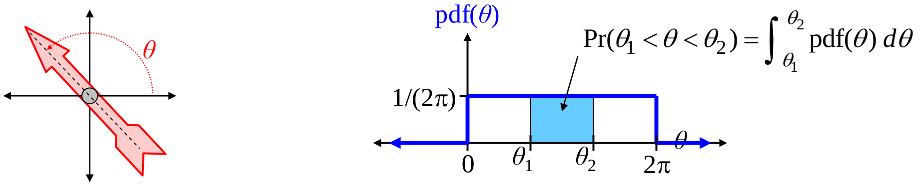


Figure 1.9 (Left) A simple spinner. (Right) Example of pdf(θ), and $\Pr(\theta_1 < \theta < \theta_2)$.

Since the angle must be somewhere in $[0, 2\pi)$,

$$\int_0^{2\pi} \text{pdf}(\theta)\, d\theta = 1 \qquad \Rightarrow \qquad \text{pdf}(\theta) = \frac{1}{2\pi} .$$

Note that the probability of measuring any given, specific value (with no interval around it) is vanishingly small, i.e. zero. Therefore, $\Pr(\theta \text{ in } [a, b]) = \Pr(\theta \text{ in } (a, b))$.

As a slightly more complicated example, let's measure the 1-dimensional position of a hypothetical particle. Position is a continuous random variable, and with many real-world measurements, the probability of finding the particle is concentrated near a point, though finite probabilities extend to infinity in both directions.

Within a differentially small region *dx*, around a value *x*,
the probability of finding the particle is proportional to *dx*.

This is the relationship between probability and probability density:

Pr(random value being in the region $[x, x+dx]$) = pdf(x) dx .

pdf(x) varies for different values of x, and is called the **probability distribution function**, or **PDF**:

$$\text{pdf}(x) \equiv \lim_{dx \to 0} \frac{\Pr\left(x < X < x + dx\right)}{dx}, \qquad \textit{where} \quad X \text{ is a random variable described by pdf}(x) .$$

Thus we can graph the PDF (proportionality factor) as a function of x (Figure 1.10).

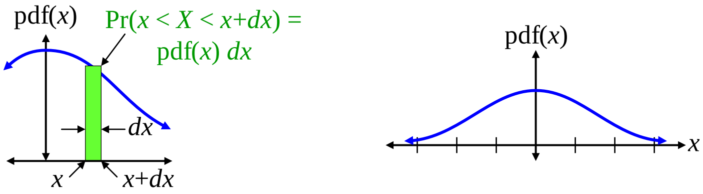


Figure 1.10 Typical probability distribution functions (PDFs).

The PDF is like a smoothed-out histogram of samples of the random variable. If you made billions of trials, and a fine-grained histogram of the results, you'd get (essentially) the PDF.

What about the probability of the random value being in a finite interval [a, b]? From:

$$\Pr(\text{random value being in the region } [x, x+dx]\ ) = \text{pdf}(x)\ dx\ ,$$

it follows that:

$$\Pr\left(\text{random value being in the region } [a,b]\ \right) = \int_a^b \text{pdf}(x)\ dx\ .$$

Furthermore, since the particle must be somewhere in ($-\infty$, $+\infty$),

$$\int_{-\infty}^{\infty} \text{pdf}(x)\ dx = 1\ .$$

When there is more than one pdf under consideration, we use a subscript to distinguish them. E.g., the PDF for a random variable X might be given as $\text{pdf}_X(x)$.

[Sometimes, one encounters a function whose value is *proportional* to the probability density, i.e., the function is some constant times the PDF. In that case, $\int_{-\infty}^{\infty} f(x)\ dx \ne 1$. Such a function is called an **unnormalized** PDF.] In this book, we always use normalized PDFs (and wave-functions).

1.7.3 Complex (or "Probability") Amplitude

In Quantum Mechanics, probabilities are given by a wave-function, such as $\psi(x)$. $\psi(x)$ is a complex-valued function of the real position x: for every real number x, $\psi(x)$ gives a complex number. $\psi(x)$ is related to pdf(x), i.e., the probability density function for measuring the particle at the point x:

$|\psi(x)|^2 = \text{pdf}(x)$, and recall $|\psi(x)|^2 = \psi^*(x)\ \psi(x)$.

Because PDFs must integrate to 1, we say $\psi(x)$ is "normalized" if:

$$\int_{-\infty}^{\infty} dx\ \psi^*(x)\ \psi(x) = 1.$$

$\psi(x)$ is the complex-valued **probability amplitude** [Bay p8t], or just **amplitude**, at each point, x.

Why do we need a complex valued function to define a simple real-valued PDF? We need it because of the way probability amplitudes combine when a wave-function is a superposition of two or more possible states.

The two states of a superposition combine as if their amplitudes were oscillating sinusoids at every point, with both an amplitude and a phase. We can think of $\psi(x)$ as a phasor-valued function of space. This means the aggregate probability is *not* the sum of the (real) probabilities of the component states. Instead, the aggregate (complex) probability amplitude is the sum of the (complex) amplitudes of the component states. An example illustrates this:

Suppose there exists a state, ψ, which at some point a is: $\psi(a) = (1 + i)$. Then $|\psi(a)|^2 = 2 = \text{pdf}(a)$. Suppose another state exists, φ, which at the point a is the same as ψ: $\varphi(a) = (1 + i)$. Then $|\varphi(a)|^2 = 2 = \text{pdf}(a)$. Now suppose a particle is in an equally weighted superposition of states ψ and φ (call this new state χ). What is the pdf of $\chi(a)$? Let's start by noting that, if $\psi(x)$ is orthogonal to $\varphi(x)$:

$$\chi(x) = \frac{\psi(x) + \varphi(x)}{\sqrt{2}} \qquad \text{(we insert the } \sqrt{2} \text{ to keep } \chi \text{ normalized)}\ .$$

Then $\chi(a) = \dfrac{(1+i)+(1+i)}{\sqrt{2}} = \sqrt{2} + \sqrt{2}i$, and $\text{pdf}(a) = |\chi(a)|^2 = 4$, which is twice the PDF of either ψ or φ. This is because ψ and φ reinforce each other at that point..

Now let's suppose that ψ and φ are different at the point a. We'll leave $\psi(a) = (1 + i)$, but let $\varphi(a) = (1 - i)$. Then $|\varphi(a)|^2 = 2 = \text{pdf}(a)$, just as before. But what happens now if a particle is in an equally weighted superposition of states ψ and φ? Then, the state χ is given by (again assuming orthogonal ψ and φ):

$$\chi(a) = \frac{\psi(a) + \varphi(a)}{\sqrt{2}} = \frac{(1+i)+(1-i)}{\sqrt{2}} = \sqrt{2}$$

Then $\text{pdf}(a) = |\chi(a)|^2 = 2$. In this case, even though the magnitudes of both ψ and φ are the same as before, the magnitude of their sum is *different* than before. This is because they only partially reinforce each other; the *phase* of the complex values of ψ and φ are different, so their magnitudes don't simply add.

Finally, suppose $\varphi(a) = (-1 - i)$. Then $|\varphi(a)|^2 = 2 = \text{pdf}(a)$, as before, but $\chi(a) = 0$ and $\text{pdf}(a) = 0$. In this case, the magnitudes of ψ and φ are still the same as before, but they are of opposite phase, so they cancel completely in an equally weighted superposition of states.

It is this adding of complex valued probability amplitudes
that accounts for most quantum weirdness.

In other words, QM is weird because of interfering sinusoids.

Note that this kind of interference (that of adding complex amplitudes) is identical to phasor computations in fluids, electromagnetics, or any other kind of wave interference. In other words:

The algebra of adding sinusoids of a fixed frequency, but arbitrary amplitude and phase,
is the same as the algebra of complex numbers.

1.7.4 Averages vs. Expectations

It can be shown that the average value of a random variable (averaged over many trials) is:

$$\langle x \rangle = \int_{-\infty}^{\infty} dx\, \text{pdf}(x)\, x, \qquad \text{or} \qquad \langle x \rangle = \sum_{i=1}^{N} \Pr(x_i)\, x_i \, .$$

It is clear that the average value of many trials may be an impossible value for any single trial; e.g., the average value of the roll of a die is 3.5, but no single role can produce that value. Often, the term "expectation value" or "expected value" is used to mean "average value." This can be confusing, because the "expected value" of a die roll is 3.5, yet you would *never* expect the value of a die roll to be 3.5. Since the term "average" is clear and precise, we don't use the term "expectation value."

1.8 The Origin of Schrödinger's Equation

The Schrödinger equation is the equation of motion (EOM) for a wave-function, as well as for other quantum systems, including spin and multi-particle systems. Like all EOMs, it predicts the future state from an initial state. The Schrödinger Equation is completely deterministic; it is the measurement of quantum systems that is probabilistic. Where does Schrödinger's equation come from? Many books note that it cannot be *derived* from anything; it is a *discovery* of Quantum Mechanics. Nonetheless, it must be *motivated* by something; Erwin Schrödinger didn't just pull it out of thin air.

1.8.1 The Wave Equation

Wave Mechanics began when experimentalists noticed that particles had wave properties; most notably, they exhibited wave-like diffraction and interference. (Davisson and Germer discovered electron waves by accident after their equipment malfunctioned.) Moreover, the wavelength of particles has the same relationship to their momentum as photons, namely:

$$p = \frac{h}{\lambda} = \hbar k \qquad \textit{where} \quad k \equiv \text{spatial frequency (aka wave-number), in rad/m}. \tag{1.1}$$

Since one wavelength, λ, is 2π radians, $k = 2\pi/\lambda$.

Since photons are traveling waves, and particles had similar wave-like properties, physicists wondered if particles could be represented by similar traveling waves. A real-valued wave, traveling in the *x*-direction, can be represented by:

$$a(t,x) = B\cos(kx - \omega t) \qquad \textit{where} \qquad \omega \equiv \text{frequency in time, measured in rad/s}$$

$$B \equiv \text{real amplitude of the wave}.$$

Thus at a fixed point in space, the amplitude varies sinusoidally over time; at a fixed point in time, the amplitude varies sinusoidally over space. Viewed over space and time, the wave travels smoothly in the positive *x* direction (Figure 1.11).

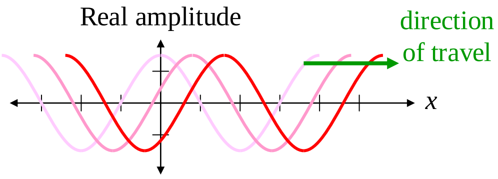


Figure 1.11 A wave traveling smoothly in the positive *x* direction.

Particles were found experimentally to interfere like sinusoids. The mathematics of interfering sinusoids is, in fact, the mathematics of complex arithmetic.

(See Phasors section earlier). This means that particles' waves are well-described by complex-valued traveling waves. Complex-valued wave-functions add as complex numbers and then square to produce probability densities. The complex wave equation then has both real and imaginary parts as traveling waves:

$$\psi(t,x) = Ae^{i(kx-\omega t)} \qquad \textit{where} \qquad A \equiv \text{complex amplitude of the wave}.$$

In other words, *A* is a phasor. The instantaneous value of a wave at a given point in space at a given point in time is sometimes called the **instantaneous amplitude**. Thus, $\psi(t, x)$ is the instantaneous (complex) amplitude of the wave-function at the point (t, x).

1.8.2 Energy, the Hamiltonian, and the Wave Equation

Further work found that for a fixed-energy quantum particle or system, the phase of its representing sinusoid shifted in time at a constant rate: $\phi = \omega t$. This is equivalent to the complex wave function rotating in the complex plane over time as $e^{-i\omega t}$ (i.e., the complex value, at each point in space, had its phase decrease continuously with time). Here again, the analogy to photons was useful. The frequency of rotation (in the complex plane) of the complex value of a wave-function is proportional to its energy, just like a photon:

$$E = \hbar\omega \qquad \textit{where} \qquad E \text{ is the } \textit{total} \text{ energy of the particle (or photon)}. \tag{1.2}$$

Total energy is kinetic energy plus potential energy: $E = T + V$. For a particle, kinetic energy is $T = p^2/2m$. The potential energy of a particle is a given function of its position in space; call it $V(x)$. (For simplicity, we've assumed *V* is independent of time, but all of this works even if $V = V(t, x)$). Therefore:

$$E = T + V = \frac{p^2}{2m} + V(x) \qquad \text{(just as in classical mechanics).}$$

We will see that Schrödinger's equation is just this equation, written as quantum operators acting on the wave-function:

$$\hat{E} = \hat{T} + \hat{V} \qquad \rightarrow \qquad \underbrace{i\hbar\frac{\partial}{\partial t}}_{\hat{E}}\psi = \left[\underbrace{-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}}_{\hat{T}} + \underbrace{V(x)}_{\hat{V}}\right]\psi \qquad \text{(Schrödinger's equation).}$$

If a particle of fixed energy is represented by a traveling wave-function, $\psi(t, x) = A\, e^{i(kx - \omega t)}$, then its derivative with respect to time is:

$$\frac{\partial}{\partial t}\psi(t,x) = -i\omega A e^{i(kx-\omega t)} = -i\omega\psi(t,x) = -i\left(\frac{E}{\hbar}\right)\psi(t,x) ,$$

and rearranging:

$$i\hbar\frac{\partial}{\partial t}\psi(t,x) = E\,\psi(t,x) . \qquad (1.3)$$

The derivative with respect to time multiplies the wave function by a value proportional to the particle's total energy.

The derivative with respect to x of a simple traveling wave function (which has definite momentum, eq. (1.1)) is:

$$\frac{\partial}{\partial x}\psi(t,x) = ikAe^{i(kx-\omega t)} = \left(\frac{ip}{\hbar}\right)\psi(t,x) \qquad \text{and} \qquad \frac{\partial^2}{\partial x^2}\psi(t,x) = -\left(\frac{p}{\hbar}\right)^2\psi(t,x) .$$

Rearranging the latter equation, to get kinetic energy:

$$-\hbar^2\frac{\partial^2}{\partial x^2}\psi(t,x) = p^2\psi(t,x), \qquad \text{and dividing by } 2m \ (m \equiv \text{particle mass})$$

$$\frac{-\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(t,x) = \frac{p^2}{2m}\psi(t,x) = T\,\psi(t,x) .$$

The derivative with respect to x multiplies the wave function ψ by a value proportional to the particle's momentum, and the second derivative multiplies ψ by a value proportional to p^2, and thus proportional to the particle's kinetic energy. Now simply plug $E = T + V(x)$ into the time derivative above, eq. (1.3):

$$i\hbar\frac{\partial}{\partial t}\psi(t,x) = \left[T + V(x)\right]\psi(t,x) = \frac{-\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(t,x) + V(x)\psi(t,x) .$$

Voila! Schrödinger's equation! It's nothing more than a Hamiltonian operator ($T + V$) applied to the wave function, using the experimental relationship between total energy and temporal frequency (ω), and momentum and spatial frequency (k). The Hamiltonian produces a particle's total energy, kinetic plus potential (even if $V(t, x)$ varies in time). On the left side of Schrödinger's equation is the time derivative, producing total energy times ψ: $E\psi$. On the right side is the spatial second derivative, producing $T\psi$, plus the potential energy, $V(x)\psi$. (Magnetic fields complicate this simple picture somewhat.)

The final piece of the puzzle is that a particle usually can not be in a state of a single traveling sinusoidal wave function (because then it would be equally likely to be found anywhere in infinite space). It must be in a superposition of many (even infinitely many) sinusoidal wave functions. But each component traveling wave function individually satisfies Schrödinger's equation. And Schrödinger's equation is a *linear* differential equation. So given a set of functions which satisfy the equation, any superposition (linear combination) of those functions also satisfies Schrödinger's equation. Thus:

In nonrelativistic QM, Schrödinger's equation applies
to all massive particles and all wave-functions.

1.9 The Meaning of the Wave-Function

1.9.1 Where Are We, and Where Do We Go From Here?

The most evident feature of the wave function is that it tells the probability density of finding the particle at any point in space; in 1D the unit of probability density is m^{-1}; in 2D the unit of probability density is m^{-2}; in 3D the unit of probability density is m^{-3}. But the wave function is much more [P&C p59m]:

> A particle's wave-function defines *all* the spatial properties of the particle, not just its position.
> It defines it momentum, orbital angular momentum, kinetic energy, etc.

To compute things involving external fields, such as potential energy, we must know (in addition to the wave-function) the potential field $V(\mathbf{r})$. Note that even the probability density of a particle is more than just the probability of measuring it to be somewhere: for many quantum purposes, a single particle behaves as if it were *actually distributed* throughout space, as defined by its wave function. [In fact, in Quantum Field Theory, they often call it "particle density" or "number density" instead of "probability density." Particle density for anti-particles can be taken as negative.]

For example, if the particle has charge q, we compute the potential produced by it from the charge distribution defined by $\rho_{\text{charge}}(x) = q\left|\psi(x)\right|^2$ [Blo p434b]. (However, the distribution of charge from a *single* particle does not push on itself.) If the particle has mass m, we distribute its mass according to $\rho_{\text{mass}}(x) = m\left|\psi(x)\right|^2$, as well. If a particle has spin, it's spin angular momentum, and therefore also magnetic dipole moment, is actually a spin (and dipole moment) *density*, spread out according to $\rho_{\text{spin}}(x) = \dfrac{\hbar}{2}\left|\psi(x)\right|^2$ [Bay 14-92 p326m].

However, there are limits to the model of a physically distributed particle. When things like mass appear in a denominator, as with kinetic energy = $p^2/2m$, we cannot consider an infinitesimal point as having an infinitesimal fraction of the mass. If it did, and the momentum were macroscopic, the KE would blow up everywhere.

You can localize a particle to a very small region of space with a good measurement, but until you do so, it may be spread out. Also, immediately after you measure it, it will start spreading out again, and become distributed over time. (The concept of actually *being* distributed in space gets trickier with entangled multi-particle states, which we address later.) [In QM, a particle can be localized to an arbitrarily small region, but in reality, as described by QFT, even a so-called "point particle" has a limit to how tightly you can localize it, but that is beyond our scope.]

The wave function does not define the particle's spin (if any), nor any angular momentum or energy associated with the spin (such as magnetic-moment/B-field interactions). Spin-1/2 and related values are defined by "spinors," described later.

Because the wave-function defines a particle's momentum, it tells us not only where it is, but where it's going. The question of where it's going is tricky in QM, though, because any real particle is in a superposition of momentum states, i.e. it's moving in multiple ways. For example, consider bound states: a particle is tied to some region by an inescapable potential. For a **stationary** state (where the particle properties don't change with time), the average momentum must be zero, otherwise the particle would be moving, and escape the potential.

Non-stationary states and unbound states can be subtle; we discuss these in the text as the need arises.

1.9.2 Gross Anatomy of a Wave Function

We now describe some important qualitative properties of wave-functions, whose understanding helps make quantum mechanics more sensible and intuitive. We first introduce a particle at a high energy above the potential, which behaves fairly classically, then discuss a more "quantum-like" energy, to reveal some of the non-classical consequences of QM.

High energy wave-function: Consider a particle in a box, with a potential step at the bottom (Figure 1.12). What does the wave function look like?

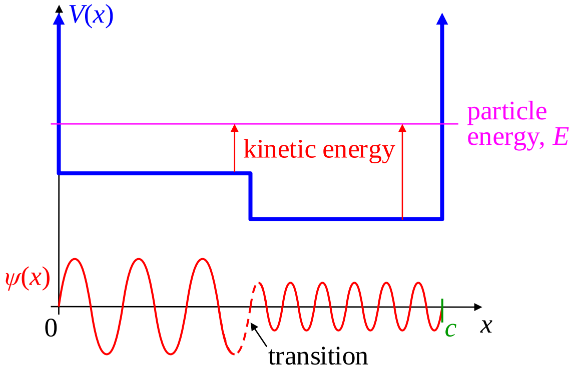


Figure 1.12 Particle of energy *E*, in a box with a step potential bottom. From the correspondence principle, we expect the wave-function to be larger on the left, where the particle is moving "slower."

The stationary state $\psi(x)$ is real, and is everywhere a superposition of positive and negative momenta. The real-valued sinusoids oscillate according to $\left|k(x)\right| = \left[2m\left(E-V\right)\right]^{1/2}/\hbar \equiv p_{cl}/\hbar$. (We will see that since p^2 can be thought of as an average, $p_{cl} \equiv (p^2)^{1/2}$ is not the same as p.) On the left, the kinetic energy ($T = E - V = p^2 / 2m$) is low, so p^2 is low, and $k = p_{cl} / \hbar$ (rad/m) is low (low spatial frequency). On the right, T is higher, p_{cl} is higher, and k is higher (higher spatial frequency). Not only is $\psi(x)$ a superposition of momenta, but it is important that we can associate *different* momenta to different *locations* on the wave-function. p_{cl} is lower on the left, and higher on the right. This introduces the concept of "local momentum," and "local properties" in general.

Concerning the amplitude of $\psi(x)$: on the left of the step-bottom box, the kinetic energy ($E - V$) is lower, so a classical particle moves "slowly." On the right, the kinetic energy is higher, so a classical particle moves "quickly." Classically, then, the particle is more likely to be measured on the left half than the right. The correspondence principle says that at high energy, these classical results must be born out quantum mechanically. The particle "spends more time" on the left side, where it is moving slowly, so the amplitude of ψ is larger (more likely to find the particle there). Conversely, the particle "spends less time" on the right, where it is moving quickly, so the amplitude of ψ is smaller (less likely to find the particle there).

As always, the allowed energies of this potential are quantized by the boundary conditions on ψ, in this case, that $\psi(x) = 0$ at the edges.

If the potential box is wide (many cycles of ψ in each half), then the potential step has only a small transient effect on ψ, indicated as a transition from the sine wave on the left to the sine wave on the right. The perturbation on ψ from the discontinuity decays rapidly away from the transition.

Low energy wave-function: Consider the ground state of a harmonic oscillator:

$$\psi(x) = \left(\frac{m\omega}{\hbar\pi}\right)^{1/4} \exp\left(-\frac{1}{2}\left(x/x_0\right)^2\right) \qquad \text{where} \quad x_0 \equiv \sqrt{\hbar/m\omega} \qquad \text{[Gos 7.18 p144]},$$

and also consider the 1st excited state (both shown in Figure 1.13).

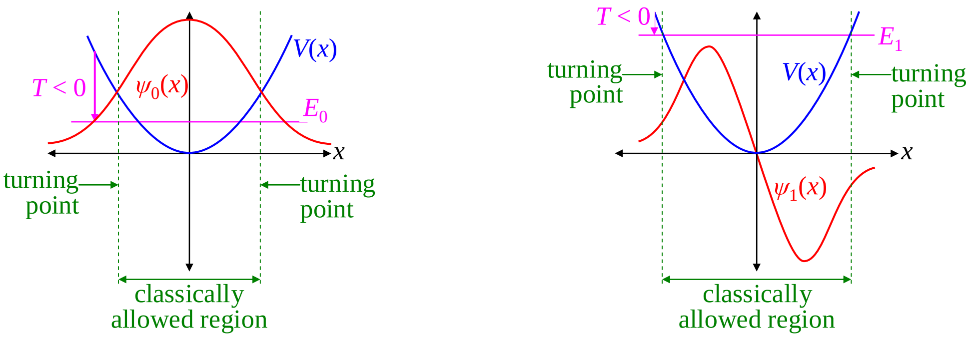


Figure 1.13 Ground state (left), and first excited state (right) of a harmonic oscillator. $E_1 = 3E_0$.

There are several important points to notice, which illustrate general quantum mechanical principles. We discuss the following points:

1. All bound states can have $\psi(x)$ real.
2. At a peak, the local momentum $p = 0$, but $p^2 \neq 0$.
3. In the classically allowed region, $\psi(x)$ turns toward zero: down when it's positive (negative curvature), and up when it's negative (positive curvature).
4. Beyond the classical turning points, $p^2 < 0$, and p is imaginary!

It is less obvious than with the previous wave-function, but each point on this wave-function also has a value of momentum associated with it. We call that the "local momentum," and define it more precisely later, when we examine operators.

Bound states can have $\psi(x)$ real: Simple hamiltonians, $\hat{H} = \hat{p}^2 / 2m + V(x)$, are symmetric with respect to p: any solution for p also has $-p$ as a solution, with the same energy. Since Schrödinger's equation is linear, then the superposition wave function for $p + (-p)$ is also a solution. Their sum is real. For example, ignoring normalization, we have:

$$\psi_p(x) = e^{ipx}, \qquad \psi_{-p}(x) = e^{-ipx} \qquad \text{are wavefunctions for } p \text{ and } (-p).$$

$$\psi(x) = e^{ipx} + e^{-ipx} = 2\cos(px), \qquad \text{is also a solution, and is real}.$$

[With magnetism, the hamiltonian is:

$$\hat{H} = \frac{1}{2m}\left(\hat{\mathbf{p}} - \frac{q}{c}\mathbf{A}(x)\right)^2 + V(x) \qquad \textit{where} \qquad \mathbf{A}(x) \text{ is the magnetic vector potential}.$$

The hamiltonian no longer commutes with $\hat{p}$, and the situation is more complicated.]

Ultimately, all hamiltonians are real, so the time-independent Schrodinger eigenvalue equation is real, and the existence theorem of differential equations says it has a real solution. Therefore all bound states can be real.

At a peak of $\psi(x)$, $\hat{p}\psi(x) = \dfrac{\hbar}{i}\dfrac{\partial\psi(x)}{\partial x} = 0$. We will see that the momentum operator at a point x multiplies the wave-function by the local value of the momentum at that point. When $\partial\psi/\partial x = 0$, there are equal contributions at x from each of $+p$ and $-p$, so they cancel. However, $(p^2)_{local}$ is *not* generally zero at such a point.

In the classically allowed region, $\psi(x)$ turns toward zero: down when it's positive, and up when it's negative. In other words, when the spatial frequency k (aka wave number) is real, ψ oscillates (i.e., we have a *wave*, given enough room). In the ground state of Figure 1.13, there's not enough room in the classically allowed region for ψ to actually cross zero, which is why the original claim is about ψ "turning toward 0," instead of "oscillating." This means every zero of $\psi(x)$ is also an inflection point. From the local momentum, we will find that:

$$\frac{\partial^2 \psi(x)}{\partial x^2} = -\frac{\left(p^2\right)_{local}}{\hbar^2}\psi(x) .$$

In the classically allowed region of space, kinetic energy $T > 0$, so $(p^2)_{local} > 0$. Then, when $\psi(x) > 0$, $\partial^2\psi/\partial x^2 < 0$, and ψ curves downward. When $\psi(x) < 0$, $\partial^2\psi/\partial x^2 > 0$, and ψ curves upward.

Negative kinetic energy and imaginary momentum: Related to wave-function curvature, we see that *beyond the classical turning point*, $(p^2)_{local} < 0$! This surprising result is consistent with $\psi(x)$ being an energy eigenstate: the energy operator, $\hat{H} = \hat{T} + \hat{V}(x)$, must evaluate to the same local energy E at every point of $\psi(x)$. But in the classically disallowed region (beyond the classical turning point), $V(x) > E$, so kinetic energy $T = p^2/2m$ must be less than zero! This is crucially important to orbital angular momentum, which we discuss later. Furthermore, the classically disallowed region adds a negative contribution to the overall averages of T and of p^2.

A negative kinetic energy implies that formally, in the disallowed region, p itself is imaginary. For an energy eigenstate, the local momentum is either real or purely imaginary. For superpositions of energy eigenstates, the local momentum may be complex. We will see later, in the discussion of probability current, that the *real* part of the local momentum has a direct physical meaning. However, negative kinetic energy and imaginary momentum have no classical or simple interpretation. Nonetheless, we accepted their validity as an axiom earlier (#6, p15), and their consequences are fully verified by experiment.

Also, at the classical turning points, $\psi(x)$ has inflection points. (Recall that a change in curvature is defined as an inflection.) In the classically allowed region, ψ curves toward the x-axis, as noted above. In the classically forbidden region, ψ is asymptotic to the x-axis, and therefore curves away from the axis.

1.10 Operators

Operators are an essential part of quantum mechanics.

> The existence of superpositions of states leads to the need for operators
> to compute the properties of such superpositions, because an operator can associate
> a different number with each component of the superposition.

In this section, we focus on operators on functions. Later, we consider operators on discrete-state systems. Before discussing the physical meaning of operators, we must first give some mathematical description.

Operators turn a function of space into another function of space (or more generally, a vector into another vector). In QM, there are three main uses for operators:

1. To extract observable (measurable) attributes of a particle from a wave function. These are *hermitian* operators. The functions resulting from such operators are *not* quantum states.
2. To compute new states from old states, such as the state of a particle or system after some time, or after a rotation. These are *unitary* operators, and their results *are* quantum states.
3. To perform mathematical operations which yield valuable information about quantum mechanics, and which are much more difficult to discover with non-operator methods. These *algebraic* operators are usually neither hermitian nor unitary, but some of them might be either. Their results are not generally quantum states.

We describe each of these uses in more detail shortly. Examples of the three uses of operators:

Observable operators	**State transformation operators**	**Algebraic operators**
$\hat{x}, \hat{p}, \hat{H}, \hat{L}^2, \hat{L}_z, \hat{s}^2, s_z ...$	$\hat{T}(a), \hat{R}(\theta, \phi), \hat{U}(t)$	$\hat{a}, \hat{a}^\dagger, \hat{J}_+, \hat{J}_-$

> An operator acting on a spatial function $\psi(x)$ produces another spatial function.

Operators are part of the algebra of quantum mechanics; operators are written in both wave-function notation, and Dirac notation (described later). Operators exist for both spatial wave-functions, and discrete states, such as particle spin. In quantum mechanics, all operators [except time-reversal] are linear operators.

> A linear operator acting on a superposition, say $\psi(x)$, produces a superposition of *results* based on the superposition of *functions* which compose $\psi(x)$.

We distinguish between a **quantum state**, which is wave-function that describes all the spatial properties of a particle, and more general functions of space which have other uses but are *not* quantum states. The result of an operator may be either a quantum state, or some other function of space that is not a quantum state. For example:

$$\hat{p}\psi(x) = \frac{\hbar}{i}\frac{d}{dx}\psi(x) \qquad \textit{where} \quad \hat{p} \text{ is a hermitian operator ,}$$

is a function of space (function of x), but is not a quantum "state," because it is not a wave-function. It is not normalized, and has the wrong units for a wave-function. In contrast, given a 2D wave-function, $\psi(x, y)$, which *is* a state, we can rotate it in the x-y plane:

$$\hat{R}(\theta)\psi_1(x,y) \equiv \psi_2(x,y) = \psi_1\left(x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta\right) \quad \textit{where} \quad \hat{R}(\theta) \text{ is a unitary operator .}$$

This *is* a new quantum state, computed from the old one.

> Some references do not properly distinguish between a "spatial function" and a quantum "state."
> Some even define an "operator" as a thing which acts on a "state" to produce another "state."
> This is incorrect. An operator acts on a function to produce another function.
> Either one or both functions may or may not be quantum "states."

For example, imagine a particle state $\psi(x)$ = 1/2 between [0,4], and 0 elsewhere (Figure 1.14, left). Consider also the position operator acting on it, $\hat{x}\psi(x) = x\psi(x) = x/2$ between [0, 4], and 0 elsewhere (Figure 1.14, middle).

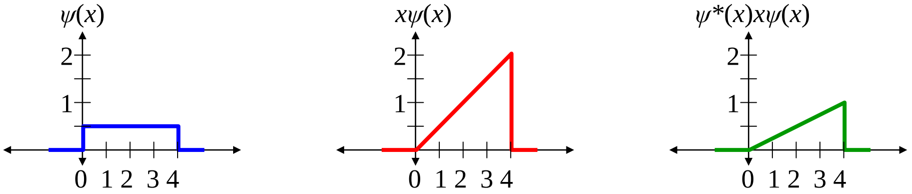


Figure 1.14 (Left) A normalized wave-function, $\psi(x)$. (Middle) $x\psi(x)$. (Right) $\psi^*(x)x\psi(x)$.

Note that, though $\psi(x)$ *is* a quantum state, $\hat{x}\psi(x)$ is not a state, because it is not a wave-function: it doesn't define the properties of a particle or system; it's not even normalized, and it has the wrong units! However, it can be used to calculate things related to the particle's position, as shown later.

Operators may also act on a given function which is *not* a quantum state, e.g. the result of some other operator on a state. Such a given function may still be a superposition of results. A second operator acting on such a function which is *not* a state produces a new superposition of new results based on the superposition of old results in the given function. For example, in two dimensions, we have wave-functions $\psi(x, y)$. The $\hat{L}_z$ operator (angular momentum about the z-axis) is a composition of $\hat{x}, \hat{p}_y, \hat{y}$, and $\hat{p}_x : \hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x$. $\hat{L}_z$ typically acts on a state, which means that in the first term, $\hat{p}_y = \frac{\hbar}{i}\frac{\partial}{\partial y}$ acts first on the given state to produce a function of space which is *not* a state. Then $\hat{x}$ acts on *that* spatial function to produce a new spatial function, which is also *not* a state. Similarly for the second term, $-\hat{y}\hat{p}_x$. The result of $\hat{L}_z$ acting on a state is a spatial function that is not a state.

1.10.1 Linear Operators

In quantum mechanics, all operators [except time-reversal] are linear operators. A linear operator produces a superposition of *results* based on the superposition of *components* in the given function.

A **linear operator** *distributes* over addition,
and *commutes* with scalar multiplication:

$$\hat{L}\left[a\psi(x)+b\phi(x)\right]=a\hat{L}\psi(x)+b\hat{L}\phi(x)\,.$$

Examples: Is multiplication by x a linear operator? Let's see:

$$x\left[a\psi(x)+b\phi(x)\right]=ax\psi(x)+bx\phi(x)\,. \qquad \text{It's linear!}$$

How about multiplication by x^2 ? $x^2\left[a\psi(x)+b\phi(x)\right]=ax^2\psi(x)+bx^2\phi(x)$. It's also linear.

How about multiplication by cos(x)? Or multiplication by any arbitrary function $f(x)$?

$$f(x)\left[a\psi(x)+b\phi(x)\right]=af(x)\psi(x)+bf(x)\phi(xs)\,. \qquad \text{It's linear!}$$

Is squaring a spatial function a linear operation?

$$\left[a\psi(x)+b\phi(x)\right]^2=a^2\psi^2(x)+2ab\psi(x)\phi(x)+b^2\phi^2(x)\neq a\psi^2(x)+b\phi^2(x)\,. \qquad \text{It's \textit{not} linear.}$$

Note that complex conjugation, the Re{}, and Im{} operators are all *nonlinear*.

Composition of operators: We frequently operate on a function with one operator, and then operate on that result with another operator. For example, we may first operate with the operator 'x', which means multiply the function by 'x' everywhere: $\psi(x) \to x\psi(x)$. We may then operate on that result by differentiating w.r.t. x: $\frac{\partial}{\partial x}x\psi(x)$. Acting with one operator and then another, is called **composition** of operators. The composition of $\hat{A}$ on $\hat{B}$ is $\hat{A}\hat{B}$. Because operators act to the right, this means act with $\hat{B}$ *first*, then act with $\hat{A}$. Many references call the composition of two operators the "product" of the two, and sometimes even "multiplying" the two. In such a case, there is generally no multiplication involved (unless one of the operators is a "multiply by" operator).

1.10.2 Operator Algebra

Operators have their own algebra, or rules, for performing mathematical operations and for being manipulated. In wave mechanics, operators are linear operators on continuous functions. (Matrix mechanics is discussed later.) Some operators are simple and straightforward:

$$\frac{\partial}{\partial x} \qquad \text{This is a linear operator: it takes a derivative with respect to '}x\text{'.}$$

However, some other operators are more involved. For example, '$\left(\frac{\partial}{\partial x}\right)x$' can be a compound operator. It means "multiply by x, and then take the derivative with respect to x." As an operator, it is *not* an arithmetic expression (where $\left(\frac{\partial}{\partial x}\right)x=1$). Instead,

$$\left(\frac{\partial}{\partial x}\right)x=1+x\frac{\partial}{\partial x} \qquad \text{(an operator equation)}\,.$$

How can this be? What happened to basic calculus? The key here is to distinguish the *operator* from the *operand*. The **operator** is the action to be taken. The **operand** is the object on which the action is taken. For example:

$\frac{\partial}{\partial x}$ ‘ $\frac{\partial}{\partial x}$ ’ is an operator. There is no operand or result here.

$\left(\frac{\partial}{\partial x}\right)x = 1$ The operator is ‘ $\frac{\partial}{\partial x}$ ’, the operand is ‘x’, and the result is ‘1’.

But in the case of the operator equation:

$\left(\frac{\partial}{\partial x}\right)x = 1 + x\frac{\partial}{\partial x}$, ‘ $\left(\frac{\partial}{\partial x}\right)x$ ’ is an operator, and ‘ $1 + x\frac{\partial}{\partial x}$ ’ is an *equivalent* operator.

In other words, this is an **operator identity**. The above is not an *arithmetic* equation; it is a statement of equivalence of two operators. There is no operand in the above equation, and no result (just like writing ‘$\partial/\partial x$’ by itself). Let’s establish the equivalence in the **operator equation** (aka **operator identity**) above:

$\left(\frac{\partial}{\partial x}\right)x\,\psi(x)$ We insert an arbitrary operand, $\psi(x)$, so we can operate on it.

$= \left(\frac{\partial}{\partial x}\right)\big(x\,\psi(x)\big)$ The operator says to first multiply by x, then differentiate w.r.t. x.

$= \left[\left(\frac{\partial}{\partial x}\right)(x)\right]\psi(x) + x\frac{\partial}{\partial x}\psi(x)$ Using the product rule for derivatives.

$= \psi(x) + x\frac{\partial}{\partial x}\psi(x)$ because $\left(\frac{\partial}{\partial x}\right)x = 1$, where x is an *operand* here.

$= \left(1 + x\frac{\partial}{\partial x}\right)\psi$ “factoring out” ψ, because linear operators distribute over addition.

Thus the operator ‘ $\left(\frac{\partial}{\partial x}\right)x$ ’ is exactly equivalent to the operator ‘ $1 + x\frac{\partial}{\partial x}$ ’.

When you see something like ‘ $\left(\frac{\partial}{\partial x}\right)x$ ’, how do you know if this is an arithmetic expression (with an operator ‘ $\frac{\partial}{\partial x}$ ’ and operand ‘x’), or just an operator? You can only tell from context. In other cases, though, operators may have a “hat” over them, to explicitly indicate they are operators.

Let’s now derive an example of operator equivalence without actually inserting an arbitrary operand. We keep the operand silently in our minds, and write our steps as we operate on this hypothetical operand:

$\left(\frac{\partial}{\partial x}\right)x^2$ What is an equivalent operator to this?

$= \left[\frac{\partial}{\partial x}x^2\right] + x^2\frac{\partial}{\partial x}$ Using product rule for derivatives; in the 1st term, x^2 is an *operand*.

$= 2x + x^2\frac{\partial}{\partial x}$ Because $\frac{\partial}{\partial x}x^2 = 2x$, where x^2 is an *operand* here.

Notice how there are fewer steps compared to the previous derivation, because we didn’t actually insert an operand, and then remove it at the end. However, this shortcut method may take some getting used to.

Beware that many pairs of operators do not commute, i.e. $\hat{A}\hat{B} \ne \hat{B}\hat{A}$., so keep operators in order. In ordinary algebra, we often write:

$$(a+b)^2 = a^2 + 2ab + b^2 .$$

When 'a' and 'b' are numbers (or functions), this is fine. But if instead of numbers, we had operators $\hat{A}$ and $\hat{B}$, this would be wrong, because we changed the order of 'a' and 'b' in the second term. Instead, we should write:

$$\left(\hat{A}+\hat{B}\right)^2 = \hat{A}^2 + \hat{A}\hat{B} + \hat{B}\hat{A} + \hat{B}^2 .$$

For example, the hamiltonian of a charged particle in a magnetic field includes:

$$\left(\hat{\mathbf{p}} - \frac{q}{c}\mathbf{A}(x)\right)^2 = \hat{\mathbf{p}}^2 - \frac{q}{c}\left(\hat{\mathbf{p}}\mathbf{A}(x) + \mathbf{A}(x)\hat{\mathbf{p}}\right) + \frac{q^2}{c^2}\mathbf{A}^2(x) \qquad \textit{where} \quad \hat{\mathbf{p}} \text{ and } \mathbf{A}(x) \text{ do not commute} .$$

This preservation of ordering is also required for matrix operators in discrete-state systems. (See the later section on commutators for more on operator identities.)

1.10.3 Operators: What Is the Meaning of This?

We continue our focus here on wave-mechanics, i.e. operating on wave-functions such as $\psi(x)$. Later, we'll discuss operators in Dirac notation, and in discrete state quantum mechanics.

1.10.3.1 Think Globally, Act Locally: Local Values, Local Operators

Just what do operators do? What is the meaning of some arbitrary operator, $\hat{o}\psi(x)$? Let's start simply, with a spatial wave function $\psi(x)$, in the position basis. At each point in space, the wave function defines the particle density (or probability density of finding the particle). Also, for a given $\psi(x)$, a given *operator* has a **local value** at each point in space (i.e. for each x), which generally depends on $\psi(x)$. In other words, at each point in space, the wave function and the operator *together* define the local value of the operator. We could write this local value as a function of position, say $o_\psi(x)$, where the subscript ψ indicates that the local value of the operator at each point x was computed from a given $\psi(x)$. The result of an operator acting on a spatial function $\psi(x)$ is to simply multiply $\psi(x)$ at each point in space by the local value of the operator at that point in space, $o_\psi(x)$:

$$\hat{o}\psi(x) = o_\psi(x)\psi(x) . \tag{1.4}$$

The operator $\hat{o}$ is a mathematical operation that turns $\psi(x)$ into $o_\psi(x)\psi(x)$.

For example, in the position basis, the simplest operator is the position operator, $\hat{x}$:

$$\hat{x}\psi(x) = x_{local}(x)\psi(x) .$$

But what is the local position associated with the position 'x'? It is simply 'x'! Therefore,

$$x_{local}(x) = x \qquad \Rightarrow \qquad \hat{x}\psi(x) = x\psi(x) .$$

We will show shortly that we can use this to compute the *overall* average value of 'x' (the particle's position), from its wave-function $\psi(x)$:

$$\langle \hat{x} \rangle = \int_\infty \psi^*(x)\, x\, \psi(x)\, dx .$$

A more complicated example is "local energy," which is an important quantity. In computational quantum chemistry, one can find numerical solutions to the time-independent Schrödinger equation by starting with a trial wave-function (a trial solution), and adjusting it according to its local energy at each point, to construct the next iteration of trial wave-function. But how can we compute $E_{local}(x)$? We use the definition of an operator:

$$\hat{H}\psi(x) = E_{local}(x)\psi(x) \qquad \Rightarrow \qquad E_{local}(x) = \frac{\hat{H}\psi(x)}{\psi(x)} .$$

For the non-magnetic hamiltonian, we have:

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(x) = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)$$

$$\Rightarrow \qquad E_{local}(x) = \frac{\hat{H}\psi(x)}{\psi(x)} = \frac{\left[-\left(\hbar^2/2m\right)\dfrac{\partial^2}{\partial x^2} + V(x)\right]\psi(x)}{\psi(x)} = \frac{-\left(\hbar^2/2m\right)\dfrac{\partial^2}{\partial x^2}\psi(x)}{\psi(x)} + V(x).$$

Given some trial wave-function $\psi(x)$, one can evaluate the local energy from this formula.

In general, then, we *define* the local value of any operator as:

$$o_{local}(x) \equiv \frac{\hat{o}\psi(x)}{\psi(x)}, \qquad \text{provided } \psi(x) \ne 0.$$

There is a subtlety if $\psi(x)$ is zero somewhere, because then the local value is undefined. This is usually not a problem, because the "local density" *is* well defined, as we will see later. However, for well-behaved wave-functions, a zero of $\psi(x)$ is a "removable singularity" in the local value, and can be "filled" by its limiting value:

$$\text{If} \qquad \psi(x_0) = 0, \qquad \text{then} \qquad o_{local}(x_0) \equiv \lim_{x \to x_0} o_{local}(x).$$

A **local operator** is an operator determined only by ψ at x, or ψ in an infinitesimal neighborhood of x. Thus $\partial/\partial x$ is a local operator. Some operators depend on more than one point of ψ, or even on *all* of $\psi(x)$; they are **nonlocal operators**.

All common observables are local operators.

In general, an operator can act on any function of space; it need not be a quantum state. If a spatial function is not a state, then what is it? It is usually the result of some other operator acting on a state, but it could instead be a given potential (e.g., $V(x)$ or $\mathbf{A}(x)$), or one of various other functions of space. Such a spatial function encodes information. To get at that information, you usually have to take a dot product with another relevant spatial function.

Summary: *An operator acting on a given spatial function* $\psi(x)$ *produces another spatial function.* The given function may or may not be a quantum *state*, and the resulting function may or may not be a quantum *state*, i.e. it may or may not represent a quantum state that a particle or system could be in.

Here is a summary of some common operators:

Operator	**Position basis representation**	**Comments**
position, $\hat{x}$	x	multiplies $\psi(x)$ by position x. The value of x does *not* depend on ψ.
potential energy, $\hat{V}$	$V(x)$	multiplies $\psi(x)$ by potential $V(x)$. The value of $V(x)$ does *not* depend on ψ.
momentum, $\hat{p}$	$-i\hbar\,\partial/\partial x$	multiplies $\psi(x)$ by momentum $p_{local}(x)$. The value of $p_{local}(x)$ *does* depend on ψ.
kinetic energy, $\hat{T}$	$\dfrac{\hat{p}^2}{2m} = \dfrac{-\hbar^2}{2m}\dfrac{\partial^2}{\partial x^2}$	multiplies $\psi(x)$ by kinetic energy $T_{local}(x)$. The value of $T_{local}(x)$ *does* depend on ψ.
Hamiltonian, $\hat{H}$	$\dfrac{\hat{p}^2}{2m} + \hat{V}(x) = \dfrac{-\hbar^2}{2m}\dfrac{\partial^2}{\partial x^2} + V(x)$	just a sum of kinetic and potential energy.

total energy, $\hat{E}$	$i\hbar\, \partial/\partial t$	multiplies $\psi(t, x)$ by energy $E_{local}(t, x)$. The value of $E_{local}(t, x)$ *does* depend on $\psi(t, x)$. Note that $\partial/\partial t$ works in the momentum basis, and other bases, as well.

1.10.3.2 Operators as Eigenfunction/Eigenvalue Weighting

Another way to think of a linear operation on a given function is this: the operator mentally decomposes the given function into a superposition (weighted sum) of the operator's own eigenfunctions, then multiplies each eigenfunction component by its eigenvalue, and then sums up the result. I.e.,

$$\text{Given:} \quad \phi_n(x) \equiv \text{eigenfunctions of } \hat{o}; \qquad o_n \equiv \text{eigenvalue for } \phi_n(x),$$

$$\text{and} \quad \psi(x) = \sum_{n=1}^{\infty} c_n \phi_n(x), \qquad \text{then} \quad \hat{o}\psi(x) = \sum_{n=1}^{\infty} c_n o_n \phi_n(x).$$

In other words, the operator "factors in" its eigenvalues as additional weights to the eigenfunction components of $\psi(x)$. We will examine more details of operator behavior in discrete bases in the chapter on matrix mechanics.

1.10.3.3 Computing Measurable Results, Local Density

Example: position: As an example of extracting a measurable attribute of a particle from a wave function, let's consider the position operator, $\hat{x}$. From our earlier discussion, we know that:

$$\hat{x}\psi(x) = x_{local}(x)\psi(x) = x\psi(x).$$

We use this to compute the *overall* average value of 'x', the particle's position, from its wave function $\psi(x)$:

$$\langle \hat{x} \rangle = \int \psi^*(x)\, x\, \psi(x)\, dx.$$

For example, in Figure 1.14 (left), we had $\psi(x) = 1/2$ between [0, 4], and 0 elsewhere. The average value of the particle's position is

$$\langle \hat{x} \rangle = \int_{-\infty}^{\infty} \psi^*(x)\, x\, \psi(x)\, dx = \int_0^4 \frac{1}{2} x \frac{1}{2}\, dx = \frac{1}{8} x^2 \Big|_0^4 = 2,$$

which we can also see by inspection.

You can think of the "average value" equation, $\langle \hat{o} \rangle = \int \psi^*(x)\, \hat{o}\, \psi(x)\, dx$, as a direct statistical computation of an average value from a probability distribution function (PDF), where $\text{pdf}(x) = |\psi(x)|^2 = \psi^*(x)\psi(x)$. Then statisticians would write

$$\langle \hat{o} \rangle = \int o_{local}(x)\, \text{pdf}(x)\, dx = \int o_{local}(x)\, \psi^*(x)\psi(x)\, dx.$$

So why do we write '$\hat{o}$' in between ψ^* and ψ? Because the operator notation of QM is that an operator acts on the function to its right (or sometimes in a different way, it acts to the left). Therefore:

$$\int \psi^*(x)\left(\frac{\hbar}{i}\frac{\partial}{\partial x}\right)\psi(x)\, dx \quad \text{is quite different from} \quad \int \left(\frac{\hbar}{i}\frac{\partial}{\partial x}\right)\big(\psi^*(x)\psi(x)\big)\, dx.$$

In QM, we must write the operator between ψ^* and ψ, so that it acts on ψ:

$$\begin{aligned}\langle \hat{o} \rangle &= \int o_{local}(x)\,\mathrm{pdf}(x)\,dx = \int o_{local}(x)\,\psi^*(x)\psi(x)\,dx = \int \psi^*(x)\,\underbrace{o_{local}(x)\,\psi(x)}_{\hat{o}\psi(x)}\,dx \\ &= \int \psi^*(x)\,\hat{o}\psi(x)\,dx\,.\end{aligned} \tag{1.5}$$

Example: momentum in position basis: An operator such as momentum includes a derivative: $\hat{p} = (\hbar/i)\partial/\partial x$. [Therefore, the value of the momentum at any point in space depends not only on the value of $\psi(x)$ at x, but on $\psi(x)$ in an infinitesimal neighborhood *near* x.] How can we understand this operator for momentum? Recall that one of the starting points of wave mechanics was that particles of definite momentum have a definite spatial frequency (wave-number), $k = p/\hbar$. This implies the eigenfunctions of momentum are (ignoring normalization):

$$\phi_p(x) = e^{ipx/\hbar} \qquad \text{for a definite momentum, } p\,.$$

This particle has the same momentum, p, everywhere, so $p_{local}(x) = p$ (a constant). Then by the definition of operators, the momentum operator must take $\phi_p(x)$ to $p\phi_p(x)$. That is:

$$\hat{p}e^{ipx/\hbar} = p_{local}(x)\psi(x) = pe^{ipx/\hbar}\,.$$

What linear operator brings down the p from the exponent, as a multiplier in front? The derivative w.r.t. x:

$$\frac{\partial}{\partial x}e^{ipx/\hbar} = \frac{i}{\hbar}pe^{ipx/\hbar}\,.$$

This is almost what we need. We fix the prefactor by simply multiplying by $\hbar/i$

$$\frac{\hbar}{i}\frac{\partial}{\partial x}e^{ipx/\hbar} = pe^{ipx/\hbar} \qquad \Rightarrow \qquad \hat{p} = \frac{\hbar}{i}\frac{\partial}{\partial x}\,.$$

We derived the momentum operator from the definition of operators,
and the empirical fact that particles of definite momentum have definite wave-number k.

Then for a general state $\psi(x)$, at each point x there is a local value of momentum: $p_{local}(x) = \dfrac{\hbar}{i}\dfrac{\psi'(x)}{\psi(x)}$, provided $\psi(x) \neq 0$. We can then compute the average value of p for the particle in the standard way:

$$\begin{aligned}\langle \hat{p} \rangle &= \int_{-\infty}^{\infty} p_{local}(x)\,\mathrm{pdf}(x)\,dx = \int_{-\infty}^{\infty} \psi^*(x)\,p_{local}(x)\psi(x)\,dx \\ &= \int_{-\infty}^{\infty} \psi^*(x)\left(\frac{\hbar}{i}\frac{d}{dx}\right)\psi(x)\,dx = \int_{-\infty}^{\infty} \psi^*(x)\,\hat{p}\psi(x)\,dx\,.\end{aligned}$$

It is instructive to see how this works when $\psi(x)$ is a superposition of two momenta, p_1 and p_2:

$$\psi(x) = A\exp(ip_1x/\hbar) + B\exp(ip_2x/\hbar), \qquad A, B \text{ complex}, |A|^2 + |B|^2 = 1 \qquad \Rightarrow$$

$$p_{local}(x) = \frac{\hbar}{i}\frac{\psi'}{\psi} = \frac{A\exp(ip_1x/\hbar)\,p_1 + B\exp(ip_2x/\hbar)\,p_2}{A\exp(ip_1x/\hbar) + B\exp(ip_2x/\hbar)}\,.$$

This is a weighted average of p_1 and p_2, where the weights depend on both the coefficients A and B, and on the values of the momentum eigenfunctions at x. The weights are complex, and given explicitly by:

$$p_{local}(x) = \frac{w_1p_1 + w_2p_2}{w_1 + w_2} \quad \textit{where} \quad w_1 = A\exp(ip_1x/\hbar), \quad \textit{and} \quad w_2 = B\exp(ip_2x/\hbar)\,.$$

Momentum squared: We noted earlier that at an extremum of a *real*-valued $\psi(x)$, $\partial\psi/\partial x = 0$, and the local momentum $p_{local}(x) = 0$. This can be thought of as a weighted average of two different momenta: one at $+p$ and one at $-p$. But when considering momentum-squared at that same extremum, $\hat{p}^2\psi(x) = -\hbar^2 \dfrac{\partial^2\psi(x)}{\partial x^2}$, the contributions from $(+p)^2$ and $(-p)^2$ add; they don't cancel. In other words:

$$\left(p_{local}(x)\right)^2 \ne p^2{}_{local}(x) \quad where \quad p^2{}_{local}(x) \equiv \frac{\hat{p}^2\psi(x)}{\psi(x)} = \frac{-\hbar^2\psi''(x)}{\psi(x)} .$$

This implies that globally, $<p>^2 \ne <p^2>$, as is well-known. A similar situation exists with angular momentum $\left(\hat{L}^2\right)$, which we will examine more closely later.

Local density: We've noted that the overall average of an operator is an integral over all space of the local value, weighted by the particle density (PDF) at each point:

$$\left\langle \hat{o} \right\rangle = \int_\infty o_{local}(x)\, \psi^*(x)\psi(x)\, dx = \int_\infty \psi^*(x)\, \hat{o}\psi(x)\, dx .$$

This suggests we define a **local density** of $\hat{o}$ as:

$$\rho_o(x) = o_{local}(x)\, \psi^*(x)\psi(x) = \psi^*(x)\, \hat{o}\psi(x), \text{e.g.,} \qquad \rho_p = \psi^*(x)\, \hat{p}\psi(x) .$$

A local density is analogous to physical densities such as mass density. Then the global average is just the integral over all space of the local density:

$$\left\langle \hat{o} \right\rangle = \int_\infty \rho_o(x)\, dx = \int_\infty \psi^*(x)\, \hat{o}\psi(x)\, dx, \qquad \text{e.g.,} \qquad \left\langle \hat{p} \right\rangle = \int_\infty \rho_p(x)\, dx = \int_\infty \psi^*(x)\, \hat{p}\psi(x)\, dx .$$

Conceptually, the local density can be thought of, in most single-particle cases, as describing the actual distribution in space of a physical quantity, such as momentum density. Also:

> Local density has an advantage over the local value, $o_{local}(x)$,
> because it is well defined, even when $\psi(x) = 0$.

When $\psi(x) = 0$, the local density is zero, regardless of the local value at that point, since (roughly) no part of the particle exists at that point.

Even non-wave properties, such as a particle's intrinsic spin, have local densities given by the wave-function. E.g., for a spin-up electron, its spin angular momentum density is:

$$\rho_{spin}(x) = \frac{\hbar}{2}\psi^*(x)\psi(x) \qquad \left[\text{from Bay 14-92 p326m}\right].$$

1.10.3.4 New States From Old States: Time Evolution, and More

An example of an operator producing a new state from an old state is the time-evolution operator. This section requires an understanding of composing an arbitrary state from energy eigenstates.

The time evolution operator takes a wave-function at some time t_0, and produces what the wave-function will be at a later time t_1. Therefore, the time-evolution operator is really a family of operators, parameterized by both t_0 and t_1. In the simple case where the hamiltonian is not explicitly time-dependent, the time evolution operator depends only on the time difference, $\Delta t \equiv t_1 - t_0$. Then, for arbitrary t:

$$\psi(t+\Delta t, x) = \hat{U}(\Delta t)\psi(t, x) \ where \quad \hat{U}(\Delta t) \equiv \text{time evolution operator,}$$
$$\hat{H} \text{ independent of time} .$$

For this simple case of time-independent hamiltonian, we can derive the time-evolution operator from the fact that energy eigenstates $u_n(t, x)$, with energy E_n, evolve in time with a particularly simple form. The spatial form of energy eigenstates does *not* change with time. By definition, energy eigenstates satisfy the time-independent Schrödinger equation $\hat{H}u_n(t,x) = E_n u_n(t,x)$. Then for the full Schrödinger equation, we have:

$$i\hbar\frac{\partial}{\partial t}u_n(t,x) = \hat{H}u_n(t,x) = E_n u_n(t,x) \qquad \Rightarrow \qquad \frac{\partial}{\partial t}u_n(t,x) = -\frac{iE_n}{\hbar}u_n(t,x) .$$

This is of the form $y'(t) = ky(t)$, with solution (by inspection) $y = y(0)e^{kt}$. Also, by a simple time-shift, $y(t+\Delta t) = y(t)e^{k\Delta t}$. Applying this form to our equation for $u_n(t, x)$, we get the time evolution for an energy eigenstate:

$$u_n(t+\Delta t, x) = e^{-iE_n\Delta t/\hbar}u_n(t,x) \qquad \text{time evolution of an energy eigenstate} . \tag{1.6}$$

Time evolution simply multiplies the wave-function by a complex phase, $\exp(-iE_n t/\hbar)$. Since phase does not affect any quantum property, all the properties of such a system are independent of time; such a state is called **stationary**. Note that stationary does not imply "static:" a stationary state can be moving (such as an electron orbiting a nucleus).

Decomposing an arbitrary state $\psi(t, x)$ into energy eigenstates yields the general time evolution. At an arbitrary time t:

$$\psi(t,x) = \sum_n c_n u_n(t,x) \qquad \Rightarrow \qquad \psi(t+\Delta t, x) = \hat{U}(\Delta t)\psi(t,x) = \sum_n e^{-iE_n\Delta t/\hbar}c_n u_n(t,x) . \tag{1.7}$$

This gives the explicit form of time evolution of a state from its energy eigenstate components. We now show how to write this as an operator, independent of components. Recall that the exponential of a linear operator is another linear operator, and is defined by the power series expansion of the exponential. For example:

$$e^{\hat{a}} = \mathbf{\hat{1}} + \hat{a} + \frac{\hat{a}^2}{2!} + \frac{\hat{a}^3}{3!} + ... \qquad \Rightarrow \qquad e^{-i\hat{H}\Delta t/\hbar} = \mathbf{\hat{1}} + \frac{-i\Delta t\hat{H}}{\hbar} + \frac{1}{2!}\left(\frac{-i\Delta t\hat{H}}{\hbar}\right)^2 + \frac{1}{3!}\left(\frac{-i\Delta t\hat{H}}{\hbar}\right)^3 + ...$$

Thus the exponential of the hamiltonian operator is a power series of hamiltonian operators. When acting on an energy eigenstate $u_n(t, x)$, each appearance of $\hat{H}$ gets replaced by the energy E_n. Therefore,

$$e^{-i\hat{H}\Delta t/\hbar}u_n(t,x) = \left[\mathbf{\hat{1}} + \frac{-i\Delta tE_n}{\hbar} + \frac{1}{2!}\left(\frac{-i\Delta tE_n}{\hbar}\right)^2 + \frac{1}{3!}\left(\frac{-i\Delta tE_n}{\hbar}\right)^3 + ...\right]u_n(t,x) = e^{-iE_n\Delta t/\hbar}u_n(t,x)$$

Thus we see that for an energy eigenstate, the exponential of the hamiltonian (times some factors) gives the simple form for time evolution of an energy eigenstate shown in (1.6). But again, any state can be written as a sum of energy eigenstates, and $e^{-i\hat{H}\Delta t/\hbar}$ is linear, so $e^{-i\hat{H}\Delta t/\hbar}$ gives the time evolution of an *arbitrary* state shown in (1.7).

$$\begin{aligned} e^{-i\hat{H}\Delta t/\hbar}\psi(t,x) &= \sum_n e^{-i\hat{H}\Delta t/\hbar}c_n u_n(t,x) = \sum_n e^{-iE_n\Delta t/\hbar}c_n u_n(t,x) = \psi(t+\Delta t, x) \\ &= \hat{U}(\Delta t)\psi(t,x) . \end{aligned} .$$

Therefore, the general time evolution operator is:

$$\hat{U}(\Delta t) = e^{-i\hat{H}\Delta t/\hbar} \qquad \text{time evolution operator for time-independent hamiltonian} .$$

This is a pure operator equation, and is independent of any representation of the wave-function.

We note in passing that, in general, the future value of ψ at some point x_0, $\psi(t, x_0)$, depends on *all* the values of $\psi(0, x)$ for all x. Therefore:

The time-evolution operator, $\hat{U}(\Delta t)$, is a nonlocal operator.

More on nonlocal operators later.

In the general case, where the hamiltonian explicitly depends on time, we must keep t_0 and t_1:

$$\psi(t_1, x) = \hat{U}(t_1, t_0)\psi(t_0, x) \quad where \quad \hat{U}(t_1, t_0) \equiv \text{time evolution operator, } \hat{H} \text{ depends on time} .$$

Time evolution for a time-dependent hamiltonian is much more complicated, and we do not address that here.

A second example of "new states from old states" is rotation operators, e.g. in 3D, $\psi_{rotated}(\mathbf{r}) = \hat{R}(\theta, \phi)\psi(\mathbf{r})$. Here again, the rotation "operator" is really a family of operators, parameterized by the rotation angles, θ and ϕ. We return to rotations later when considering generators.

1.10.3.5 Simplifying Calculations

A third use for operators is to aid and simplify calculations. For example, raising and lowering operators ($\hat{a}^\dagger$, $\hat{a}$, $\hat{J}_+$, $\hat{J}_-$) do not represent observables, and have nothing to do with physically changing a state, or adding/removing energy (or angular momentum). They are used for computing matrix elements, for analysis, proving relationships and theorems, perturbation theory, etc.

The harmonic oscillator raising and lower operators, $\hat{a}$ and $\hat{a}^\dagger$ (often written without hats), are usually defined as acting on entire states:

$$\hat{a}|n\rangle = \sqrt{n}\,|n-1\rangle \qquad\qquad \hat{a}^\dagger|n\rangle = \sqrt{n+1}\,|n+1\rangle .$$

From this definition, you might think that they are nonlocal operators (since they seem to depend on the entire state). However, it turns out that these definitions are achieved with local operators. We can see this from the formulas for $\hat{a}$ and $\hat{a}^\dagger$ in terms of $\hat{x}$ and $\hat{p}$, which are both local operators:

$$\hat{a} = \sqrt{\frac{m\omega}{2\hbar}}\hat{x} + \frac{i}{\sqrt{2m\hbar\omega}}\hat{p}, \qquad \hat{a}^\dagger = \sqrt{\frac{m\omega}{2\hbar}}\hat{x} - \frac{i}{\sqrt{2m\hbar\omega}}\hat{p} \quad \Rightarrow \quad \hat{x} = \sqrt{\frac{\hbar}{2m\omega}}\left(\hat{a} + \hat{a}^\dagger\right).$$

These definitions and formulas allow us to evaluate common inner products (aka "matrix elements") without any integration:

$$\langle 2|\hat{x}^2|0\rangle = \langle 2|\left[\sqrt{\frac{\hbar}{2m\omega}}\left(a + a^\dagger\right)\right]^2|0\rangle = \frac{\hbar}{2m\omega}\langle 2|\left(a^\dagger\right)^2|0\rangle = \frac{\hbar\sqrt{2}}{2m\omega}\cancel{\langle 2|2\rangle} = \frac{1}{\sqrt{2}}\frac{\hbar}{m\omega}.$$

This would be substantially more work to evaluate by integrating products of Hermite polynomials and exponentials.

1.10.3.6 Non-Local Operators

The time evolution operator, $\hat{U}(t_1, t_0)$ (described above), takes a quantum state (a function of space) into another quantum state (another function of space). Recall that for time-independent hamiltonians:

$$\hat{U}(\Delta t) = e^{-i\hat{H}\Delta t/\hbar} = \hat{\mathbf{1}} + \frac{-i\Delta t\hat{H}}{\hbar} + \frac{1}{2!}\left(\frac{-i\Delta t\hat{H}}{\hbar}\right)^2 + \frac{1}{3!}\left(\frac{-i\Delta t\hat{H}}{\hbar}\right)^3 + \ldots \qquad (\text{time-independent } \hat{H}).$$

For a finite time interval, the new value of the wave-function at a given point x depends not only on the infinitesimal neighborhood of x, but on values of ψ far away from x. Therefore, the time evolution operator for a finite time interval is a *non*local operator. It may be surprising that the time evolution operator is nonlocal, since it is written as a Taylor series sum of (local) hamiltonian operators. This is a subtle issue that requires careful examination, but briefly, the hamiltonian requires an infinitesimal *neighborhood* around the point x to evaluate the momentum. Then, integrating over an infinite number of such infinitesimal neighborhoods makes the time evolved state at any point x depend on distant points of the original state. (The space translation operator has a similar characteristic.)

[This nonlocality is evident in the more advanced concept of "propagators." Propagators are the time evolution of a localized particle (delta-function) in space, and for all finite times, they extend to infinity.] We do not consider this nonlocality of operators further.

1.10.4 Commutators

The commutator of two operators is also an operator. Commutators are compound operators built from other operators. A commutator is written with square brackets, and defined as the difference between the operators acting in both orders:

$$\left[\hat{A},\hat{B}\right] = \hat{A}\hat{B} - \hat{B}\hat{A} .$$

We usually think of operators as acting to the right, so the first term above has $\hat{B}$ acting first.

Commutators are often a convenient way to specify operator identities:

$$[\hat{x},\hat{p}] = i\hbar \qquad \leftrightarrow \qquad \hat{x}\hat{p} - \hat{p}\hat{x} = i\hbar \qquad \leftrightarrow \qquad \hat{x}\hat{p} = i\hbar + \hat{p}\hat{x} .$$

As with all operators, their algebraic forms are generally *dependent* on the representation basis (bases are described in some detail later). The only commutators that are *independent* of representation are constants, e.g., $[\hat{x},\hat{p}] = i\hbar$. It is mildly interesting to evaluate this explicitly in both the x (position) and p (momentum) bases:

$$x \text{ basis: } [\hat{x},\hat{p}] = x\frac{\hbar}{i}\frac{\partial}{\partial x} - \frac{\hbar}{i}\frac{\partial}{\partial x}x = \frac{\hbar}{i}\left(x\frac{\partial}{\partial x} - \mathbf{1}_{op} - x\frac{\partial}{\partial x}\right) = i\hbar, \qquad \textit{where} \quad \mathbf{1}_{op} \equiv \text{the identity operator} .$$

$$p \text{ basis: } [\hat{x},\hat{p}] = i\hbar\frac{\partial}{\partial p}p - p\left(i\hbar\frac{\partial}{\partial p}\right) = i\hbar\left(\mathbf{1}_{op} + p\frac{\partial}{\partial p} - p\frac{\partial}{\partial p}\right) = i\hbar .$$

We've noted that a commutator is an operator, so is there a physical meaning to the function $[\hat{x},\hat{p}]\psi(x)$? Let's see: the first term of the commutator is $\hat{x}\hat{p}$. Recall that $\hat{p}$ multiplies a function by the (local) momentum at each point, and produces a function of space which is *not* a quantum state. Similarly, $\hat{x}$ multiplies a function by the position 'x' everywhere, and also produces a function of space which is *not* a state. What does it mean then for $\hat{x}$ to act on $\hat{p}\psi(x)$, since $\hat{p}\psi(x)$ is not a state? In other words, what is the physical meaning of $\hat{x}\hat{p}$? Answer: nothing! This kind of commutator, taken as an operator, has no direct physical interpretation. It is simply an algebraic relationship that is used to derive extremely important results (such as the uncertainty principle).

Many commutators, taken as operators, have no direct physical interpretation;
they are simply algebraic relationships that are used to help derive physical results.

A second use for commutators relates to simultaneous eigenstates. The uncertainty principle says that some pairs of dynamic quantities cannot both have definite values. It is important to know which physical properties a quantum system can have with simultaneously definite values. A set of quantum numbers, one for each possible definite value, then fully specifies the state of the system, and therefore *all* its properties, definite or not. When two operators **commute**, it means they produce the same result acting in either order:

$$\hat{a}\hat{b} = \hat{b}\hat{a} \quad \Rightarrow \quad \left[\hat{a},\hat{b}\right] = \mathbf{0}_{op} \qquad \text{(the zero operator)} .$$

Note that the "zero" on the right hand side is an operator, not exactly a number. It means to multiply the function on which it acts by the number 0, which always returns the zero-valued function of space. Similarly, in a general vector space (say, spin states), it means to multiply the ket (i.e., vector) on which it acts by the number 0, which always returns the zero-vector, $\mathbf{0}_V$.

When two operators commute, they have a common set of eigenvectors, and this set is complete: it forms a basis for constructing any vector in the vector space.

Units of commutators: The units of a composition of operators is the product of the units of the constituent operators. Therefore, the units of each term of a commutator are the same: the product of the commutator's constituent operators. For example, the units of $\left[\hat{x},\hat{p}\right]$ are:

$$\left[\hat{x},\hat{p}\right]=\left[\hat{x}\hat{p}\right]=\text{m kg m/s}=\text{kg m}^2\text{/s} .$$

Commutators are linear: Commutators are linear in both arguments (where the arguments are themselves operators). Recall that linearity implies commuting with scalar multiplication, and distributivity over addition:

$$\left[k\hat{a},\hat{b}\right]\equiv k\hat{a}\hat{b}-\hat{b}k\hat{a}=k\left[\hat{a},\hat{b}\right], \qquad \left[\hat{a},j\hat{b}\right]\equiv\hat{a}j\hat{b}-j\hat{b}\hat{a}=j\left[\hat{a},\hat{b}\right] \qquad \text{(scalar multiplication),}$$

$$\left.\begin{aligned}&\left[\hat{a}+\hat{c},\hat{b}\right]\equiv\left(\hat{a}+\hat{c}\right)\hat{b}-\hat{b}\left(\hat{a}+\hat{c}\right)=\hat{a}\hat{b}-\hat{b}\hat{a}+\hat{c}\hat{b}-\hat{b}\hat{c}=\left[\hat{a},\hat{b}\right]+\left[\hat{c},\hat{b}\right]\\&\left[\hat{a},\hat{b}+\hat{d}\right]\equiv\hat{a}\left(\hat{b}+\hat{d}\right)-\left(\hat{b}+\hat{d}\right)\hat{a}=\hat{a}\hat{b}-\hat{b}\hat{a}+\hat{a}\hat{d}-\hat{d}\hat{a}=\left[\hat{a},\hat{b}\right]+\left[\hat{a},\hat{d}\right]\end{aligned}\right. \quad \text{(distributes)} \qquad (1.8)$$

There is also a composition rule (loosely, a "product rule") for commutators:

$$\left[\hat{a}\hat{c},\hat{b}\right]=\hat{a}\hat{c}\hat{b}-\hat{b}\hat{a}\hat{c}=\hat{a}\hat{c}\hat{b}\underbrace{-\hat{a}\hat{b}\hat{c}+\hat{a}\hat{b}\hat{c}}_{=0}-\hat{b}\hat{a}\hat{c}=\hat{a}\left[\hat{c},\hat{b}\right]+\left[\hat{a},\hat{b}\right]\hat{c}, \qquad \text{and similarly,}$$

$$\left[\hat{a},\hat{b}\hat{d}\right]=\hat{b}\left[\hat{a},\hat{d}\right]+\left[\hat{a},\hat{b}\right]\hat{d} .$$

You can remember the commutator composition rule as the sum of two terms, which you create as follows: "pull the left factor out to the left, and pull the right factor out to the right" (Figure 1.15).

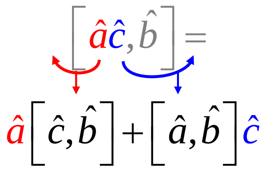


Figure 1.15 The two terms of the composition rule: pull the left factor out the left, and the right factor out the right.

1.10.5 External Fields

A quantum system may include "external fields," e.g. an externally applied magnetic fields whose value at all points is a given, and *not* part of the state vector $|\psi>$. Such fields are good approximations to the more precise quantized EM field. In such a case, the value of the field is embedded in the operators we define for observables of the system. For example, the energy operator of an electron in a given B-field would be:

$$\hat{H}_{magnetic}=-\boldsymbol{\mu}\bullet\mathbf{B}(\mathbf{r})=-\left(g_e\mu_B\frac{\hat{\mathbf{s}}}{\hbar}\right)\bullet\mathbf{B}(\mathbf{r})=-\left(-2\mu_B\frac{1}{2}\boldsymbol{\sigma}\right)\bullet\mathbf{B}(\mathbf{r})=+\mu_B\boldsymbol{\sigma}\bullet\mathbf{B}(\mathbf{r}), \qquad \text{using} \qquad g_e\approx-2 .$$

Similarly, the (canonical) momentum operator would include the given potential function $\mathbf{A}(\mathbf{r})$.

Including a "given" field in the operators of the system is in contrast to descriptions of a system where the field itself is quantized, and its value is part of the system state. Quantizing such fields involves Quantum Field Theory. Then, a function such as $\mathbf{B}(\mathbf{r})$ is replaced by the field operator, $\hat{\mathbf{B}}$,which supplies the value of the field from the quantized field state vector.

1.11 From Schrödinger to Newton: Deriving F = ma

We now show the quantum mechanical statement of Newton's 2nd law, $dp = F\,dt$ [which is still true relativistically], and which is (non-relativistically) equivalent to $F = ma$. Consider a 1D particle with definite momentum $p(t_0) = p_0$, at some time t_0. Its wave-function is $\psi(t_0, x) = \exp[(i/\hbar)p_0 x]$ (dropping the normalization factor, for simplicity).

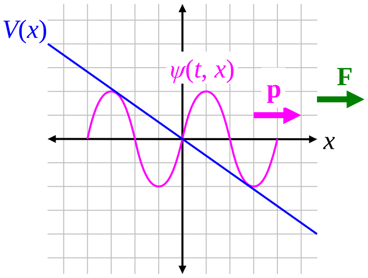


Figure 1.16 $\psi(t, x)$ is a 1D wave-function moving right. The force is also to the right.

Now apply a force $F = -dV(x)/dx$ to it for a time dt. From Newton's 2nd law, we expect that its new momentum and wave-function are:

$$p(t_0 + dt) = p_0 + F\,dt \qquad \Rightarrow \qquad \psi(t_0 + dt, x) = \exp\left[(i/\hbar)\left(p_0 + F\,dt\right)x\right]. \tag{1.9}$$

We now derive that result from quantum dynamics. Recall that the Schrödinger equation implies the time evolution operator, which tells how a wave-function evolves in time:

$$\psi(t_0 + dt) = \exp\left[-(i/\hbar)\hat{H}\,dt\right]\psi(t_0) = \exp\left[-(i/\hbar)\left(\frac{{p_0}^2}{2m} + V(x)\right)dt\right]\exp\left[(i/\hbar)\,p_0 x\right]$$

$$\textit{where} \qquad p_0 \equiv \text{initial momentum}\,.$$

The force is constant everywhere along the wave-function, so $V(x) = -Fx$. Rearranging the exponentials, and replacing $V(x)$ with $-Fx$, we get (to first order in dt):

$$\begin{aligned}\psi(t_0 + dt) &= \exp\left[-(i/\hbar)\frac{{p_0}^2}{2m}dt\right]\exp\left[+(i/\hbar)Fx\,dt\right]\exp\left[(i/\hbar)\,p_0 x\right]\\ &= \exp\left[-(i/\hbar)\frac{{p_0}^2}{2m}dt\right]\exp\left[(i/\hbar)\left(p_0 x + Fx\,dt\right)\right]\\ &= \exp\left[-(i/\hbar)\frac{{p_0}^2}{2m}dt\right]\exp\left[(i/\hbar)\left(p_0 + F\,dt\right)x\right].\end{aligned}$$

This is of the form expected in (1.9). The first factor is the usual complex rotation in time according to the energy, and does not affect the observable properties of the particle. The second factor is the wave function throughout space of momentum ($p_0 + F\,dt$). Thus quantum mechanics has proven the impulse law, $dp = F\,dt$. Then non-relativistically, as in classical mechanics:

$$F = \frac{dp}{dt} = m\frac{dv}{dt} = ma\,.$$

Thus, Schrödinger's equation, applied to a momentum eigenstate, reproduces Newton's 2nd law. Isn't it amazing that Newton could find this 300 years before quantum mechanics was discovered?

1.12 Measurement and Loss of Coherence

We now discuss the important phenomena of measurement and observation. In theoretical QM, we usually focus on perfect systems, and pure states. We frequently say that a measurement "collapses" the quantum state vector to one agreeing with the measurement, and this is often a useful simplification of the measurement process. However, in practice, the measurement process is more complicated than that, because most measuring equipment, and all observers, are macroscopic. The "decohered" state is the norm; you must work hard to achieve even an approximately pure entangled state. We show here that elementary QM can explain some of the features of real measurements, however, the full explanation of decoherence is beyond our scope. (The term "decoherence" has a specific meaning: the process of a system becoming entangled with its environment in irreversible ways, resulting in the loss of a consistent phase relationship between components of the system state. We therefore use the more general term "loss of coherence" for both decoherence and other processes.)

Most *macroscopic* measurements do not show quantum interference. Why not? One reason is that macroscopic bodies suffer unknowable, and unrepeatable energy interactions, i.e. they gain or lose an unknowable amount of energy due to uncontrollable interactions with their environments. In other words, they are subject to simple "noise." This results in the loss of a consistent phase relationship between components of a superposition state. We discuss below how such a loss of consistent phase leads to classical probabilities [e.g., Bay p26-7].

Walk-Through of a Real Measurement: Let us walk through a plausible measurement, and consider the elementary quantum mechanics involved. The system of Figure 1.17 is a macroscopic version of Figure 1.1, the quantum experiment that demonstrated the need for new concepts of measurement in quantum mechanics.

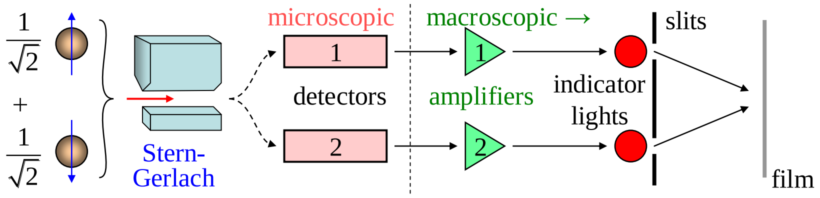


Figure 1.17 A typical measurement process: is the spin up or down?.

Suppose we start with a particle which can be in either of two states, $|s_1>$ or $|s_2>$, such as polarization (horizontal or vertical), or spin (up or down). A general particle state is then:

$$|\psi\rangle = a|s_1\rangle + b|s_2\rangle \qquad \textit{where} \quad a,b \equiv \text{complex coefficients}, |s_1\rangle, |s_2\rangle \equiv \text{basis states} .$$

This is called a **coherent superposition**, because a and b have definite phases. (This is in contrast to a **mixed state** or **incoherent mixture**, where a and b have unknown phases.) All that is required for loss of coherence is for the relative phases of a and b to become unknown. For simplicity, we take $|s_1>$ and $|s_2>$ to be energy eigenstates, and the particle is spread throughout our measurement system.

According to the Schrödinger equation, every state time-evolves with a complex phase, $\alpha(t)$, determined by its energy:

$$\alpha(t) = \alpha(0) + \frac{E}{\hbar} t \qquad \text{(phase accumulation of a system according to its energy)} .$$

Then our 2-state system time evolves according to:

$$|\psi(t)\rangle = a \exp(-iE_1 t/\hbar)|s_1\rangle + b \exp(-iE_2 t/\hbar)|s_2\rangle .$$

Since the energies E_1 and E_2 are quantized, the complex phases multiplying $|s_1>$ and $|s_2>$ maintain a precise (aka coherent) relationship, though the relative phase varies with time.

When we measure the particle state, the state of the measuring device becomes entangled with the measured particle. Let $|M_1>$ and $|M_2>$ be states of the whole measuring system in which either detector 1 detected the particle, or detector 2. If we look directly at the indicator lights, we will observe only state 1 or state 2, but never both. This means $|M_1>$ and $|M_2>$ are orthogonal. As the measuring system first detects the particle, the combined state of the particle/measuring-device starts out as a coherent superposition:

$$\left|\Psi\right\rangle = c\left|M_1\right\rangle\left|s_1\right\rangle + d\left|M_2\right\rangle\left|s_2\right\rangle \qquad \textit{where} \quad c, d \equiv \text{complex coefficients} .$$

The combined system time evolves according to its new energies:

$$\left|\Psi(t)\right\rangle = c\exp\left[-i\left(E_1 + E_{M1}\right)t/\hbar\right]\left|M_1\right\rangle\left|s_1\right\rangle + d\exp\left[-i\left(E_2 + E_{M2}\right)t/\hbar\right]\left|M_2\right\rangle\left|s_2\right\rangle .$$

If the energies of the two measuring device states fluctuate even a tiny bit, the two components of the superposition will rapidly wander into an unknown phase relation. They will lose coherence.

Every macroscopic system suffers unrepeatable and unknowable energy fluctuations due to its environment.

We estimate a typical coherence loss rate shortly.

Let us examine the effects of various kinds of energy transfers between a system and its environment. In our two-path experiment, the interference pattern is built up over many trials, by recording detections on film. Now suppose one path suffers an energy transfer to/from its environment before recombining and interfering. There are four possibilities:

1. The energy transfer is knowable and repeatable. Then one can predict and see an interference pattern in the usual way.
2. The energy transfer is unknowable, but repeatable. Then we can record an interference pattern, and from it, determine the relative phases of the two paths (mod 2π), and therefore the relative energies (mod $2\pi\hbar/t$) from $\Delta\alpha = \omega t = \left(\Delta E/\hbar\right)t$.
3. The energy transfer is knowable for each trial, but not repeatable. Essentially, each trial has its own position for the interference pattern. One can then divide the detection region into intervals of probability calculated for each trial, and then show consistency with QM predictions, but contrary to classical probability.
4. The energy transfer is unknowable and unrepeatable. Then there will be no interference pattern, and repeated trials do not allow us to measure any quantum effects, since the phase is unknown on each trial. Therefore, the measurements are equivalent to classical probabilities: it is as if a single path was chosen randomly, and we simply don't know which path it was.

This fourth condition, of unknowable and unrepeatable energy transfer, causes loss of coherence, the randomization of phase of components of a superposition. Loss of coherence makes measurements look like the system behaves according to classical probabilities, with no "wave" effects. Loss of coherence destroys the interference pattern when we try to measure through "which slit" a particle passes.

Full loss of coherence leads to classical probabilities.

Our example process leading to loss of coherence follows directly from the Schrödinger equation and unknown energy transfers. There is no need to invoke any "spooky" quantum effects.

Note that even accounting for loss of coherence, quantum theory *still* requires the axiom of collapse of the wave-function upon observation. When a particle's wave splits, then passes through both detector 1 and detector 2, and then loses coherence because of entanglement with a macroscopic measuring device, the system is still left in a superposition of both slits:

$$\left|\Psi\left(t_{after}\right)\right\rangle = f\left|M_1\right\rangle\left|s_1\right\rangle + g\left|M_2\right\rangle\left|s_2\right\rangle ;$$

we just don't know f or g. We can't generate an interference pattern from multiple trials, because each trial has a different phase relation between f and g, putting the peaks and valleys of any hoped-for interference pattern in a random place on each trial. These shifts average over many trials to a uniform distribution. Nonetheless, each trial evolves in time by the Schrödinger equation, which still leaves the system in a superposition. Once we "see" the result, however, the unobserved component of the wave-function disappears, i.e. the wave-function **collapses**.

Collapse of the wave-function is outside the scope of the Schrödinger equation, but *within* the scope of QM, because collapse is a *part* of QM theory. It is one of our axioms. Some references confuse this issue: they try to avoid assuming such a collapse as an axiom, but cannot derive it from other axioms. From this, they conclude that QM is "incomplete." In fact, what they have shown is that the axiom of collapse completes QM.

Note that once the measuring system *fully* loses coherence, we could just as well say that the wave-function has then collapsed, because from then on the system follows classical probabilities (equivalent to a collapsed, but unknown, wave-function). However, we now show that a binary model of "collapse or not" cannot explain partial coherence.

Partial coherence: What if we start with a microscopic system, such as that in Figure 1.1, but replace our microscopic atoms with *mesoscopic* things: bigger than microscopic, but smaller than macroscopic. Mesoscopic things might be a few hundred atoms. These are big enough to lose coherence much faster than single atoms, but still slowly enough that *some amount* of interference is observed. However, the interference pattern is weaker: the troughs are not as low, and the peaks are not as high. A superposition leading to a *weak* interference pattern is called **partially coherent**. We describe partial coherence in more detail in section 8.4. The simple model that the wave-function either collapsed or didn't cannot describe the phenomenon of partial coherence.

The larger the mesoscopic system, the more uncontrollable interactions it has with its environment, the faster it loses coherence, and the less visible is any resulting interference pattern. We can estimate the time-scale of coherence loss from our example energy fluctuations as follows: a single 10 μm infrared photon is often radiated at room temperature. It has an energy of ~ 0.1 eV = 1.6×10^{-20} J. This corresponds to $\omega = E/\hbar \sim 2\times10^{14}$ rad/s. When the phase of the resulting system has shifted by an unknowable amount > $\sim 2\pi$, we can say the system has completely lost coherence. At this ω, that takes $\sim 4\times10^{-14}$ s. In other words, thermal radiation of a single IR photon causes complete loss of coherence in about 40 femtoseconds. In practice, other effects cause macroscopic systems to lose coherence in dramatically shorter times.

Summary: A measurement entangles a measuring device with the measured system. The entangled state of device and system time evolves according to the SE. Macroscopic devices lose coherence, due to interactions with the environment. Lack of coherence prevents any interference pattern within the system. Therefore, measurement by a macroscopic device produces subsequent results that are classical, as if the system collapsed into a definite state upon measurement, but observers only "see" which state when they look at the measuring device. Any observation by a person is necessarily macroscopic, because people are big. Such an observation collapses the (incoherent) device/system/world state to that observed. Quantum interference can only be seen if it occurs *before* any entanglement with a macroscopic system (and therefore before any loss of coherence in the system).

The model of "collapse of the wave-function" is a binary concept: either the wave-function collapses or it doesn't. Such a model cannot account for the phenomenon of partial coherence. Loss of coherence is a continuous process, taking a fully coherent state through less and less partially coherent states and eventually to incoherent (aka "mixed") states. Continuous loss of coherence fully explains partial coherence and the varying visibility of interference patterns.

Some quantum effects, such as the spectrum of atoms, do not rely on interference, and are therefore macroscopically observable. In fact, measurement of such effects led to the development of QM.

1.13 Why Do Measurements Produce Eigenvalues?

This section is preparation for the next, which discusses wave-function collapse. Quantum Mechanics is consistent in this way:

> If you make a measurement, and then quickly repeat the same measurement,
> you will get the same result (quickly so the system doesn't "move" in between).

(Similarly, if you make a measurement, then wait and account for the system time evolution, and then make another measurement, your second result will be consistent with the first.) If you believe the postulates of QM, specifically that the wave function defines the spectrum (the set) of possible results, and their probabilities, then you must conclude that the wave function after a measurement must allow only a *single* result for a re-measurement (the result already measured).

Since the wave function before a measurement may be a superposition of possibilities, it must be that the act of measuring "collapses" the wave-function to a state that can produce only one result. Recall that a spatial wave-function (i.e., a wave-function of position in space) associates a local value for an observable with each point of the wave-function. Assume a wave function, $\psi(x)$, and an observable operator, $\hat{o}$. '$\hat{o}$' has some local value associated with each position in space; call it $o_{local}(x)$. After a measurement, the particle has a definite value of the observable, so it is reasonable that all points of the wave-function should contribute the same local value: that which was already measured. This means $o_{local}(x)$ is constant throughout space; call it o_{obs}. Then from the definition of local values, eq. (1.4), we have:

$$o_{local}(x)\psi(x) = o_{obs}\psi(x) = \hat{o}\psi(x)$$

The last equality requires that the wave function after a precise measurement be an eigenfunction of the observed operator, and the value measured must be its eigenvalue.

This phenomenon of a measurement causing a "collapse of the wave function" into an eigenstate is directly related to the uncertainty principle. The collapse is an inevitable part of taking a measurement; it derives from the consistency of measuring the same observable twice. It has nothing to do with clumsy photons, or other "probing" particles. Since non-commuting observables rarely have simultaneous eigenstates, a single wave function usually cannot have definite values of non-commuting observables. The lack of a common eigenfunction means that a definite value of one observable forces some uncertainty in the other. [See later section on Generalized Uncertainty for more details.]

Observable Operators Do *Not* Produce States Resulting From Measurements

The mathematical result of an observable operator acting on a state is very different from the state resulting from actually measuring that observable.

Many people confuse the result of an observable operator on a state with the act of measuring that observable. These are very different things!

The action of an observable operator on a state produces a mathematical result, a function of space (a "ket" in Dirac notation), which is one step in calculating the statistics of possible measurements of that observable on a system in that state. Such a result is *not* a quantum state. As we saw earlier, $\hat{x}\psi(x)$ yields a function of space (*not* a state) which is part of calculating possible values of the position of a particle in state $\psi(x)$, and their probabilities. This is quite different from the resulting *state* after an actual measurement of the particle's position. Recall the example $\psi(x) = 1/2$ between [0, 4], and 0 elsewhere (Figure 1.18, left).

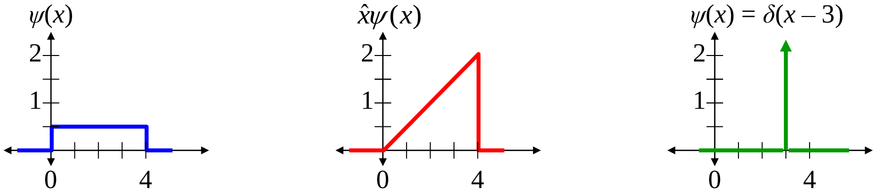


Figure 1.18 (Left) A normalized wave-function, $\psi(x)$. (Middle) $\hat{x}\psi(x)$. (Right) $\psi(x)$ after measuring the particle to be at position $x = 3$.

Figure 1.18, middle, is the function $\hat{x}\psi(x)$, which is a mathematical calculation. If we actually measure such a particle to be at position $x = 3$, the resulting state of the particle is (essentially) a delta function at $x = 3$:

$\psi(x) = \delta(x-3)$ (Figure 1.18, right).

(This is *not* square integrable, and so uses "delta-function normalization." More on this later.)

Note that the act of measurement is a *nonlinear* operation on the wave function; it can *not* be represented by a linear operator acting on the wave function. Recall that the whole point of a linear operator is to produce a superposition of *results* based on the superposition that composes the given function (or ket). In contrast, the consequence of a measurement is to *choose one specific state* out of a superposition of eigenstates.

A measurement *eliminates* a superposition, in favor of a more definite state.
Therefore, a measurement is *not* a linear operation on the state; it is inherently nonlinear.

We must be careful in interpreting this statement, because the term "superposition" is relative: it depends on the basis in which we write our function. A state of definite energy is a superposition of position states. A state of definite position is a superposition of energy states. *Any* spatial function is always a superposition in some bases, and a single component in other bases. However, when we take a measurement, one basis is special: the eigenbasis of the operator whose value we are measuring. If we measure energy, then the energy basis is special. If we measure position, then the position basis is special. A better statement of the effect of a measurement is then:

A measurement of a property eliminates a superposition of eigenstates of that property. The resulting ("collapsed") state is then an eigenstate of the operator corresponding to the measurement.

We now describe some ways of thinking about this effect.

1.14 The Collapse of Quantum Mechanics

Or, "Shut Up and Calculate." [Thanks to Andrew Cooksey for that sympathetic advice. This section is inspired by Sidney Coleman, Heinz R. Pagels, and all the other dedicated physicists trying to make Quantum Mechanics seem sensible.]

What is *really* going on inside the collapse of the wave-function? We must be careful with the word "really:" science is the ability to predict future observations based on past ones. Science uses **theories** (detailed quantitative models) to make these predictions. Theories usually include a *conceptual* model to help us remember the theory, but theories *always* include a mathematical model to compute predictions.

Any theory which computes predictions that agree with experiment is a "valid" theory.

Sometimes, more than one theory is valid. Because we are simple-minded, between two competing valid theories, we choose the simpler one. (This criterion is known as Occam's Razor.) We now compare two models, both of which agree with the following facts. These facts are verified countless times by experiment, and are undisputed by serious physicists:

1. If Alice takes a measurement, and now wants to predict future measurements, she *must* use, as her starting point, the quantum state which is consistent with her first measurement. No other quantum state will work for her. Calculations starting from any state other than that demanded by her first measurement will be *wrong*. Absolutely, completely, no "ifs, ands, or buts" wrong.

2. If Alice measures an observable for which the system is known *not* to be in an eigenstate, then her answer is determined by probability, and Alice cannot predict for certain the result of any single such measurement.

3. Alice prepares a quantum system in state A. She allows it to time-evolve into a superposition of two states, B or C (perhaps, (B) = decayed and (C) = not-decayed), but she does not measure the system to determine which state it is in. She waits further, and finally measures it in state D. To predict the probability of starting in A and ultimately measuring D, Alice *must* use as the intermediate state a coherent superposition of B and C, i.e. she must use the complex amplitudes of states B and C, to predict the measurement D. She *cannot use* classical probabilities.

Fact #1 says that quantum mechanics is self-consistent: no measurement will contradict another. It is why measurements always produce eigenvalues as their result (see "Why Do Measurements Produce Eigenvalues?" on page 48).

Fact #2 says that *from the point of view of an observer*, quantum mechanics is inescapably probabilistic.

Fact #3 says that probability for an observer *only comes into play when that observer actually observes something*. Until then, there is deterministic time evolution of the quantum state, including fully complex-valued amplitudes, and therefore (critically) *the possibility of quantum interference*. [However, as noted earlier, large systems inevitably suffer from decoherence: uncontrollable interactions with their environments.]

What is the "explanation" of these facts? There is no single, definitive explanation. There are many competing, valid theories. Here are two:

Dr. Xavier E. Rox believes in predictability. There can be no collapse, no random events. Physics is deterministic, just like the old classical physicist, Dr. Diehard, says. Dr. Rox observes, "Diehard's only problem is that his math is wrong. Physics follows the Schrödinger equation." Of course, to be consistent with experiment, Dr. Rox must assume that at every instant, the quantum state of the entire universe, including himself(!), splits into a new superposition of all possible results. "Better complexity and confusion than uncertainty," he declares, much to the dismay of Werner Heisenberg. On Sunday, Dr. Rox goes to the Church of Duplicity, and worships a rapidly growing list of very similar gods.

Dr. Ophelia C. Cam retorts, "Stuff and nonsense! I can only ever perceive *one* world, so it is unscientific to talk about others. They are, by definition, outside the possibility of observation, and therefore, also by definition, outside the realm of science." She believes that each observer, with each observation, collapses her own wave-function of the universe. That is, to be consistent with experiment, she must assume that each observer has her own wave-function for the universe, which collapses only when its owner makes an observation. This means the quantum state of the universe is *different for different observers*. How can a wave-function collapse? How can a wave-function be subjective, and not absolute? "I don't know, and I don't care," says Dr. Cam. "Like it or not, it is what it is. The *measured* results provide a single reality for all observers, so there is no physical consequence of personalized wave-functions." On Wednesdays, Dr. Cam goes to the Church of One Mind, where she prays to a very lonely God.

Who is right, then, Dr. Rox or Dr. Cam? This is not a scientific question, since both professors make the same experimental predictions. Whom you believe depends on which church you attend.

[There are many other conceptual models ("churches") of QM which agree with the facts, but since they have no observable consequences, we maintain that they have no scientific substance.]

2 Riding the Wave: More on Wave Mechanics

2.1 Units of Measure

Many QM references ignore the units of measure of the components of a calculation. For example, the most common element of QM calculations is the 1-dimensional *x*-representation wave-function, $\psi(x)$. What are its units? Answer: m$^{-1/2}$, or per-square-root-meters. Surprised? So were we.

In the following, we use square brackets to mean "the units of." E.g., [x] means "the units of x."

Let's start with the basics: in the macroscopic universe there are exactly four fundamental quantities: distance, mass, time, and charge. [One can reasonably argue for a fifth: angle.] In the MKSA system, the corresponding units are meters (m), kilograms (kg), seconds (s), and coulombs (C). We stick mostly with MKSA in this text. As is common, we use the terms "units" and "dimensions" interchangeably in this context.

For the units of $\psi(x)$, recall that the dot product of a normalized wave function with itself is a dimensionless 1:

$$\int_{-\infty}^{\infty} \psi^*(x)\psi(x)\,dx = 1 \qquad (\text{dimensionless})\,.$$

Since dx is in meters (m), and the units of ψ^* are the same as ψ, then $\psi^*\psi$ must be in m^{-1}, and thus ψ is in m$^{-1/2}$.

Equivalently, if x is in meters, and we compute the average of x:

$$\langle x \rangle = \int_{-\infty}^{\infty} \psi^*(x)\, x\, \psi(x)\,dx\,,$$

then the units of ψ must be m$^{-1/2}$.

What about the momentum representation, $a(p)$? The same normalization process starts with

$$\int_{-\infty}^{\infty} a^*(p)a(p)\,dp = 1 \quad (\text{dimensionless}) \qquad where \quad p \text{ is in } \frac{kg\cdot m}{s}\,.$$

Then $a(p)$ must be in $\left[\text{momentum}\right]^{-1/2} = \left[\dfrac{s}{kg\cdot m}\right]^{1/2}$, or s$^{1/2}$ kg$^{-1/2}$ m$^{-1/2}$, or "inverse square-root momentum."

Recall that mathematically, exponentials and logarithms are dimensionless, and their arguments must be dimensionless. Also, the unit "radian" is equivalent to dimensionless, because it is defined as arc-length/radius = m/m = dimensionless.

What about 3-dimensional wave functions? Given $\psi(x, y, z)$, its units are m$^{-3/2}$. Why? We refer again to the normalization integral, which says that the particle must be somewhere in the universe, i.e.

$$\Pr(\text{particle is somewhere in the universe}) = 1 = \iiint_{universe} \psi^*\psi\,dx\,dy\,dz \qquad (\text{dimensionless})\,.$$

The units of $dx\,dy\,dz$ are m^3, so ψ must be in m$^{-3/2}$. Often, for spherically symmetric potentials, ψ is a function of r, ϕ, and θ: $\psi(r, \theta, \phi)$. Then it must have units of m$^{-3/2}$ rad^{-1}:

$$\iiint_{universe} \psi^*\psi\, r^2 \sin\theta\,dr\,d\phi\,d\theta = 1 \qquad \text{and} \qquad r^2\,dr\,d\phi\,d\theta \text{ is in m}^3\text{-rad}^2\,.$$

However, since rad is dimensionless, this is the same as before: m$^{-3/2}$. Thus, as expected, the units of ψ are independent of the units of its arguments.

The unit of 2D $\psi(x, y)$ is left as an exercise for the reader.

Dimensions of Operators

Operators also have dimensions.

Let us consider the momentum operator. $\dfrac{d}{dx}$ is like dividing by x, so it has units of (1/m), or m^{-1}. Planck's constant h, or $\hbar \equiv h/2\pi$, is a quantum of action, (energy)(time), or of angular momentum (distance)(momentum); the units are thus joule-seconds (J-s), or in purely fundamental terms, kg-m^2/s. Then:

$$\hat{p} = \frac{\hbar}{i}\frac{d}{dx} \qquad \Rightarrow \qquad \text{units of } \left(\frac{kg \cdot m^2}{s}\right)\left(\frac{1}{m}\right) = \frac{kg \cdot m}{s}, \qquad \text{consistent with } p = mv.$$

The momentum operator has units of momentum.

In fact, all observable operators have the units of the observable. We can see this from the average value formula:

$$\langle \hat{o} \rangle = \int_{\infty} \psi^* \hat{o} \psi \, dx, \qquad \text{and} \qquad \left[\psi^* \psi \, dx\right] = \text{dimensionless} \qquad \Rightarrow \qquad [\hat{o}] = [o].$$

When composing operators, their units multiply. Thus we see that $\dfrac{d^2}{dx^2}$ has units of m^{-2}, etc.

Commutators are compositions of other operators, so the units of commutators are the composition of the units of the constituent operators. [More on commutators elsewhere.] Perhaps the most famous quantum commutator is:

$$[\hat{x}, \hat{p}] = \hat{x}\hat{p} - \hat{p}\hat{x} = i\hbar .$$

The units of $\hat{x}$ are m. (Note that the units of $\hat{x}\psi(x)$ = m(m$^{-1/2}$) = m$^{1/2}$, *not* m.) The units of $\hat{p}$ are kg-m/s. The units of $\hat{x}\hat{p}$ are simply the product of the units of $\hat{x}$ and $\hat{p}$: (m)(kg-m/s) $= \dfrac{kg \cdot m^2}{s}$. This must be, because the commutator in this case works out to a constant, $i\hbar$, with those units.

Note that the units of operators don't change with the representation basis. For example, $\hat{x}$ in the momentum representation is still meters:

$$\hat{x} = i\hbar \frac{d}{dp} \qquad \Rightarrow \qquad \text{units of} \left[\frac{\;\dfrac{\text{kg m}^2}{\text{s}}\;}{\dfrac{\text{kg m}}{\text{s}}}\right] = \text{m} .$$

2.2 The Dirac Delta Function

The Dirac delta function is used heavily all over physics, engineering, and mathematics. Thoroughly understanding it is essential for anyone in those fields. Read more about the δ-function in *Funky Mathematical Physics Concepts* (http://physics.ucsd.edu/~emichels/FunkyMathPhysics.pdf). The δ-function is also called an "impulse" or "impulse function."

The Dirac delta function is really a pseudo-function: it implies taking a limit, but without the bother of writing "$\lim\limits_{\Delta x \to 0}$" all the time. The Dirac delta function is often formally defined as the limit of a Gaussian curve of (a) infinitesimal width, (b) unit integral $\left(\int_{-\infty}^{\infty} \delta(x)\, dx = 1\right)$, and thus (c) infinite height. This is somewhat overkill for our purposes, and it may be simpler to think of the delta function as a *rectangular* pulse of (a) infinitesimal width, (b) unit area, and thus (c) infinite height, located at zero (Figure 2.1, left):

Figure 2.1 (Left) The δ-function can be written as one-sided in most cases. However, (right) it is usually considered even. In any case, we take the limit as $\Delta x \to 0$.

Mathematically, we could write this *simplified* (asymmetric) delta function as

$$\text{simplified } \delta(x) = \lim_{\Delta x \to 0} f(x) \qquad \textit{where} \quad \begin{array}{ll} f(x) = 0, & x < 0 \\ f(x) = 1/\Delta x, & 0 \le x \le \Delta x \\ f(x) = 0, & x > \Delta x\,. \end{array}$$

Though the above works for any well-behaved (i.e., continuous) function, the delta function is usually considered an even function (symmetric about 0), so it is sometimes better to write (Figure 2.1, right):

$$\text{simplified } \delta(x) = \lim_{\Delta x \to 0} f(x) \qquad \textit{where} \quad \begin{array}{ll} f(x) = 0, & x < -\Delta x/2 \\ f(x) = 1/\Delta x, & -\Delta x/2 \le x \le \Delta x/2 \\ f(x) = 0, & x > \Delta x/2\,. \end{array}$$

However, in spherical polar coordinates, the radial delta function at zero requires the *asymmetric* form, and cannot use the symmetric form (see *Funky Mathematical Physics Concepts*).

Both of the above simplified versions of the delta function require special handling for more advanced applications where we need to take derivatives of $\delta(x)$; we will not use such derivatives in this book.

Units of the delta function: Another surprise: the δ-function is *not* dimensionless.

The Dirac delta-function has units!

Usually, such mathematically abstract functions are dimensionless, but the key property of the delta function is that its *area* is 1 and dimensionless. This means:

$$\int_{-\infty}^{\infty} \delta(x)\, dx = 1 \quad \text{(dimensionless)}\,.$$

So if x (and thus dx) is in m, $\delta(x)$ must be in m^{-1}. But we use the delta function for all sorts of measures, not just meters: radians, momentum, etc. So by definition, the delta function assumes units of the inverse of its argument. Given a radian, the units of $\delta(\theta)$ are inverse radians (rad^{-1}, equivalent to dimensionless); given a momentum, the units of $\delta(p)$ are $\dfrac{s}{kg \cdot m}$; and so on. Also, $\delta^{(3)}(\mathbf{r})$ has units of the inverse cube of the units of $\mathbf{r}$ ($\mathbf{r}$ in m $\Rightarrow \delta^{(3)}(\mathbf{r})$ in m^{-3}), and $\delta^{(4)}(x^{\mu})$ has units of the inverse 4th power of the units of x^{μ}.

An important consequence of the definition of $\delta(x)$ is that, because $\delta(x) = 0$ except near $x = 0$,

$$\int_{-\varepsilon}^{\varepsilon} \delta(x)\, dx = 1, \qquad \forall \varepsilon > 0\,.$$

Note that $\delta(x)$ is *not* square integrable, because

$$\int_{-\infty}^{\infty} \delta^2(x)\, dx = \lim_{\Delta x \to 0} \int_0^{\Delta x} \left(\frac{1}{\Delta x}\right)^2 dx = \lim_{\Delta x \to 0} \left[\frac{x}{(\Delta x)^2}\right]_0^{\Delta x} = \lim_{\Delta x \to 0} \frac{1}{\Delta x} \to \infty\,.$$

Interestingly, though the delta function is often given as a Gaussian curve, the precise form doesn't matter, so long as it is analytic (i.e., infinitely differentiable, or has a Taylor series), unit integral, and infinitely narrow [P&C p479b]. Other valid forms are:

$$\delta(x) = \lim_{\lambda \to 0} \frac{1}{\pi} \frac{\lambda}{\lambda^2 + x^2}, \qquad \delta(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} dk .$$

This latter form is extremely important in QM in the momentum representation, and in quantum field theory.

Integrals of δ-Functions of Functions

When changing variables, we sometimes need to know what is $\int dx\, \delta\left(f(x) \right)$?

Let $\quad u = f(x), \qquad du = f'(x)\, dx, \qquad$ and define x_0 $s.t.$ $f(x_0) = 0$.

Then $$\int_{-\varepsilon}^{\varepsilon} dx\, \delta\left(f(x) \right) = \int_{-\varepsilon'}^{\varepsilon'} du \frac{1}{\left| f'(x_0) \right|} \delta\left(u \right) = \frac{1}{\left| f'(x) \right|} .$$

We must take the magnitude of the derivative, because $\delta(x)$ is always positive, and always has a positive integral. The magnitude of the derivative scales the area under the delta function.

If the interval of integration covers multiple zeros of $f(x)$, then each zero contributes to the integral:

Let $\quad x_i =$ zeros of f, i.e. $f(x_i) = 0, \qquad i = 1,..n$.

Then $$\int_{-\varepsilon}^{\varepsilon} dx\, \delta\left(f(x) \right) = \sum_{i=1}^{n} \frac{1}{\left| f'(x_i) \right|} .$$

3D δ-function in Various Coordinates

See *Funky Mathematical Physics Concepts* for a more complete description, but note that $\delta^3(\mathbf{r})$ has a simple form *only* in rectangular coordinates:

$$\delta^3(x, y, z) = \delta(x)\delta(y)\delta(z),$$

but:

$$\delta^3(r, \theta, \phi) \ne \delta(r)\delta(\theta)\delta(\phi) . \qquad \text{(It's more complicated than this.)}$$

2.3 Dirac Notation

Dirac notation is a way to write the **algebra** of QM bras, kets, and operators. It is widely used, and essential to all current quantum mechanics. It applies to both wave-mechanics, and discrete-state mechanics (discussed in a later chapter).

You are familiar with the ordinary algebra of arithmetic. You may be familiar with Boolean algebra. There are also algebras of modular arithmetic, finite fields, matrix algebra, vector spaces, and many others. All algebras are similar to arithmetic algebra in some ways, but each is also unique in some ways. In general, an **algebra** is a set of rules for manipulating symbols, to facilitate some kind of calculations. We here describe Dirac notation and its associated Dirac algebra. Included in Dirac algebra is the algebra of operators (covered in a later section). Dirac algebra also brings us closer to the concept of kets and bras as vectors in a vector space (see page 63).

2.3.1 Kets and Bras

For wave mechanics, **kets** and **bras** are complex-valued functions of space (spatial functions), such as quantum states, and the results of operators on states. In Dirac notation, kets are written as |*name*>, where 'name' identifies the ket. The ket is a short-hand for the spatial wave function, say $\psi(\mathbf{r})$. The 'name' is arbitrary, much like the choice of letters for variables in equations. However, there are some common conventions for choosing ket names, again similar to the conventions for using letters in equations. In this section, we discuss only the spatial kets.

As a ket example, suppose we have a 1D spatial wave function, $\psi(x)$. Since any wave function can be written as a ket, we might write the ket for $\psi(x)$ as $|\psi>$ (assuming some notational license for now):

$$|\psi> \; \equiv \; \psi(x) = \text{complex-valued function of } x\ .$$

Note that the ket $|\psi>$ stands for the *whole* wave function; it does *not* represent the value of the wave function at any particular point. One of the key benefits of Dirac notation is that *kets, bras, and operators are independent of any representation basis*. Since they always represent the *entire* spatial function, there's no question of "what is the basis for a ket?" More on representations (decomposition in different bases) later.

Some might object to equating a ket to a function, as we did above: $|\psi> = \psi(x)$. More specifically, $\psi(x)$ is a particular representation of the quantum state $|\psi>$, so it would perhaps be more explicit to say "$|\psi>$ can be represented as $\psi(x)$," but that seems pedantic. We all agree that "5 = 4 + 1," yet the symbol "5" is different than the symbol "4 + 1." They are two representations of the same mathematical quantity, 5. Furthermore, since any function of position, say $\psi_x(x)$, can be written as a function of momentum, $\psi_p(p)$, our flexible notation would say that $\psi_x(x) = \psi_p(p)$, which is OK with us. This simply means that $\psi_x(x)$ and $\psi_p(p)$ both represent the same mathematical entity. I am therefore content to say:

$$\left|\psi\right\rangle = \psi_x(x) = \psi_p(p) = \text{any other representation of the ket } \left|\psi\right\rangle .$$

Dual to kets are **bras**. Bras are written as <*name*|, where '*name*' identifies the bra. Bras are also a short-hand for complex-valued functions of space. The same function of space can be expressed as either a ket or a bra. The difference is the ket is short-hand for the spatial function itself; the bra is shorthand for the complex conjugate of the function. Thus (continuing our flexible notation):

$$\left\langle\psi\right| = \psi^*(x) \qquad (\text{complex conjugate})\ .$$

For example, suppose we have two wave-functions over all space, $\psi(x)$ (in one dimension), and $\varphi(x)$. (The generalization to higher dimensions is straightforward.) It is frequently useful to determine the **dot product** of two wave functions, which is a single complex number, defined as:

$$\psi \cdot \varphi = \int_{-\infty}^{\infty} \psi^*(x)\varphi(x)\, dx \qquad \text{(a complex number)}\ .$$

Notice that the first wave function, ψ, is conjugated. Now the bra representation of ψ^* is just $<\psi|$, and the ket representation of φ is $|\varphi>$. When written next to each other, bra-ket combinations are defined as the dot product integral, i.e.

$$\left\langle\psi\middle|\varphi\right\rangle \equiv \psi\cdot\varphi = \int_{-\infty}^{\infty} \psi^*(x)\varphi(x)\, dx\ . \qquad \text{(A bra-ket combination is a bra-c-ket, < >. Get it?)}$$

When writing a bra-ket combination, use only one vertical bar between them: $<\psi|\varphi>$, *not* $<\psi||\varphi>$.

As a related example, using our new Dirac shorthand, we can write the "squared-magnitude" (sometime called "squared-length") of ψ as the dot product of ψ with itself:

$$magnitude^2 = \int_{-\infty}^{\infty} \psi^*(x)\psi(x)\, dx = \left\langle\psi\middle|\psi\right\rangle .$$

Summary of Kets and Bras

The ket shorthand for $\psi(x)$ is $|\psi>$. The bra shorthand for $\psi^*(x)$ is $<\psi|$.
Combining a bra with a ket, $<\psi|\varphi>$, invokes the dot product operation.

Quite simply, a **ket** is a function of space; a **bra** is the complex conjugate of such a function. A bra-ket is the dot product of the bra and ket, yielding a complex number. Recall that the QM dot product is *not* commutative (discussed elsewhere):

$$\langle\psi|\varphi\rangle \equiv \int_{-\infty}^{\infty} \psi^*(x)\varphi(x)\,dx = \langle\varphi|\psi\rangle^* \qquad \text{(reversing the operands conjugates the dot product).}$$

We've seen that kets and bras can be wave functions which are quantum states, but as noted earlier, kets and bras are more general than that. A ket or a bra can be either a quantum state, or the result of *operations* on a state. In other words, a ket or bra can be most any function of space. (Recall that a **quantum state** defines everything there is to know about a particle, including probabilities of finding it anywhere in space. A particle spatial quantum state (i.e., excluding its spin part), can be expressed as a complex-valued function of position, say $\psi(x, y, z)$.) Therefore:

All states are kets, but not all kets are states.

For example, a particle can be in a *state* $|\psi\rangle$, but no particle state can be given by the *ket* $\hat{p}|\psi\rangle$.

A note about spin: wave-functions alone may not fully define a quantum state, because they do not define the **spin** of a particle, i.e. its intrinsic angular momentum. Therefore, a full quantum state, for a particle with spin, is a combination of the wave-function (spatial state) and its **spin-state**. More on this later.

2.3.2 Operators in Dirac Notation

I hope this manuscript draft excerpt will encourage you to buy the final book:
Locally at the UCSD bookstore: Quirky Quantum Concepts

From Springer: Quirky Quantum Concepts

From Amazon: Quirky Quantum Concepts

2.3.3 Units Revisited

You can almost think of kets and bras as wave-functions (but kets and bras are independent of any representation, such as position or momentum representations). We can therefore apply our knowledge of units of wave-function to Dirac notation. We cannot ascribe any units to a ket or a bra by itself, since those units depend on the representation. However, we've seen that the squared-magnitude is dimensionless in any representation, so for a quantum state:

$$\langle\psi|\psi\rangle \quad \text{is dimensionless}.$$

We've also seen that the dimensions of operators are the same in any representation, so operators in Dirac notation have the same units as operators in "function notation:"

$$\hat{x} \qquad \text{has units of m,}$$

$$\langle\psi|\hat{x}|\psi\rangle \text{ has units of m,}$$

$$\langle\psi|\hat{p}|\psi\rangle \text{ has units of kg-m/s, etc.}$$

2.3.4 And Now, the Confusing Part

There is a subtle and confusing ambiguity in common Quantum Mechanics notation. The symbol '$\psi(x)$' has two different meanings, in different contexts. First, $\psi(x)$ means a function of space, the whole function. Second, sometimes $\psi(x)$ means the value of ψ at some specific point x, i.e., one specific point of ψ, and *not* the whole function. You can only tell from the context which meaning of the symbol '$\psi(x)$' is intended.

An example of the first meaning is (in our flexible notation):

$\psi(x) = |\psi>$, '$\psi(x)$' means the whole function. A ket is *always* a whole function. (2.1)

Another way to write "the whole function" is just "ψ"; thus, the following are equivalent:

"A particle is in a state, ψ."	(This does not imply any particular basis of ψ.)
"A particle is in a state, $\|\psi>$."	(Nor does this.)
"A particle is in a state, $\psi(x)$."	(This emphasizes the position representation of ψ.)

An example of the second meaning is

$\psi(x) = <x|\psi>$, '$\psi(x)$' means the value at some point x. (2-2)

A dot product is a single complex number, and $<x|\psi>$ is the value of ψ at the point x.

'$<x|$' is a *variable* bra: for each value of x there is a different bra (a different wave-function). Thus, $<x|$ is one of a set of functions of space. It is variable because there are *different spatial functions* for each value of x. '$<x|$' is the (conjugate of the) wave-function of a particle completely localized to the position x, specifically, an impulse (Dirac δ-function) at some position 'x'. Given x, we can write the wave function for $<x|$, call it $\eta(a)$, where we use 'a' as the position variable (since 'x' is already taken), see Figure 2.2.

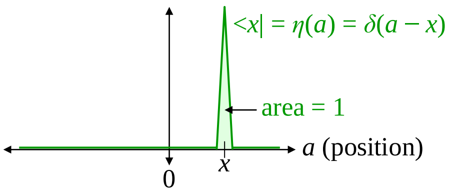


Figure 2.2 The bra $<x|$ for one particular value of x.

Since $|x>$ is real, the conjugation of the bra has no effect:

$<x| \equiv \eta(a) = \delta(a-x)$ for all a, $-\infty < a < \infty$, where 'a' is the position variable (not x).

Then:

$$\langle x|\psi\rangle = \int_{-\infty}^{\infty} da\ \delta(a-x)\psi(a) = \psi(x), \qquad \text{the value of } \psi \text{ at position } x \quad \text{[Sak 1.7.5 p52m].}$$

Recall that integrating through an impulse function times some function "picks out" that function at the impulse position; in this case, we pick out ψ at x, or $\psi(x)$.

Variable kets and bras are often written with x' as the position, so that x (instead of a) is the usual coordinate axis:

$$|x'\rangle = \delta(x-x') \qquad \text{and} \qquad \langle x'| = \delta(x-x') .$$

In other coordinates, there are other variable bras. In two dimensions, $<x, y|$ is a wave-function localized to a point in the plane:

$$<x, y| \equiv \eta(a, b) = \delta(a-x)\ \delta(b-y) .$$

In polar coordinates, $<r, \phi|$ is a wave-function localized to a particular angle. Similarly, $<\theta, \phi|$ is a direction in 3D space, though the normalization is tricky in non-rectangular coordinates. In particular, for $\theta > 0$:

$$\langle\theta,\phi| \equiv \eta(c,d) = \frac{1}{\sin\theta}\delta\left(c-\theta\right)\delta(d-\phi) \ne \delta\left(c-\theta\right)\delta(d-\phi), \qquad \theta > 0 .$$

(Some references forget the (1/sin θ) factor.) See *Funky Mathematical Physics* for more on multi-dimensional δ-functions in non-rectangular coordinates. (We address more issues of normalization on page 60.) This concept of multi-dimensional localized states extends directly to multi-particle states.

2.4 Covering Your Bases: Decomposition

Any wave function can be written as the (possibly infinite) sum of **basis functions:** defined as a set of functions which can be linearly combined to construct any function in the system of interest (i.e., in the vector-space). This is a common concept in linear system analysis throughout engineering and physics. For example, any (reasonably well-behaved) function can be written as the sum of an infinity of sine waves (a Fourier series or transform). The sine waves are the basis functions, which when weighted with coefficients, sum to compose the original function. In QM, Fourier *transforms* come up for converting between the position-representation and the momentum representation of a wave-function, which is a special case of the more-general concept of changing bases. (Fourier *series* don't come up much in QM.)

2.4.1 Countably Infinite Basis Sets

I hope this manuscript draft excerpt will encourage you to buy the final book:
Locally at the UCSD bookstore: Quirky Quantum Concepts

From Springer: Quirky Quantum Concepts

From Amazon: Quirky Quantum Concepts

2.4.2 A Funny Operator

Any bra or ket (i.e., wave-function or function of space) can be decomposed into basis functions (or basis kets):

$$\left|\psi\right\rangle = \sum_n c_n \left|b_n\right\rangle = \sum_n \left\langle b_n \middle| \psi \right\rangle \left|b_n\right\rangle, \qquad \textit{where} \quad \left|\psi\right\rangle \text{ is a ket, } c_n = \left\langle b_n \middle| \psi\right\rangle, \text{ and } \left|b_n\right\rangle \equiv \text{basis kets} .$$

We can rewrite the last expression by putting the coefficient of each basis function on the right, and applying parentheses:

$$\left|\psi\right\rangle = \sum_n \left|b_n\right\rangle\left\langle b_n \middle| \psi\right\rangle = \left(\sum_n \left|b_n\right\rangle\left\langle b_n\right| \right)\left|\psi\right\rangle .$$

This is simply decomposing a vector into its components, and then summing all the components, which returns the original vector. Therefore, the parenthesized operator is the identity operator:

$$\sum_n \left|b_n\right\rangle\left\langle b_n\right| = \mathbf{1}_{op} \qquad \text{(the identity operator, aka the } \textbf{completeness operator}\text{)} .$$

In other words, we can insert a sum over a set of arbitrary basis functions, $\sum_n \left|b_n\right\rangle\left\langle b_n\right|$, before any ket. That might seem pointless, but it's actually an important tool of Dirac algebra. It works just as well acting to the left on a bra:

$$\left\langle\psi\right| = \sum_n \left\langle \psi \middle| b_n\right\rangle\left\langle b_n\right| = \left\langle\psi\right| \left(\sum_n \left|b_n\right\rangle\left\langle b_n\right| \right) .$$

This is also evident because the identity operator is hermitian. The completeness operator is dimensionless.

This "completeness operator" can also be considered a special case of the prior section on operators "mentally decomposing" a vector into the operator's eigenvectors, multiplying those components by their eigenvalues, and summing the results. Here, the operator is the identity operator; any basis is an eigenbasis of the identity operator; and the eigenvalues are all 1.

This idea also generalizes to continuous bases in a straightforward way: the sum becomes an integral. For example:

$$\int_{-\infty}^{\infty} dx\, |x\rangle\langle x| = \mathbf{1}_{op}\ .$$

2.4.3 Countable or Uncountable Doesn't Mean Much

You may be worried that a countably infinite basis set (say, the energy basis) can compose a function of space, with an uncountably infinite position basis. (If this doesn't bother you, skip this section.) But uncountable infinities are "bigger" than countable ones, so what is the dimensionality of our vector space?

Let's consider states of a simple harmonic oscillator, in both the energy and position bases. In the energy basis, the basis functions are the eigenstates of energy (often labeled $|u_0>$, $|u_1>$, $|u_2>$, ...). Any bound-state wave function can be composed from the *countably* infinite set of energy eigenstates. In the position representation (basis), the bases (basis functions) are Dirac deltas at all values of x (position). Any wave function can be composed from the *uncountably* infinite set of position impulses. In short, the $|u_n>$ are countably infinite, the $|x>$ is uncountably infinite, yet both sets are (complete) basis sets for bound states. In fact, the bound states are a subset of the full vector space spanned by $|x>$, which explains why the dimension of the function space spanned by $|x>$ is "bigger" than the bound state space. However, any bound state function can be represented as the sum of energy bases, or the sum of position bases. Thus, the degree of infinity of the dimension of a basis is not always significant.

Finally, consider the energy states of a hydrogen atom: for bound states the electron energies are below the "lip" of the well [usually the lip is set to $E = 0$, so this means for negative electron energies]. There are a countably infinite set of such energy eigenstates ($E_n = Ry/n^2$). But for ionized hydrogen, where the electron is unbound and can be anywhere in space, there are an uncountably infinite set of energy eigenstates. Thus, the complete basis of energy eigenstates for a (possibly ionized) hydrogen atom is the union of a countably infinite set and an uncountably infinite set (Figure 2.3). C'est la vie.

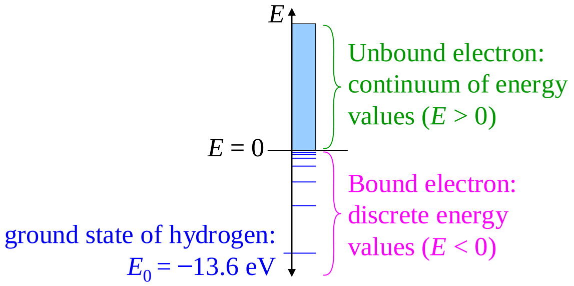


Figure 2.3 Allowed electron energies near a proton.

2.5 Normalize, Schmormalize

We normalize quantum states to make it easier to compute quantities such as probabilities, averages, cross sections, etc. Usually, we normalize our basis functions so that the probability of a given basis in a superposition is given by the squared-magnitude of that basis function's coefficient:

$$\text{Given} \quad |\psi\rangle = c_1|\phi_1\rangle ... + c_j|\phi_j\rangle + ... \qquad \text{then} \qquad \Pr(\text{measuring}\,|\phi_j\rangle) = |c_j|^2\ .$$

But this guideline has different implications for different circumstances, and sometimes other criteria are chosen, so there are at least 6 different normalization methods in common use. We summarize them first, and then describe most of them in more detail. All normalization methods apply to any number of spatial dimensions (not vector-space dimensions); we show some examples as 1D, but they apply just as well to 2D and 3D. Recall that $|\psi|^2 = \psi^*\psi$. Common normalizations:

1. Bound states usually use wave-function normalization:

$$\langle\psi|\psi\rangle = \int_{-\infty}^{\infty} \psi^*(x)\psi(x)\,dx = 1, \qquad \text{or in 3D:} \quad \langle\psi|\psi\rangle = \int_{\infty} \psi^*(\mathbf{r})\psi(\mathbf{r})\,d^3r = 1 .$$

2. Unbound states (e.g., free particle states) usually use delta-function normalization:

$$\langle x|x'\rangle = \int_{-\infty}^{\infty} \psi_x^*(a)\psi_{x'}(a)\,da = \delta(x-x'), \qquad \text{or} \qquad \langle\mathbf{r}|\mathbf{r}'\rangle = \int_{\infty} \psi_{\mathbf{r}}^*(\mathbf{a})\psi_{\mathbf{r}'}(\mathbf{a})\,d^3a = \delta^3(\mathbf{r}-\mathbf{r}') .$$

3. Box normalization: $\langle\psi|\psi\rangle = \int_{Length} \psi^*(x)\psi(x)\,dx = 1, \quad \text{or} \qquad \int_{Volume} \psi^*(\mathbf{r})\psi(\mathbf{r})\,d^3r = 1 .$

4. Radial wave-functions, and the like: $|R(r)|^2$ is the radial part of the 3D probability density function pdf$_{volume}(\mathbf{r})$; it is *not* the 1D pdf$_{radius}(r)$. The pdf$_{radius}(r) = r^2|R(r)|^2$. Therefore:

$$\int_0^{\infty} r^2 R^*(r)R(r)\,dr = 1 \qquad \left[\text{Gri 4.31 p138}\right] .$$

5. For perturbation theory, with orthonormal basis state $<n|n> = 1$, we normalize the perturbed state $|n'>$ such that $<n|n'> = 1$, which means $|n'\rangle = |n\rangle + \sum_{j\neq n} c_j\,|j\rangle$ [Bay 11-17 p227b] .

6. In quantum field theory, energy normalization is often used, because it is Lorentz covariant. E.g.:

$$\int_{Volume} \bar{\psi}\psi\,d^3r = 2E \qquad \text{[from P\&S 2.36 p23m]} .$$

We now discuss some of these further.

2.5.1 Wave-function normalization

I hope this manuscript draft excerpt will encourage you to buy the final book:
Locally at the UCSD bookstore: Quirky Quantum Concepts

From Springer: Quirky Quantum Concepts

From Amazon: Quirky Quantum Concepts

2.5.2 Delta-function normalization

(excerpted)

2.5.3 Box normalization

(excerpted)

2.5.4 Funny Normalization of Radial Wave Functions (and the Like)

(excerpted)

2.6 Adjoints

The "adjoint" of an operator is another operator. Adjoints are an essential part of Dirac algebra, being used in proofs, and allowing the crucial concept of a **self-adjoint operator**: an operator which is its own adjoint (aka "hermitian" operator). All quantum operators corresponding to observable properties of systems are self-adjoint. Self-adjoint operators have real eigenvalues, which are required since observable properties are real. If you are comfortable with matrices, you may want to preview the section on finite dimensional adjoints in the Matrix Mechanics chapter, since adjoints are easier to understand in finite dimensions.

We usually think of operators as acting to the right on a ket. However, operators can also be thought of as acting to the left on a bra. [In matrix mechanics, these correspond to (respectively) pre-multiplying a vector (ket) by a matrix, and post-multiplying a row vector (bra) by a matrix.] Recall that, when acting to the right, an operator produces a new ket from a given ket. Since we can alternatively think of operators as acting to the left, we can also say that an operator produces a new bra from a given bra. However, in general, the effect of an operator acting to the left on a bra is different than the effect of the operator acting to the right on a ket. We now show that we can define the left action of an operator to make inner products associative.

I hope this manuscript draft excerpt will encourage you to buy the final book:
Locally at the UCSD bookstore: Quirky Quantum Concepts

From Springer: Quirky Quantum Concepts

From Amazon: Quirky Quantum Concepts

2.6.1 Adjoint Summary

We've shown that we can define a left-action of an operator, in terms of its right-action, as that required to preserve the value of all inner products regardless of which way we think of the operator acting. In other words, we make inner products $<a|\hat{o}|b>$ associative. The left action then defines the "adjoint" operator, which acts to the right the way the original operator acts to the left. With respect to adjoints, there are two special classes of operators: "hermitian" operators are self-adjoint: they are their own adjoint operator. "Unitary" operators have an adjoint which is also the operator inverse. We have much to say about these two special classes of operators throughout the rest of this book.

2.7 The WKB Approximation

Since few quantum mechanics bound states can be computed exactly, approximation methods are very important for computing practical quantities. In particular, knowing the quantized energy levels for the bound states of a given potential is quite important, because it gives us the radiation spectrum. For bound states with moderate to high energies, the WKB approximation provides a good estimate of quantized energy levels [Sch p277]. (WKB can also be used to estimate tunneling probabilities, but we do not address that here [Sch p278].) WKB is named for Gregor Wentzel, Hans Kramers, and Léon Brillouin, three of many people who described the method early on. WKB goes by other acronyms, as well.

The WKB approximation can be used for any 1D problem, including a separated variable from a 2D or 3D problem, such as the radial equation from a central potential. The energy spectrum is our concern here; we do not focus on approximating the wave-function itself. WKB also introduces the semi-classical approximation for momentum, which is used in many situations.

For WKB to apply, the potential energy function which binds the particle must be either smooth near the classical turning points, or essentially infinite at a hard edge (more later). We take the following approach:

1. Review of qualitative consideration of high energy wave functions.
2. Understand the wave-function cycles in the bowl of a bound state.
3. Understand the wave-function near the edges (turning points).
4. Match the bowl solution to the edge solutions, to arrive at the quantization condition.
5. From the quantization condition, we determine the quantized energies for integer n, $E(n)$.
6. Examine the validity conditions: when is WKB a valid approximation?

I hope this manuscript draft excerpt will encourage you to buy the final book:
Locally at the UCSD bookstore: Quirky Quantum Concepts

From Springer: Quirky Quantum Concepts

From Amazon: Quirky Quantum Concepts

2.8 Long-Distance Operators

This section attempts to clarify some subtleties of spatial operators.

The position and momentum bases (basis functions) are parameterized by continuous observables: position and momentum. Between any two eigenvalues (position or momentum) there are an infinite number of basis functions and eigenvalues, i.e. there is a continuum of basis functions. (In fact, there are an uncountably infinite set of basis functions.) *Observable* operators in continuous bases are simple multiplier or differential operators; e.g., in the position basis:

$$\hat{p} = \frac{\hbar}{i}\frac{\partial}{\partial x}, \qquad \hat{p}^2 \equiv \hat{p}\hat{p} = -\hbar^2 \frac{\partial^2}{\partial x^2}, \qquad \hat{x} = x \qquad \text{(position basis)}.$$

For these simple operators, at any given point, the only contribution to the value of the operation on a spatial function comes from that point, or at most, a differentially small neighborhood of that point. Distant points do not contribute, and therefore distant places in the spatial function do not interact.

Mathematical operators (as opposed to observable operators) do not necessarily have such simple differential forms. We now consider some important operators: space translation, rotation, time evolution, and parity. Later, we will use the concepts of rotations, angular momentum, and generators, developed here, to deduce the quantization rules for angular momentum, without the need to solve the Schrödinger equation.

2.8.1 Introduction to Generators

(excerpted)

2.9 Vector Spaces and Hilbert Spaces

Linear vector spaces are used throughout science and engineering. Quantum Mechanics organizes cleanly into vector spaces, so it will make your life much easier if you understand them. Vector spaces also form a foundation for the future study of group theory. However, for spatial wave functions alone, you probably don’t need all of this section, because you can work most of the math without it. However, it is critically important to understand the concept of a “zero vector” (defined later), and its distinction from the number 0. For example, the zero vector is an essential part of how the Dirac algebra derives energy quantization in the harmonic oscillator.

For angular momentum states, and other finite dimensional quantum states, you really *need* to understand vector spaces, including the zero vector. Similarly to the harmonic oscillator, the zero vector is essential to deriving the requirement of half-integer multiples of $\hbar$ for all angular momenta.

Vector spaces derive from the common notion of vectors in elementary physics: such a vector is 2 or 3 real values, representing a magnitude and direction. The vectors are drawn in either a 2-dimensional space (the space of a plane), or a 3D space (the space of volumes). The number of real values in such a vector is the **dimension** of the vector: 2D vectors have two real numbers (x, y). 3D vectors have 3 real numbers (x, y, z).

Such physically based vector spaces have some important properties, which we detail below. But it turns out that many types of mathematical elements share important properties with physical vector spaces. Thus it becomes very useful to define *general* vector spaces, called simply **vector spaces**, which are both more general and more abstract than spatial vectors in physical space. We describe vector spaces as follows:

- Overview: a vector space comprises fields, groups, and more.
- Fields.
- Groups.
- Vector spaces, plus inner products, and operators.

2.9.1 Vector Space Overview

(excerpted)

3 Introduction to Scattering

Scattering is important for several reasons: lots of physics was discovered by scattering one thing off another. Rutherford and his students discovered atomic nuclei by scattering alpha particles off gold foil; Davisson and Germer discovered the wave nature of electrons by scattering them off a nickel block, thus ushering in the era of quantum mechanics. Today, the most advanced quantum theories are tested by scattering (aka "colliding") particles off each other, and measuring the angles and energies of particles in the shower of debris that results. (Scattering theory is also called "collision theory.") Furthermore, 1D scattering essentially includes the topic of tunneling, which is of great practical importance. Tunneling has many experimental applications, such as the scanning tunneling microscope (STM), and everyday practical uses, such as high-frequency tunnel diodes.

The math for scattering is covered in most books, but the big picture preceding the math is often skipped. We start here with the big picture, before delving into a few selected computations. Our goal is to convey the concepts, so that the presentations in standard texts are more accessible. ([Gos chapter 23] provides an unusually clear introduction to scattering.)

Recall that a "stationary" state has properties that do not change with time; however, it is not "static." For example, a particle in a pure momentum eigenstate is "stationary," but it is also moving (it has momentum). A particle in an orbital angular momentum eigenstate is also stationary, but "moving:" it is revolving around the center.

In the following analysis, we use a variant of delta function normalization that is often used in scattering with momentum eigenstates. We drop the prefactor, so:

$$\left| p \right\rangle \equiv \psi_p(x) = e^{ipx/\hbar} \qquad \text{(unit magnitude delta-function normalization)} .$$

In other words, we are simply choosing our normalization such that the (implicit) prefactor of "1" before $e^{ipx/\hbar}$ represents one particle. This can also be considered normalizing to one particle per unit volume, but scattering is generally analyzed one incident particle at a time.

We present a 5-step program for scattering, in increasing order of complexity:

1. 1D scattering: Solving Schrödinger's equation.
2. 3D classical scattering: defining cross sections.
3. 3D quantum scattering: Solving Schrödinger's equation, and defining quantum cross sections.
4. Born approximation: good for mild scattering (particle energy >> scattering potential).
5. Partial wave expansion, including the low-*l* approximation (good for particle wavelength >> target size).

Alternatively, one can view scattering as a time-dependent perturbation problem, and use Fermi's Golden Rule [Bay p252b], but we do not discuss that development here.

This whole chapter assumes you are familiar with solving the time-independent Schrödinger equation, wave-packets. The 3D scattering sections assume understanding angular momentum, spherical Bessel functions, and solutions to the Schrodinger equation in spherical coordinates.

3.1 1D Quantum Scattering

3.1.1 1D Scattering: Finite Width Barrier

Suppose we have a scattering potential, $V(x)$, which is localized (Figure 3.1, left). This potential is the "target." The force between the target and incident particle is short range, as indicated by the potential being zero outside a small region. Therefore, the particle is a free particle outside the range of the target. This assumption is also important in 3D scattering.

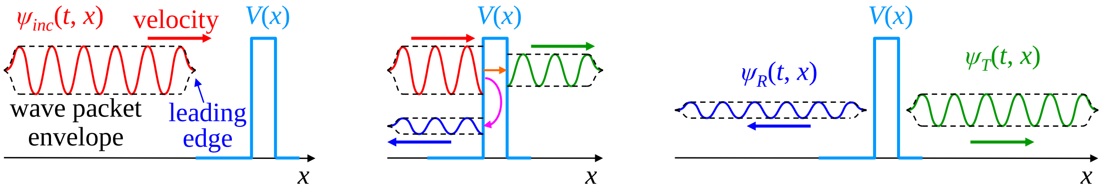


Figure 3.1 Not to scale. (Left) The particle wave-packet approaches from the left. (Middle) The particle interacts with the barrier potential, and is partially transmitted and partially reflected. (Right) The wave-function has split into spatially separate transmitted and reflected parts.

A *single* particle is incident from the left, with a very well-defined momentum, so it is described by a wave-packet that is very wide (narrow $\Delta p \Rightarrow$ wide Δx). The particle is "free" until it hits the target. A realistic wave-packet is much wider than the target width, so the diagram is not to scale. The wave-packet comprises waves of a narrow range of momenta (a superposition of momenta), and can be normalized. However, since the spread in momentum is very small, we approximate it as a momentum eigenstate, and use unit-amplitude δ-function normalization (the wave function is nearly a δ-function in the momentum basis).

The particle energy may be above or below the target height. Either way,
some of the wave-function is reflected, and some is transmitted through the target.

The particle interacts with the target ("collides", Figure 3.1, middle). This interaction splits the wave-packet into two parts: a transmitted part that continues past the target, and a reflected part that moves back to the left.

Classically, we're used to thinking of scattering as a *dynamic* process: a particle enters, collides, and leaves. In QM, however, the process can be computed as *stationary* (steady-state). Because the wave-packet is wide, when its center reaches the target, the packet envelope is nearly constant in space and time, but the wave-function phase changes with time. In other words, we have (essentially) a stationary quantum state, with the incident part of the wave-function moving to the right. The leading and trailing edges of the wave-packet are far from the barrier. Therefore, to a good approximation, Schrödinger's time-independent equation applies. Note that "stationary" does not mean "static." The particle is always moving, but during the collision, far from the wave-packet edges, the wave-function amplitudes do not significantly change with time. Therefore, the incident particle wave-function is closely approximated by that of a definite momentum:

$$\psi_{inc}(t,x) = e^{ipx/\hbar} e^{-iEt/\hbar} = e^{ikx} e^{-i\omega t} \qquad \text{where} \quad p = \hbar k, \ \text{ and } \quad E = \hbar\omega .$$

This is where all the math comes in: we solve Schrödinger's time-independent equation, and find the amplitudes for the transmitted piece and the reflected piece. We have to match the wave-function value and slope at both interfaces (front and back of barrier). Most standard references do this, so we don't discuss it here (see for example, [Gos chapter 4]).

After the collision is done, the particle has a probability to be found on the left, and moving left, and some probability to be found on the right, and moving right (Figure 3.1, right). That is, it has a probability to be reflected, and a probability to be transmitted.

Things to note: scattering amplitudes, and therefore the transmission and reflection probabilities, depend on p, the incident particle momentum, or equivalently on k, the incident particle spatial frequency (wave number). Furthermore:

In 1D barrier scattering, the energy (and therefore momentum, spatial frequency, and speed)
is the same on both sides of the barrier. Therefore,
the widths of the incident, transmitted, and reflected wave-packets are all the same.

If we take the amplitude of the incident wave-packet as unity (unit amplitude δ-function normalization), then the probabilities of transmission and reflection are simply:

$$|\psi_{inc}|^2 \equiv 1 \qquad \Rightarrow \qquad \Pr(\text{transmission}) = |\psi_T|^2, \quad \text{and} \qquad \Pr(\text{reflection}) = |\psi_R|^2 .$$

(excerpted)

3.2 Introduction to 3D Scattering

> The basic idea of 3D classical and quantum scattering is that we send a uniform spray of particles at a target, and they bounce off in all kinds of directions. However, we consider each incident particle individually, assuming no interaction between incident particles.

By "uniform spray" we mean a fixed number of particles per second per unit area: a uniform flux density. A **flux** is particles per second (*N*/s); **flux density** is particles per second per area (*N*/s/area). Flux density is also called "luminosity," especially in experimental physics. (Many references use the term "flux" to mean "flux density.") The incident beam is wide, much wider than the target, so its exact width is not a factor in the scattering. In realistic quantum scattering, the incident particles are far enough apart that they do not interact.

Before considering 3D *quantum* scattering, we start with an overview of *classical* scattering, which is the foundation on which we build.

3.2.1 3D Hard-Target Classical Scattering

"Hard target" scattering has rigid, impenetrable particles and a rigid, impenetrable target. (This situation does not exist at the quantum level.) Therefore, an incident particle either hits or misses the target. We consider the following topics in classical hard-target scattering:

- Incident particles.
- Scattered particles.
- Differential cross-section.
- Total cross-section.
- Measuring differential cross-section.
- Axially symmetric targets.
- Reduction to one body.

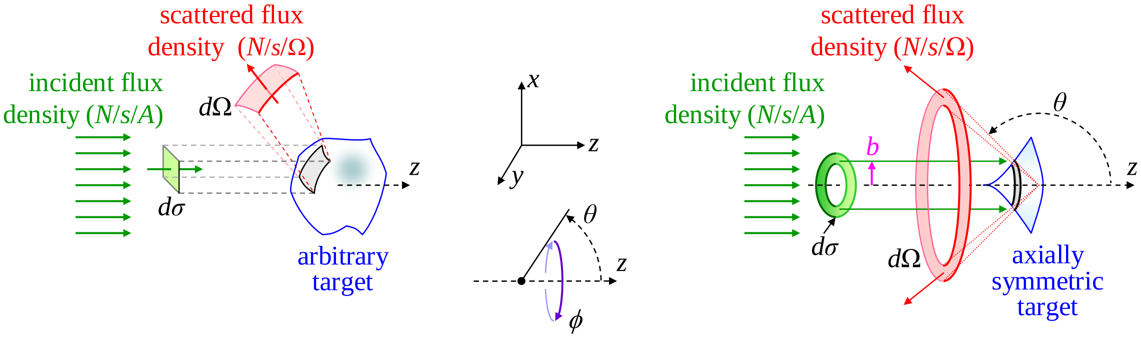


Figure 3.2 (Left) Classical scattering off an arbitrary target. (Right) Classical scattering off an axially symmetric target. '*b*' is the classical impact parameter.

Incident particles: Because the incident flux density is constant over the *x*-*y* plane, the flux through any area normal to the flux is just area times flux density (Figure 3.2):

$$\text{flux} \equiv \text{particles/s} = (\text{flux\_density})(\text{area}) \qquad \text{or}$$

$$F = JA \qquad \textit{where} \quad J \equiv \text{flux density} = \text{particles/s/area} .$$

A **hard-target** is a lump of hard stuff which is impenetrable, but has no effect outside the target boundary. (We discuss "soft-targets" later.) The scattering is usually elastic: the target is fixed (heavy) and acquires negligible energy; therefore the scattered particles have the same energy as the incident particles. In classical mechanics, the cross-section of a hard-target is literally the cross-sectional area of the target perpendicular to the flux, measured in, say, cm^2. By analogy, in quantum scattering, the *effective* area capturing flux is called a **cross-section**.

Scattered particles: The scattered particles are measured differently. Far from the target, it appears to be a point source of scattered particles. Since particles are conserved, the total outward flux through a large sphere centered on the (effectively point-source) target, including particles which "miss" the target and continue past it, equals the incident flux. For a differential solid angle $d\Omega$ measured far from the target, the flux per steradian is constant along the radius away from the target. Therefore, for *scattered* particles, the flux density is measured as particles per second per solid angle (particles/s/sr), rather than particles/s/area.

Differential cross section: Consider Figure 3.2, left. Each infinitesimal area, $d\sigma$, scatters into an infinitesimal solid angle, $d\Omega$, in the direction (θ, ϕ).

$$\text{particles/s into } d\sigma = \text{particles/s out of } d\Omega\ .$$

For infinitesimal $d\sigma$, if we vary its size, the $d\Omega$ into which it scatters will vary proportionately: if we double $d\sigma$, we will double $d\Omega$. In other words, in an infinitesimal region around any direction (θ, ϕ), the ratio $d\sigma/d\Omega$ is constant. We here call it $s(\theta, \phi)$:

$$s(\theta,\phi) \equiv \frac{d\sigma}{d\Omega} \qquad \text{(units of area/sr)}\ . \tag{3.1}$$

Generally, though, we think in the reverse direction: we can control the size ($d\Omega$) and position (θ, ϕ) of particle detectors counting scattered particles. Then we can ask (and measure): how big a $d\sigma$ contributes particles to the given $d\Omega$. Since the incident flux density j is uniform, we don't care *where* $d\sigma$ is, we only care how big it is. In general, the size (and position) of the incident cross section, $d\sigma$, that fills a given small solid angle, $d\Omega$, varies in different scattered directions, i.e. it is a function of (θ, ϕ).

$$J_{inc}\ d\sigma = \text{flux out of } d\Omega \quad \textit{where} \quad J_{inc} \equiv \text{incident flux density (particles/s/area).} \quad \text{Then:}$$

$$d\sigma = \frac{\text{flux out of } d\Omega}{J_{inc}} \equiv s(\theta,\phi)\ d\Omega\ .$$

The function $s(\theta, \phi)$ is (somewhat inappropriately) called the **differential cross-section**. Of course, the "differential cross-section" is literally just $d\sigma$, and is proportional to $d\Omega$, but this misnomer is universally used. [Gri p395b] notes that $s(\theta, \phi)$ is a derivative, not a differential, which we see from (3.1). The units of differential-cross section are cm^2/steradian, or just cm^2, because the steradian (like the radian) is equivalent to dimensionless.

Note that more than one location for $d\sigma$ could scatter into the same solid angle, $d\Omega$; this is only possible because far from the target, it appears to be a point source.

The concept of multiple incident regions scattering into the same solid angle is critical for quantum scattering.

In this case, the differential cross-section is simply the sum of all small incident areas that scatter into the given $d\Omega$.

The differential cross-section does not care which region of the incident beam the particles come from; it tells only how much total incident *area* scatters into a solid angle in a given direction.

Total cross-section: The **total cross-section** is the cross-sectional area of the incident flux that is scattered (in any direction). Therefore, we can find the total cross-section by integrating $d\sigma$:

$$\sigma = \int_{\text{sphere}} \frac{d\sigma}{d\Omega}\, d\Omega = \int_0^{\pi} d\theta \int_0^{2\pi} d\phi\; s(\theta,\phi)\,\sin\theta \qquad \textit{where} \quad d\Omega = d\theta \sin\theta\, d\phi\,.$$

Measuring differential-cross section: We measure differential cross-sections exactly the same way for all scattering (hard classical, soft classical, and quantum). Send in a known number of particles/cm^2 toward a target. Put small detectors around the target in all directions. Measure the detection rate in each direction, knowing the solid angle $d\Omega$ covered by each detector. Then simply divide:

$$d\sigma = \frac{\text{particles/s}}{j_{inc}} \quad \left(\text{in cm}^2\right), \qquad \frac{d\sigma}{d\Omega} \equiv s(\theta,\phi) = \frac{d\sigma}{\text{solid\_angle\_of\_detector}}\,.$$

Note that the detection rate (particles/s) of any detector is proportional to the incident flux density; therefore, the flux density cancels, and the cross-sections $d\sigma$ are independent of flux density; $d\sigma/d\Omega$ is a function solely of a single incident particle (including its energy) and the target.

Axially symmetric targets: Many targets are axially symmetric (Figure 3.2, right). In particular, spherically symmetric targets are also axially symmetric. Because of this symmetry, there is no ϕ dependence. Then the differential cross section is a function of θ only:

$$\frac{d\sigma}{d\Omega} = s(\theta) \qquad \text{(axially symmetric target)}\,.$$

We must still include ϕ when integrating to find total cross-section, but it integrates trivially to 2π:

$$\sigma = \int_0^{\pi} d\theta\; s(\theta)\,\sin\theta \int_0^{2\pi} d\phi = 2\pi \int_0^{\pi} d\theta\; s(\theta)\,\sin\theta \qquad \text{(axially symmetric target)}\,.$$

(excerpted)

3.3 3D Quantum Scattering

3.3.1 3D Quantum Scattering Overview

Quantum scattering combines wave scattering concepts with quantum probabilities. Quantum scattering is wave-function scattering, so interference is important. In QM, each small incident area has a probability of scattering into most any solid angle. Conversely, a given solid angle usually has contributions from a large (or infinite) incident area, i.e., a given $d\Omega$ has probabilities of coming from a large set of $d\sigma$'s. The overall QM $d\sigma$ for a given $d\Omega$ is a weighted sum of every possible infinitesimal incident area da, each weighted by its amplitude for scattering into $d\Omega$, and including the interference of all such contributors. Recall that:

We define (and measure) cross-sections by counting particles.

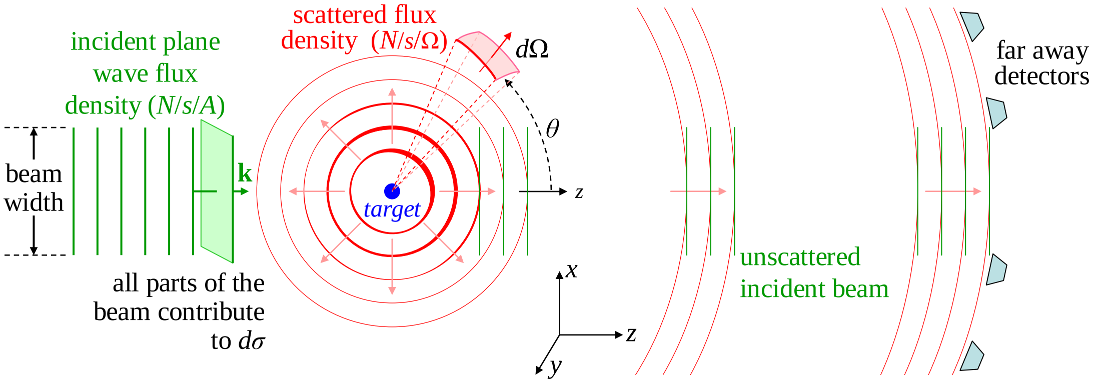


Figure 3.3 Quantum scattering: each small incident area scatters in all directions;. the contribution to a given solid angle is the incident area times the amplitude for scattering into that solid angle.

In particle-target quantum scattering, we usually consider the target fixed, and it absorbs negligible energy from the incident beam. (As noted earlier, 2-body collisions can be converted to this form.) We assume the interaction force is short-range (defined soon). Therefore, outside the range of the scattering potential, the outbound waves have (essentially) the same kinetic energy as the inbound waves, and the scattering is elastic. This means the particle is a free particle outside the range of the interaction force. The outbound wave is proportional to the incident wave-function, which is a plane wave, written as:

$$\psi_{inc}(z) = Be^{ikz} \qquad where \qquad B \equiv \text{normalization factor}.$$

From far away, the target appears to be a point source of scattered waves, so the scattered wave fronts must be spherical, i.e. proportional to exp(*ikr*). We can deduce the mathematical form by considering conservation of particle flux (sometimes called "conservation of probability"). In steady state, the particle flux *into* any thick (or thin) spherical shell must equal the flow *out*. In general, through any area in space, flux = flux-density times area, and flux-density equals volume-density times velocity:

$$F = Ja = \rho va = \left|\psi_{out}\right|^2 \frac{\hbar k}{m} a \qquad where \quad m \equiv \text{mass of particle}, \quad k \equiv \text{spatial frequency},$$

$$a \equiv \text{area},$$

$$v = \frac{p}{m} = \frac{\hbar k}{m} \equiv \text{quantum velocity of particle}.$$

Therefore, the scattered wave-function $\psi_{out}(\mathbf{r})$ must decrease as $1/r$, so that $|\psi_{out}|^2$ decreases as $1/r^2$, and the total flux through spheres of any radius is constant:

$$\text{flux} = v\int_{\text{sphere}} \left|\psi_{sc}(r,\theta,\phi)\right|^2 r^2\, d\Omega = v\int_{\text{sphere}} \left|\frac{e^{ikr}}{r}\right|^2 r^2\, d\Omega = const.$$

The outbound wave-function ψ_{sc} is thus a spherical wave (but not spherically symmetric). ψ_{sc} is also a solution to Schrödinger's equation with $V = 0$, since V is zero outside the range of the scattering potential. Therefore, placing the origin of our coordinates at the target, we define a scattering amplitude $f(\theta, \phi)$:

$$\psi_{sc}(r,\theta,\phi) = f(\theta,\phi)B\frac{e^{ikr}}{r}, \qquad \text{and}$$
$$\psi(x,y,z) = \psi_{inc}(z) + \psi_{sc}(r,\theta,\phi) = B\left[e^{ikz} + f(\theta,\phi)\frac{e^{ikr}}{r}\right]. \tag{3.2}$$

$f(\theta, \phi)$ is implicitly a function of the incident momentum (or k), as well, but we omit k for brevity, as is conventional. $f(\theta, \phi)$ has units of $\text{m}^{-1/2}$. B is arbitrary (it will cancel from all calculations), so we take $B = 1$ for simplicity (unit amplitude delta-function normalization). We have written the stationary-state wave-function $\psi(x, y, z)$ in the conventional manner on the RHS, which mixes both rectangular coordinates (z), and spherical (r, θ, ϕ) coordinates.

We show below that $d\sigma/d\Omega = |f(\theta, \phi)|^2$. Then the total cross-section is found by integrating over all angles, just as for classical scattering:

$$\sigma = \int_{\text{sphere}} \frac{d\sigma}{d\Omega}\, d\Omega = \int_0^{\pi} d\theta \int_0^{2\pi} \left|f(\theta,\phi)\right|^2 \sin\theta\, d\phi \qquad where \quad d\Omega = d\theta \sin\theta\, d\phi.$$

We can define "hard" quantum scattering as the limit of a potential that is infinite inside its boundary, and zero outside. But:

> Unlike classical scattering, the hard-target quantum scattering cross-section is *not* equal to the physical target cross-sectional area.

Scattering calculations may include many conditions on the incident and outbound particles, such as their spin, polarization, etc. The differential cross section $d\sigma/d\Omega$ for a given set of conditions is a statistical description of quantum scattering, and defines everything there is to know about such scattering.

Inelastic scattering: You sometimes hear about "inelastic" quantum scattering: this refers to scattering where the outbound particles are different than the inbound particles, and may have different total mass. For example, photons may be radiated, and carry away energy and momentum. Similarly, in condensed matter, scattering may radiate phonons, which carry away energy and momentum. Inelastic scattering includes the case where particles are "absorbed" into the target, such as photon absorption or nuclear capture. It also includes cases where the target changes state, and may *give* energy to the incident particle, *increasing* its outbound energy. Whenever the outbound kinetic energy may be different than the inbound kinetic energy, the scattering is **inelastic**.

Inelastic scattering can be modeled by a potential $V(\mathbf{r})$ which is complex [Sch sec. 20 p129], much like EM wave propagate in an absorbing medium with a complex wave-vector $\mathbf{k}$, but we do not address inelastic scattering further.

3.3.2 3D Scattering Off a Target: Schrödinger's Equation

(excerpted)

3.4 Partial Wave Method

In contrast to our discussion so far, the method of partial waves requires that the scattering target be spherically symmetric, so $V(\mathbf{r}) = V(r)$, where $r \equiv |\mathbf{r}|$. In principle, the method of partial waves gives an exact series solution to the spherically symmetric scattering problem [Blo p272b], though in general, neither the series, nor any of its terms, can be evaluated analytically. This section assumes you understand orbital (spatial) angular momentum (but not necessarily spin), and how and why the spherical harmonics are angular momentum eigenstates of spherically symmetric potentials. (We discuss orbital angular momentum in a later chapter.) This section is rather long. We proceed as follows:

- Angular momentum of a plane wave; inbound and outbound spherical waves.
- Conservation of angular momentum.
- Scattering of an angular momentum eigenstate.
- Calculating δ_l.
- Small k approximation.

Angular momentum of a plane wave: What is the angular momentum of the incident plane wave? It might seem at first that it is zero, and indeed its *average* is zero: $<L_x> = <L_y> = <L_z> = 0$. But we now show that it has nonzero squared magnitude, i.e. $<L^2> \ne 0$. This leads to an important method for computing scattering for *spherically symmetric* targets: partial waves. With partial waves, we decompose the incident plane wave into its L^2 eigenstate components (with quantum numbers l), and then treat each component separately. A single L^2 eigenstate can be easier to analyze than a plane wave, and we get a series solution in the quantum numbers l. Finally, we can estimate the magnitudes of the scattering as a function of l, and keep only the significant contributors. This allows an arbitrarily good approximation in a finite sequence of terms.

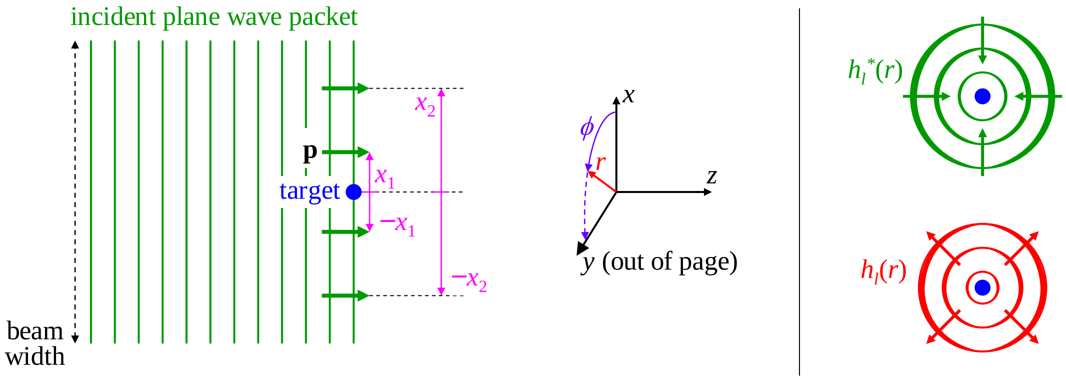


Figure 3.4 (Left) An incident plane wave comprises angular momentum components up to arbitrary *l*. The momentum, **p**, is constant everywhere, but there are many component impact parameters, e.g. x_1, x_2. (Right) Spherical inbound and outbound waves.

Consider the incident plane wave in Figure 3.4, left. Recall that $\hat{L}_z = \frac{\hbar}{i}\frac{\partial}{\partial \phi}$. We easily show that $\langle L_z \rangle = 0$ using *cylindrical* coordinates:

$$\hat{L}_z \psi_{inc}(r,\phi,z) = \frac{\hbar}{i}\frac{\partial}{\partial \phi} e^{ikz} = 0 \qquad \Rightarrow \qquad \langle L_z \rangle \equiv \langle \psi_{inc} | L_z | \psi_{inc} \rangle = 0 .$$

The left hand equation says that the local $L_z(\mathbf{r})$ is everywhere 0. Physically, as we picture the particle traveling to the right, it has no momentum in the *x* or *y* directions, and therefore has no reason to spiral around the *z* axis as it propagates. Hence $L_z = 0$.

Now consider L_y of the incident wave. Recall that $\mathbf{L} = \mathbf{r} \times \mathbf{p}$. Since $\mathbf{p} = p\hat{\mathbf{z}}$ is along *z*, $L_y = -xp$, i.e. the local angular momentum *about the target*, due to a point $\mathbf{r} = (x, y, z)$ on the wave-function $\psi(\mathbf{r})$, is $-xp$. We see from the diagram that each contribution to positive L_y from some $x < 0$ is canceled by an equal contribution to negative L_y from its mirror image with $x > 0$. Hence $\langle L_y \rangle = 0$. Similarly, $\langle L_x \rangle = 0$.

What about $L^2 = L_x^2 + L_y^2 + L_z^2$? Our prior result for L_z shows that $L_z^2(\mathbf{r}) = 0$ everywhere. However, L_y^2 is not zero. The two local components of L_y^2 now *add* rather than cancel. Similarly for L_x^2. Furthermore, for an infinitely wide plane wave, there are infinitely large values of *x* contributing, and therefore angular momentum components out to infinite magnitude.

Now recall that the spherical Hankel functions $h_l(\rho)$ are spherical radial functions propagating outward from the origin, and their complex conjugates $h_l^*(\rho)$ are spherical radial functions propagating inward. [h_l and h_l^* are sometimes written as $h_l^{(1)}$ and $h_l^{(2)}$.] The angular momentum eigenstates $h_l(kr)Y_{lm}(\theta, \phi)$ and $h_l^*(kr)Y_{lm}(\theta, \phi)$ are free-particle ($V = 0$) solutions to the Schrödinger equation ($l \equiv$ angular momentum quantum number), and so form an orthonormal basis. Therefore, we can write the incident plane wave as a sum of these angular momentum eigenstates. Since $L_z \propto \partial\psi/\partial\phi = 0$ everywhere, only the $m_l = 0$ functions contribute. Though it is traditional to write the expansion in terms of P_l rather than Y_{l0}, it is more consistent with QM notation to use the spherical harmonics Y_{l0}. Furthermore, we write it in terms of the Hankel functions, because we are interested in the outgoing and ingoing components separately:

$$\psi_{inc}(\mathbf{r}) = e^{ikz} = \sum_{l=0}^{\infty} i^l \sqrt{4\pi(2l+1)}\,\frac{1}{2}\left(\underbrace{h_l(kr)}_{out} + \underbrace{h_l^*(kr)}_{in}\right) Y_{l0}(\theta) \qquad \text{[Gos 23.34 p500b]}. \qquad (3.3)$$

This is a superposition of inbound and outbound *spherical* waves of *all* angular momenta *l*. They interfere in just the right way to add up to a plane wave. Note that while h_l and h_l^* are both irregular at the origin, their sum is the usual spherical Bessel function, and *is regular* at the origin:

$$\frac{1}{2}\left(h_l(kr)+h_l^*(kr)\right)=j_l(kr)\equiv \text{spherical Bessel function of order } l\,.$$

This explains why the expansion (3.3) is regular at the origin. However, we actually do *not* need regularity at the origin, since we will only use the Hankel functions in the vacuum outside the range of the scattering potential ($r > r_0$), and therefore excluding the origin.

(excerpted)

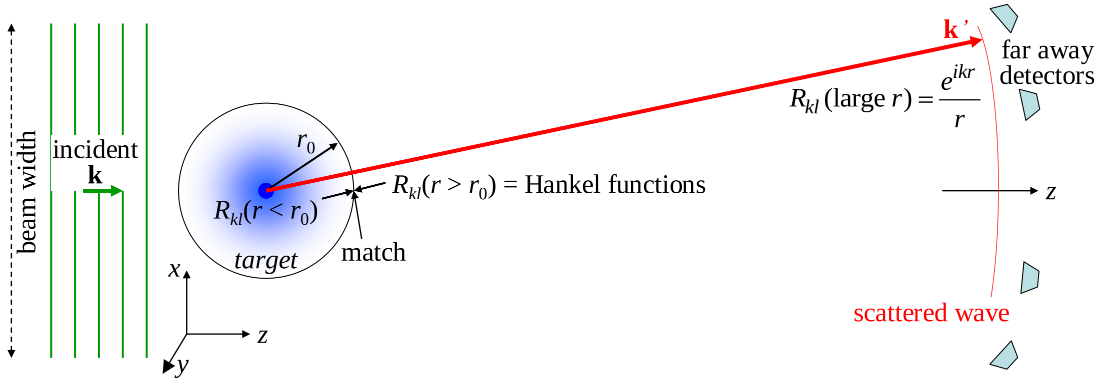


Figure 3.5 There are three regions to the solution for $\psi = \psi_{inc} + \psi_{sc}$: (1) inside the target, we can take $\psi(r < r_0)$ to be real. (2) In vacuum just outside the target, we describe ψ with Hankel functions. (3) At large r, ψ_{sc} is closely approximated by e^{ikr}/r.

(excerpted)

3.5 The Optical Theorem

The optical theorem is true for any coherent wave scattering, including electromagnetic waves, and a quantum particle. It holds with or without absorption, whenever the scattered wavelength is the same as the incident, even for asymmetric targets [Bay p202b]. However, the optical theorem requires that waves, as they propagate through the medium, maintain their phase coherence over a long distance. Therefore the medium must be uniform, and the incident particles must have a very narrow bandwidth.

Many quantum references demonstrate the optical theorem only for one or two limited cases, but it is very general. The theorem illustrates a substantive application of interference, as well as some important mathematical methods. We prove the theorem here by direct calculation of the interference. The derivation is exact, as each step is a rigorous limit as $z \to \infty$, and then as $\Delta\theta \to 0$.

This section relies on the general principles of scattering described earlier, especially the idea of cross-section. As always, we distinguish "flux" (particles/time) from "flux density" (particles/area/time).

Overview: Essentially, the optical theorem says that the flux scattered from a target exactly equals the flux depleted from the incident wave passing the target. For EM waves, the flux is energy/time or power, and the theorem is a statement of conservation of energy (power in = power out). For quantum mechanical scattering, the flux is particles/time, and it is a statement of conservation of particles (sometimes called conservation of probability): particles in = particles out. This simple idea has a surprisingly odd mathematical form:

$$\sigma_{total}=\frac{4\pi}{k}\operatorname{Im}\left\{f(\theta=0)\right\} \qquad \text{(the optical theorem)}\,.$$

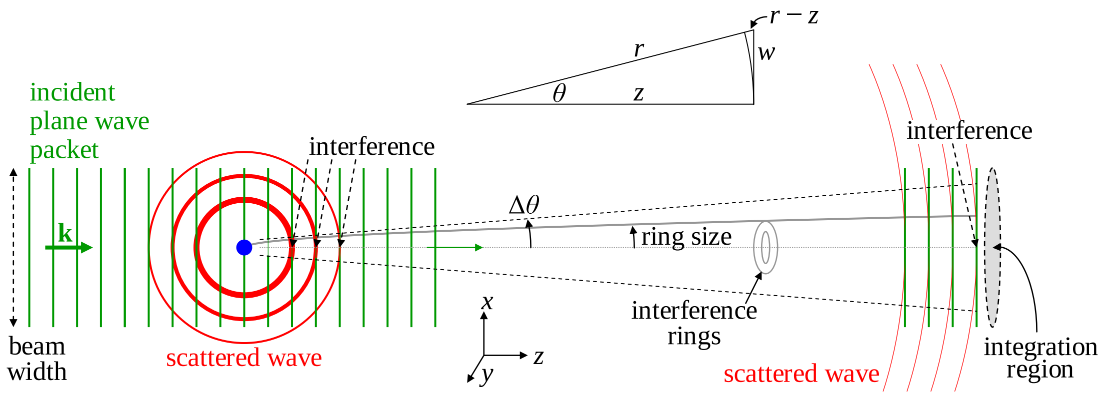


Figure 3.6 The forward scattered wave gets weaker with distance, but the "shadow" rings compensate by getting bigger. (Many wavefronts are omitted here for clarity.) The beam width is actually much bigger than the rings.

Figure 3.6 depicts the depletion of flux from the incident beam by the target. While this is sometimes called the "shadow" of the target, it bears little resemblance to an everyday macroscopic shadow. The "shadow" consists of circular interference rings across its face (Figure 3.6, though the ring edges are actually gradual, not sharp). The rings are circular even for asymmetric targets, because (as required for our previous scattering analysis), any short-range target looks like a point source from large distances, and hence scatters spherical waves (but not spherically symmetric waves). We choose a small angular width $\Delta\theta$ so that the scattering amplitude $f(\theta, \phi)$ is essentially constant. The scattered wave has a wavefront at $\theta = 0$, called the **forward scattered wave**. As the propagating forward scattered wave amplitude weakens with distance ($\sim 1/r$), its destructive interference of the incident wave decreases, but the interference *area* grows with distance. The two effects exactly compensate, creating a decrease in flux that is independent of distance. This missing flux must go somewhere, and it goes into the total scattered flux.

We consider the intensity across the x-y plane at some large z. As shown in Figure 3.6, the x and y distances covered by some small $\Delta\theta$ increase linearly with z. However, the radius of the interference rings increases more slowly, as $\sqrt{z}$. The two growth rates are shown in the figure. For a given $\Delta\theta$, there are more and more rings included as $z \to \infty$. This is important later, in choosing our limits of integration.

(excerpted)

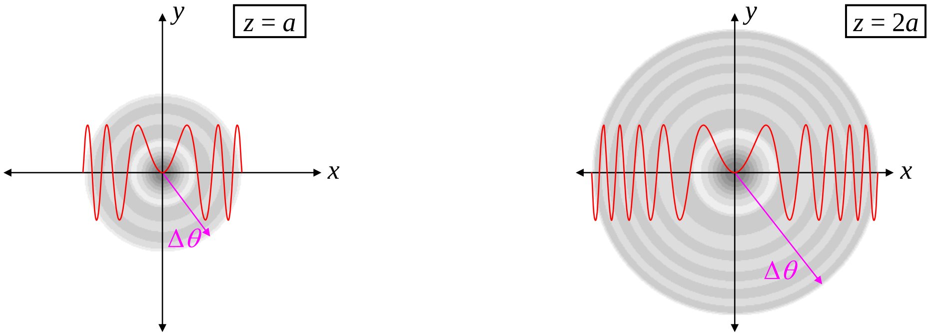


Figure 3.7 (Left) At moderate z, the "shadow rings" of the target, overlaid with the imaginary component of $\exp(ix^2)$. (Right) At twice the z, more interference rings are included within $\Delta\theta$.

(excerpted)

3.6 Identical Particle Scattering

We mention briefly here some considerations in 2-identical-particle scattering. Parts of this section rely on a knowledge of complete spin-space quantum states.

The case of identical particles requires more care regarding reduction to 1-body. As we will see in Multi-Particle Quantum Mechanics, identical particles always have a symmetry in the multi-particle wave-function. Identical fermions (matter particles such as electrons, muons, protons, quarks, ...) must have 2-particle wave-functions that are antisymmetric under particle exchange. If the two spins are in the symmetric state of total $S = 0$:

$$\psi_{12}(\mathbf{r}_1, \mathbf{r}_2) = -\psi(\mathbf{r}_2, \mathbf{r}_1) \qquad \text{(fermions with } S = 0) .$$

For example, e^--e^- scattering must satisfy this anti-symmetry. When reducing to 1-body, this anti-symmetry gets lost. For correct calculations, we must preserve the symmetry by imposing an equivalent one on the 1-body scattering analysis. In particular, when the one body moves from a point to a diametrically opposite point (across the target), that is equivalent to exchanging the two actual electrons, so for the 1-body wave-function, we must satisfy:

$$\psi(\mathbf{r}) = -\psi(-\mathbf{r}) \qquad \text{or} \qquad \psi(r, \theta, \phi) = -\psi(r, \pi - \theta, \phi + \pi) .$$

This says the 1-body wave-function must be odd parity.

As an example of the measurable consequence of this antisymmetry, consider e^--e^- scattering, where the two electrons collide "head on." For scattering into the *x-y* plane, $\theta = \pi/2$. In the 1-body picture, we have the usual axial symmetry of the incident particle around the *z*-axis. This means $\psi(\pi/2, \phi) = \psi(\pi/2, \phi + \pi)$. But antisymmetry requires that $\psi(\theta, \phi) = -\psi(\theta, \phi + \pi)$. The only way to satisfy both requirements is for the outbound wave to be zero along $\theta = \pi/2$. In terms of $f(\theta, \phi)$:

$$f(\theta = \pi/2) = 0 \qquad \text{(identical fermions with } S = 0) .$$

On the other hand, if the two fermions are in an antisymmetric spin state ($S = 1$), then the 2-particle wave-functions must be *symmetric* under particle exchange:

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = +\psi(\mathbf{r}_2, \mathbf{r}_1) \qquad \text{(fermions with } S = 1\text{, or bosons)} .$$

This is also true of identical bosons with *symmetric* spin states. For example, alpha particles are bosons of zero spin, and so require symmetric 2-particle wave-functions. This says the 1-body wave-function must be even parity.

Further discussion is outside our scope.

3.7 Scattering Conclusion

Our goal has been to familiarize you with the conceptual background needed to delve into existing mathematical developments of quantum scattering. There are many different approaches to scattering, and every book has its own pedagogical methods. It is probably most beneficial to consult several different references, so as to obtain a broad view of various aspects of scattering theory.

4 Matrix Mechanics

We've seen how quantum mechanical kets (states and other vectors) can be represented by wave-functions. Kets are the vectors of the vector space which encompasses a quantum system. Kets are either states, or the result of operators on states. Most of the previous chapters deal with wave mechanics, where the kets are continuous functions of space, and therefore the vector space has infinite dimension.

However, kets are often represented as finite-dimensional vectors, sometimes for convenience (e.g. orbital angular momentum), and sometimes by necessity (e.g., spin). Such a ket is called a **discrete ket** or "finite dimensional ket." It may be written as a column of N complex numbers, where the vector space has finite dimension, N. Most notably:

All of angular momentum, both orbital and spin,
can be described by finite dimensional vector spaces.

In addition, lots of other things can be represented or well-approximated by finite dimensional states, including an ammonia atom's nitrogen position, electron configurations in atoms and molecules, or excitations of an oscillator where only a finite number of states are likely.

Column vectors, row vectors, and matrices are perfect for QM, since they are *defined* to be the elements of linear transformations, and so represent the fundamental axiom of quantum mechanics: systems exist in a linear superposition of states. As such, no description of quantum mechanics is complete without **matrix mechanics**: the quantum mechanics of systems which can be represented by finite dimensional vectors and matrices. Just as wave-functions are vectors in a vector space, finite dimensional vectors are also vectors in a vector space. (Some quantum references call a finite dimensional vector space a "Hilbert space," but mathematicians insist a Hilbert space must be infinite dimensional. We therefore use the generic term "vector space" for finite dimensional cases.)

Because of the simple representation of discrete kets and operators as finite vectors and matrices, many QM concepts are easier to describe in finite dimensional systems, even if they also apply to continuum systems. For example, density matrices are much easier to visualize and understand in finite dimensional vector spaces.

Note that the dimension of the quantum state vector space describing a system has nothing to do with the dimension of the physical space of that system. Most of the systems we will consider exist in ordinary 3D space, but are described by quantum state spaces of many different dimensions. For example, particles orbit in 3D space, but the state space of orbital angular momentum for $l = 0$ is 1D, for $l = 1$ is 3D, for $l = 2$ is 5D, and for arbitrary l is $(2l + 1)$D.

4.1 Finite Dimensional Kets, Bras, and Inner Products

There is a strong analogy between continuous state space (wave-function space) and discrete state spaces. When written in Dirac notation, all the formulas of wave mechanics apply equally well to matrix mechanics, which illustrates again the utility of Dirac notation. Most of the wave mechanics algebra of kets, bras, and operators have simple analogs in finite dimensions. We describe those analogies as we go. Note that discrete space QM uses the standard mathematics of linear algebra, which is *not derived from* the continuous spaces, but is *analogous to* continuous spaces.

Finite dimensional kets have N components and N basis kets, i.e. any ket can be written as a linear combination of N basis vectors. E.g.,

$$\text{For } N = 3, \qquad |\psi\rangle = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = a_1|\phi_1\rangle + a_2|\phi_2\rangle + a_3|\phi_3\rangle .$$

(We often use 3D quantum state spaces as examples, because they are nontrivial and illustrative. However, this has nothing to do with ordinary 3D space.) In wave mechanics, a ket $|\psi> \leftrightarrow \psi(x)$, where given x, $\psi(x)$

is the complex value of the ket $|\psi>$ at position x. In finite dimensions, $|\psi> \leftrightarrow \psi_j$, where given an index j, ψ_j is the complex value of the j^{th} component of $|\psi>$. For general N:

$$\left|\psi\right\rangle = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix} = \sum_{k=1}^{N} a_k \left|\phi_k\right\rangle = a_1\left|\phi_1\right\rangle + a_2\left|\phi_2\right\rangle + ..., \quad where\ \left|\phi_k\right\rangle \equiv \text{basis kets, and } a_k \text{ are complex} .$$

Inner products: Inner products and bras are analogous to wave mechanics:

Let $\left|\chi\right\rangle = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$. Then:

$$\left\langle\chi\middle|\psi\right\rangle = \sum_{j=1}^{N} c_j^* a_j , \qquad \text{analogous to} \qquad \left\langle\alpha\middle|\beta\right\rangle = \int_\infty \alpha^*(x)\beta(x)\,dx .$$

Therefore, bras are written as row vectors, with conjugated components, so that an inner product is given by ordinary matrix multiplication:

$$\left\langle\chi\right| \equiv \left(\left|\chi\right\rangle\right)^\dagger = \begin{bmatrix} c_1^* & c_2^* & c_3^* \end{bmatrix} \qquad \Rightarrow \qquad \left\langle\chi\middle|\psi\right\rangle = \begin{bmatrix} c_1^* & c_2^* & c_3^* \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = c_1^* a_1 + c_2^* a_2 + c_3^* a_3 .$$

Recall that the dagger symbol acting on a ket produces the dual bra: $(|\chi>)^\dagger \equiv <\chi|$.

Kets are written as column vectors, and bras are written as row vectors.

The squared magnitude of a vector is then:

$$\left\|\psi\right\rangle^2 = \left\langle\psi\middle|\psi\right\rangle = \sum_{j=1}^{N} a_j{}^* a_j = \sum_{j=1}^{N} \left|a_j\right|^2 .$$

All of these definitions comply with standard mathematical definitions.

4.2 Finite Dimensional Linear Operators

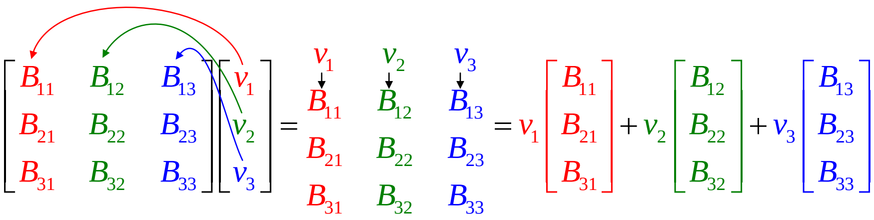
Operators acting on kets (vectors): Matrix multiplication is *defined* to be the *most general* linear operation possible on a discrete vector. Therefore:

Any discrete linear operator can be written as a matrix,
which operates on a vector by matrix multiplication.

The matrix elements are, in general, complex. For example, an operator in a 3D quantum state space (not physical 3D space) can be written:

$$\hat{B} = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{bmatrix} .$$

It is important to have a good mental image of a matrix multiplying a vector (Figure 4.1).



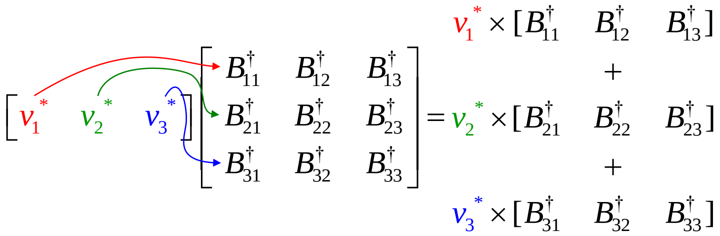
$$\begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{matrix} v_1 & v_2 & v_3 \\ \downarrow & \downarrow & \downarrow \\ B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{matrix} = v_1\begin{bmatrix} B_{11} \\ B_{21} \\ B_{31} \end{bmatrix} + v_2\begin{bmatrix} B_{12} \\ B_{22} \\ B_{32} \end{bmatrix} + v_3\begin{bmatrix} B_{13} \\ B_{23} \\ B_{33} \end{bmatrix}$$

Figure 4.1 Visualization of a matrix pre-multiplying a vector, yielding a weighted sum of the matrix columns.

Each component of the vector multiplies the corresponding column of the matrix. These "weighted" columns (vectors) are then added (horizontally) to produce the final result. Thus, when used as a linear operator, matrix multiplication of a vector converts each vector *component* into a *whole vector* with a full N components, and sums those vectors. Matrix multiplication is linear, which means:

$$\hat{B}\left(a|v\rangle + |w\rangle\right) = a\hat{B}|v\rangle + \hat{B}|w\rangle \qquad \text{for all} \quad a, |v\rangle, |w\rangle .$$

(excerpted)



$$\begin{bmatrix} v_1^* & v_2^* & v_3^* \end{bmatrix}\begin{bmatrix} B^\dagger_{11} & B^\dagger_{12} & B^\dagger_{13} \\ B^\dagger_{21} & B^\dagger_{22} & B^\dagger_{23} \\ B^\dagger_{31} & B^\dagger_{32} & B^\dagger_{33} \end{bmatrix} = \begin{matrix} v_1^* \times [B^\dagger_{11} \quad B^\dagger_{12} \quad B^\dagger_{13}] \\ + \\ v_2^* \times [B^\dagger_{21} \quad B^\dagger_{22} \quad B^\dagger_{23}] \\ + \\ v_3^* \times [B^\dagger_{31} \quad B^\dagger_{32} \quad B^\dagger_{33}] \end{matrix}$$

Figure 4.2 Visualization of a matrix post-multiplying a row vector, yielding a weighted sum of the matrix rows.

(excerpted)

4.3 Getting to Second Basis: Change of Bases

Quantum mechanics is a study of vectors, and vectors are often expressed in terms of components in some basis vectors:

$$\mathbf{r} = a\mathbf{e}_x + b\mathbf{e}_y + c\mathbf{e}_z \qquad \textit{where} \quad \mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z \text{ are basis vectors,} \quad \text{or}$$

$$|\chi\rangle = a|z+\rangle + b|z-\rangle \qquad \textit{where} \quad |z+\rangle, |z-\rangle \text{ are basis vectors} .$$

Our choice of basis vectors (i.e., our **basis**) is, in principle, arbitrary, since all observable calculations are independent of basis. However, most times one or two particular bases are significantly more convenient than others. Therefore, it is often helpful to *change our basis*: i.e., we transform our components from one basis to another. Note that such a transformation does *not* change any of our vectors; it only changes how we write the vectors, and how the internals of some calculations are performed, without changing any observable results.

A basis change transforms the *components* of vectors and operators;
it does *not* transform the vectors or operators themselves.

Angular momentum provides many examples where changing bases is very helpful. The infamous Clebsch-Gordon coefficients are used to change bases.

Many references refer to the "transformation of basis vectors," but this is a misnomer. We don't transform our basis vectors; we choose new ones. We can write the new basis vectors as a superposition of old basis vectors, and we can even write these relations in matrix form, but this is fundamentally a different mathematical process than transforming the components of a vector.

(excerpted)

4.4 Density Matrices

Density matrices and mixed states are important concepts required for many real-world situations. Experimentalists frequently require density matrices to describe their results. Density matrices are the quantum analog of the classical concept of ensembles of particles (or systems). Ensembles are used heavily in Statistical Mechanics.

Up until now, we have described systems of distinct "particles," where the particles are in definite quantum states. Even in a definite state, though, the measurable (i.e., observable) properties may be statistical (and thus not definite). This latter property, quite different from classical mechanics, gives rise to a striking new quantum result for the classical concept of ensembles. An **ensemble** is just a bunch of identical particles (or systems), possibly each in a different state, but we know the *statistics* of the distribution of states. For example, a particle drawn from a thermal bath is in an unknown quantum state, due to the randomness of thermal systems. However, if we draw many such particles (an ensemble) from the bath, we can predict the statistics of their properties from the bath temperature. While a known quantum state of a particle may be given by a ket, the state of an ensemble, or of a single particle drawn from it, is given by a density matrix.

We consider here ensembles only for finite dimensional (and therefore discrete) quantum systems, though the concept extends to more general (continuous) systems. We use some examples from the quantum mechanics of angular momentum, which is a topic discussed later in this book.

Instead of having a single particle in a definite state, suppose we have an ensemble of particles. If all the particles in the ensemble are in identical quantum states, then we have nothing new. All our QM so far applies to every particle, and extends to the ensemble as a whole. But suppose the ensemble is a *mixture* of particles in several *different* quantum states. What then? Can we compute average values of measurable quantities? If we know the fractions of all the constituent states in the ensemble, then of course we can compute the average value of any observable, and we do it in the straightforward, classical way. We will see that this idea of a classical mixture of quantum particles leads to a "density matrix:" a way of defining all the properties of such a mixture. However, we will also see that quantum ensembles have a highly *non-classical* nature.

The density matrix is essentially the quantum state of an ensemble.

For example, suppose we have an ensemble of electrons, 3/4 are spin |*z*+>, and 1/4 are spin |*x*+>. If we measure spin in the *z*-direction of many particles from the ensemble, we'll get an average which is the simple weighted average of the two states:

$$\text{For } |z+\rangle: \qquad \langle s_z\rangle_{|z+\rangle} = \frac{\hbar}{2}, \qquad \text{For } |x+\rangle: \qquad \langle s_z\rangle_{|x+\rangle} = 0. \qquad \text{Then:}$$

$$[s_z]_{ensemble} = \frac{3}{4}\langle s_z\rangle_{|z+\rangle} + \frac{1}{4}\langle S_z\rangle_{|x+\rangle} = \frac{3}{4}\cdot\frac{\hbar}{2}.$$

Following [Sak], we use square brackets [*B*] to explicitly distinguish the *ensemble* average of an observable $\hat{B}$, from a *"pure" state* average $\langle B\rangle \equiv \langle\psi|\hat{B}|\psi\rangle$, the average of an observable for a particle in a known quantum state (which may be a superposition). This average [*B*] is a number, distinct from the matrix for the operator $\hat{B} \equiv \left[\hat{B}\right]$. Expanding the average values above into bra-operator-ket inner products, we get:

$$\langle s_z \rangle_{|z+\rangle} = \langle z+|\hat{s}_z|z+\rangle = \frac{\hbar}{2}, \qquad \langle s_z \rangle_{|x+\rangle} = \langle x+|\hat{s}_z|x+\rangle = 0. \qquad \text{Then:}$$

$$[s_z]_{ensemble} = \frac{3}{4}\langle z+|\hat{s}_z|z+\rangle + \frac{1}{4}\langle x+|\hat{s}_z|x+\rangle = \frac{3}{4}\cdot\frac{\hbar}{2}.$$

So far, this is very simple.

Now let's consider a more general case: an ensemble consists of a mix of an arbitrary number of quantum states, each with an arbitrary fraction of occurrence (i.e., probability). Note that even in finite dimensional systems, there are an ***infinite*** number of quantum states, because the N basis vectors can be combined with complex coefficients in infinitely many ways. Therefore, the number of states in the mixture is unrelated to the Hilbert space dimension, N. Say we have a mix of M states, $|\psi^{(k)}>$, $k = 1, \ldots M$, each with a fraction of occurrence in the ensemble (or weight) w_k. As in the spin example, we can simply compute the average value of many measurements of particles from the ensemble by taking a weighted average of the quantum averages:

$$[B] = \sum_{k=1}^{M} w_k \left\langle \psi^{(k)} \middle| \hat{B} \middle| \psi^{(k)} \right\rangle \qquad \textit{where} \quad w_k \text{ are real, and } \sum_{k=1}^{M} w_k = 1. \qquad \text{[Sak 3.4.6 p177]}$$

A mixed state is quite different from a superposition. For one thing, a mixed state has no phase information relating the constituent states $|\psi_k>$: the w_k are real.

Everything we've done so far is independent of basis: the $|\psi^{(k)}>$ are arbitrary states, and will be superpositions in some bases, but not others. We use the term **constituent** to mean one of the states, $|\psi^{(k)}>$, of the mixture. This is distinct from "component," which refers to the complex coefficient of a basis function in a superposition. The constituents in a mixture are quantum states, independent of basis.

An ensemble with only one constituent ($M = 1$) is called a **pure state**: each particle is in a definite quantum state.

4.4.1 Development of the Density Matrix

(excerpted)

5 Angular Momentum

Angular momentum is a critical part of quantum mechanics, with applications throughout spectroscopy, magnetism, and solid state physics. This chapter assumes you are somewhat familiar with quantum angular momentum, and therefore we focus on the more conceptually challenging aspects.

5.1 Rotation Notation

We introduce some notation here, but all of it should become more clear as we use it throughout the text. "$\hat{J}$" is a common symbol for an arbitrary angular momentum operator (orbital, spin, combination, ...). To distinguish between operators and unit vectors, we use hats for operators, and bold **e** for unit vectors:

$\hat{L}_\theta$ is the operator for angular momentum in the θ direction.

$\mathbf{e}_\phi$ is the unit vector in the ϕ direction.

A quantum **vector operator** is a set of three operators, associated with the three basis directions in real space. For angular momentum vector operators we have:

$$\hat{\mathbf{J}} \equiv \hat{J}_x \mathbf{e}_x + \hat{J}_y \mathbf{e}_y + \hat{J}_z \mathbf{e}_z \qquad \text{the angular momentum vector operator,}$$

$$\textit{where} \quad \mathbf{e}_x \equiv \text{unit vector in } x\text{-direction of 3-space } (\hat{\mathbf{x}} \text{ in non-quantum language), etc.}$$

$$\hat{J}^2 \equiv \left|\hat{\mathbf{J}}\right|^2 = \left(\hat{J}_x\right)^2 + \left(\hat{J}_y\right)^2 + \left(\hat{J}_z\right)^2 \qquad \text{the magnitude-squared operator}.$$

The eigenstates (orbital or spin or combination) eigenstates of $\hat{J}^2$ and $\hat{J}_z$ are written:

$$|j\ m\rangle \quad \text{or} \quad |j, m\rangle \qquad \textit{where} \quad j \text{ is the quantum \# for } \hat{J}^2; \ \ m \text{ is the quantum \# for } \hat{J}_z .$$

For spin-1/2 particles in the z-basis, the following seven notations in common use are equivalent:

$$\chi_+ \equiv |\uparrow\rangle \equiv |+\rangle \equiv |z+\rangle \equiv |+z\rangle \equiv \left|\frac{1}{2}, \frac{1}{2}\right\rangle \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\chi_- \equiv |\downarrow\rangle \equiv |-\rangle \equiv |z-\rangle \equiv |-z\rangle \equiv \left|\frac{1}{2}, -\frac{1}{2}\right\rangle \equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

For combining two angular momenta (of arbitrary nature: spin, orbital, ...), the following notations are in common use, and use capital letters for total angular momentum, and lower-case letters for the constituent angular momenta:

$$\mathbf{J} = \mathbf{j}_1 + \mathbf{j}_2 \qquad \text{(the vectors in 3-space)} \qquad \textit{where} \quad \mathbf{J} \equiv \text{total angular momentum,}$$

$$\hat{\mathbf{J}} = \hat{\mathbf{j}}_1 + \hat{\mathbf{j}}_2 \qquad \text{(the quantum operators)}.$$

In the "uncoupled" basis, we write states of the combined system of angular momenta as combinations of the two original (uncoupled) angular momenta. There are several common, equivalent notations:

$$|j_1\ m_1\rangle \otimes |j_2\ m_2\rangle \equiv |j_1\ m_1\rangle |j_2\ m_2\rangle \equiv |j_1\ m_1; j_2\ m_2\rangle \equiv |j_1, j_2; m_1, m_2\rangle$$

$$\equiv |m_1, m_2\rangle \qquad \textit{where} \quad j_1, j_2 \text{ are "understood"}.$$

The alternative basis is the "coupled" basis (combined total angular momentum), with quantum numbers J and M, and eigenstates:

$$|J\ M\rangle, \quad \text{eigenstates of} \quad \hat{J}^2, \hat{J}_z .$$

Note the distinction between **J** and *J*: **J** is an angular momentum vector in 3-space; *J* is an angular momentum quantum number of the combined total angular momentum. In particular, *J* is *not* the magnitude of **J**:

$$J \ne |\mathbf{J}|, \quad \text{but} \qquad \hat{J}^2 = \hat{\mathbf{J}} \bullet \hat{\mathbf{J}} = \left|\hat{\mathbf{J}}\right|^2 .$$

In all of the above definitions, systems of purely orbital angular momentum might use **L**, *L*, *l*, etc. in place of **J**, *J*, *j*, etc. Systems of purely spin angular momentum might use **S**, *S*, *s*, etc.

Example: For spin ½:

$$\hat{\mathbf{s}} = \frac{\hbar}{2}\left(\sigma_x \mathbf{e}_x + \sigma_y \mathbf{e}_y + \sigma_z \mathbf{e}_z\right) \qquad \Rightarrow \qquad \hat{s}_x = \frac{\hbar}{2}\sigma_x, \qquad \hat{s}_y = \frac{\hbar}{2}\sigma_y, \qquad \hat{s}_z = \frac{\hbar}{2}\sigma_z .$$

Eigenstates: $\left|\uparrow\right\rangle, \ \left|\downarrow\right\rangle .$

5.2 Dissection of Orbital Angular Momentum

5.2.1 Orbital Angular Momentum as Fluid

Imagine water circulating in a closed circular loop of frictionless pipe. The system is stationary (no property changes with time), however it has angular momentum. If we try to tilt the loop, we will experience the resistance typical of tilting a gyroscope (bicycle wheel, etc.). Thus, even though the system is stationary, we can measure its angular momentum. The system is only dynamic if we look on a microscopic scale at the individual water molecules.

An electron in orbit around a nucleus is similar: it is stationary, i.e. no property changes with time. However, it has angular momentum. The electron is like a compressible fluid, distributed in space, with a particle-density and momentum-density given by its wave-function. In contrast to the water example above, so far as anyone knows, electrons have no smaller microscopic scale to look at: there are no “molecules” composing the electron. The electron appears to be a sort of “perfect fluid,” no matter how closely we look.

5.2.2 Spherical Harmonics as Motion

(excerpted)

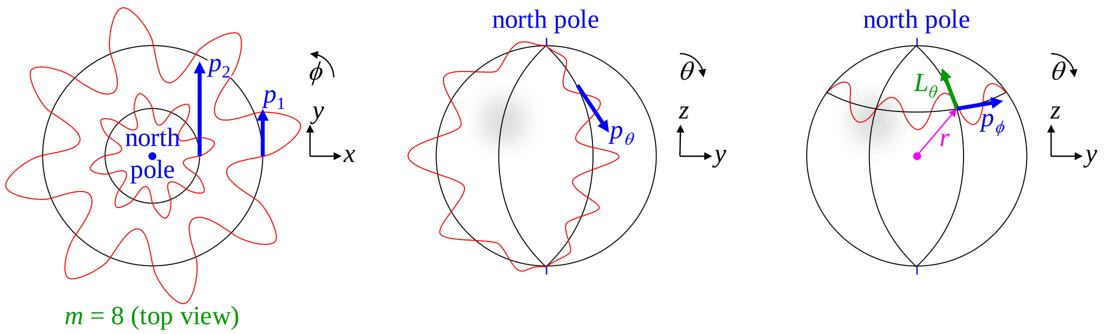


Figure 5.1 (Left) Linear momenta in the ϕ-direction at two points for m = 8: $e^{i8\phi}$. (Middle) Linear momentum in the θ-direction. (Right) Angular momentum in the $-\theta$ direction.

(excerpted)

5.3 Generating Station: A Different Spin on Angular Momentum

Particle “spin” is a kind of angular momentum with extremely important observable effects across a wide range of phenomena, but it has no classical analog. The magnitude of a particle’s spin can never be changed, but its direction can. For charged particles, angular momentum comes with a magnetic dipole moment, which is a critical element of a chemical compound's radiation spectrum. (Even neutrons have a magnetic dipole moment, due to the charged quarks inside them.) Spin-related spectroscopy is an essential tool for understanding chemical structure, as well as many other microscopic phenomena. Also, spin magnetic moment is the essence of MRI (Magnetic Resonance Imaging), an important medical diagnostic. And someday, spintronics, based on recording information in electron spins, may improve our electronic devices.

To facilitate understanding spin, and to illustrate its close association with rotations, as well as with orbital angular momentum, we demonstrate here some of the mathematical tools that are universally used for analyzing spin. Motivated by our discussion of generators (section 2.8.1), we follow this course:

- Orbital angular momentum, and its commutation relations.
- Rotations in real space, and the commutators of rotation operators.
- Quantum rotations, which must have the same commutators as classical.
- Quantum generators of rotation.
- Commutation relations of quantum generators of rotation, derived from rotation commutators.

We follow a similar approach to [S&W p257-60].

Orbital angular momentum: We review here the commutation relations of orbital angular momentum, as derived from the commutation relations of position and momentum, $[x, p_x] = i\hbar$. Classically, angular momentum is $\mathbf{L} \equiv \mathbf{r} \times \mathbf{p}$. Quantum mechanically, $\hat{\mathbf{r}}$ and $\hat{\mathbf{p}}$ are observable operators, which means they are local (they depend only on a point, or an infinitesimal neighborhood around a point). New local operators can be derived from old local operators using the classical relationships. Thus:

$$\hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x .$$

Since there is nothing special about our choice of labeling the axes x, y, and z, the same relation must hold for any right-handed set of axes, such as a circular permutation of coordinate labels:

$$\hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x \quad \Rightarrow \quad \hat{L}_x = \hat{y}\hat{p}_z - \hat{z}\hat{p}_y, \quad \text{and} \quad \hat{L}_y = \hat{z}\hat{p}_x - \hat{x}\hat{p}_z .$$

From these definitions, and

$$\left[\hat{x}, \hat{p}_x\right] = \left[\hat{y}, \hat{p}_y\right] = \left[\hat{z}, \hat{p}_z\right] = i\hbar, \qquad \text{and}$$

$$\left[\hat{x}, \hat{p}_z\right] = \left[\hat{y}, \hat{p}_z\right] = \left[\hat{x}, \hat{p}_y\right] = \left[\hat{z}, \hat{p}_y\right] = \left[\hat{y}, \hat{p}_x\right] = \left[\hat{z}, \hat{p}_x\right] = 0 ,$$

we find the well-known commutators of angular momentum:

$$\begin{aligned}
\left[\hat{L}_x, \hat{L}_y\right] &= \left(\hat{y}\hat{p}_z - \hat{z}\hat{p}_y\right)\left(\hat{z}\hat{p}_x - \hat{x}\hat{p}_z\right) - \left(\hat{z}\hat{p}_x - \hat{x}\hat{p}_z\right)\left(\hat{y}\hat{p}_z - \hat{z}\hat{p}_y\right) \\
&= \left(\hat{y}\hat{p}_z\hat{z}\hat{p}_x - \cancel{\hat{y}\hat{p}_z\hat{x}\hat{p}_z} - \cancel{\hat{z}\hat{p}_y\hat{z}\hat{p}_x} + \hat{z}\hat{p}_y\hat{x}\hat{p}_z\right) - \left(\hat{z}\hat{p}_x\hat{y}\hat{p}_z - \cancel{\hat{z}\hat{p}_x\hat{z}\hat{p}_y} - \cancel{\hat{x}\hat{p}_z\hat{y}\hat{p}_z} + \hat{x}\hat{p}_z\hat{z}\hat{p}_y\right) \\
&= \left(\hat{y}\hat{p}_z\hat{z}\hat{p}_x - \hat{z}\hat{p}_x\hat{y}\hat{p}_z\right) + \left(\hat{z}\hat{p}_y\hat{x}\hat{p}_z - \hat{x}\hat{p}_z\hat{z}\hat{p}_y\right) = \left(\hat{p}_z\hat{z} - \hat{z}\hat{p}_z\right)\hat{y}\hat{p}_x + \left(\hat{z}\hat{p}_z - \hat{p}_z\hat{z}\right)\hat{x}\hat{p}_y \\
&= \left(-i\hbar\right)\hat{y}\hat{p}_x + \left(i\hbar\right)\hat{x}\hat{p}_y = i\hbar\left(\hat{x}\hat{p}_y - \hat{y}\hat{p}_x\right) = i\hbar\hat{L}_z .
\end{aligned}$$

Again, the labeling of axes is arbitrary, so any circular permutation of indexes must also follow:

$$\left[\hat{L}_x, \hat{L}_y\right] = i\hbar\hat{L}_z \quad \Rightarrow \quad \left[\hat{L}_y, \hat{L}_z\right] = i\hbar\hat{L}_x, \quad \text{and} \quad \left[\hat{L}_z, \hat{L}_x\right] = i\hbar\hat{L}_y . \tag{5.1}$$

These angular momentum commutation relations turn out to be extremely important, and they describe more than just *orbital* angular momentum.

To expand on their significance, we now derive these relations in a more general way, from the mathematics of physical objects rotated in real space.

Rotations in real space: Before considering quantum rotations, we first consider classical rotations of a macroscopic body, such as a soccer ball. We imagine a point on the ball (perhaps where the inflation hole is), at coordinates (x, y, z). We then rotate the ball (about its center), and find the point has moved to a new position, (x', y', z'). More generally, we can replace the inflation hole by an arbitrary position vector, from the origin to the point (x, y, z). We consider "lab frame" rotations, where the x, y, z axes are fixed in the "lab," and do not rotate with the body.

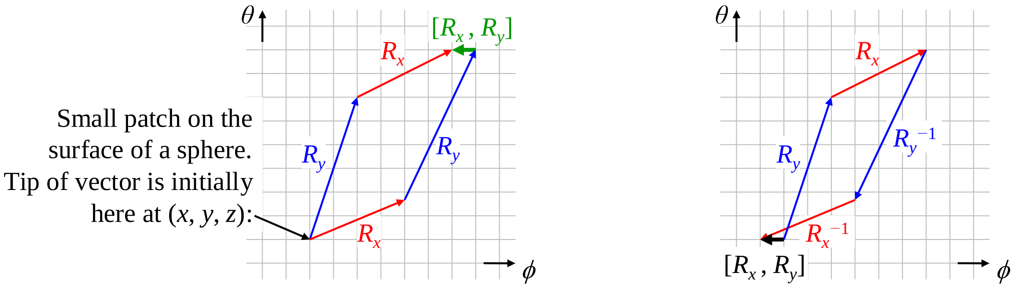
Rotations are linear operators: (1) if I double the size of a vector, and then rotate it, that's the same as rotating first, then doubling its size; and (2) if I add two vectors, and rotate the sum, that's the same as rotating both vectors, and then adding them. All linear operators (in finite dimensional vector spaces) can be written as matrices. From elementary trigonometry, we find for any angle α (the right hand rule defines positive α):

$$R_x(\alpha)=\begin{pmatrix}1 & 0 & 0\\ 0 & \cos\alpha & -\sin\alpha\\ 0 & \sin\alpha & \cos\alpha\end{pmatrix},\quad R_y(\alpha)=\begin{pmatrix}\cos\alpha & 0 & \sin\alpha\\ 0 & 1 & 0\\ -\sin\alpha & 0 & \cos\alpha\end{pmatrix},\quad R_z(\alpha)=\begin{pmatrix}\cos\alpha & -\sin\alpha & 0\\ \sin\alpha & \cos\alpha & 0\\ 0 & 0 & 1\end{pmatrix} \qquad (5.2)$$

$$[\text{S\&W } 12.45 \text{ p}258].$$

From our prior quantum experience, we know that commutators are important things (we also know it from Group Theory). Since even quantum systems can be rotated like classical systems (e.g., angular momentum can be made to have a definite value in any chosen direction), the commutators of quantum rotation operators must be the same as the commutators of classical rotation operators.

We first illustrate the commutator of rotations by an *infinitesimal* angle ε (Figure 5.2, left). Consider the effect of rotating a unit vector by $R_y(\varepsilon)$ first, then $R_x(\varepsilon)$, compared to rotating by $R_x(\varepsilon)$ first, then $R_y(\varepsilon)$. The difference is a small vector of order ε^2 .



Small patch on the surface of a sphere. Tip of vector is initially here at (x, y, z):

Figure 5.2 (Left) $[R_x, R_y]$ acting on a vector produces a small, 2nd order displacement (green). (Right) $\mathbf{1}_{op}+[R_x, R_y]$ displaces the original vector by the amount of $[R_x, R_y]$ acting on it (black).

(excerpted)

5.4 Spin ½

We reserve the term "wave-function" for spatial quantum states. We use the term **spin-state** to refer to the intrinsic spin of a particle (or particles). The combination of a wave-function and a spin-state is a complete quantum description of a particle, a "quantum state."

5.4.1 Spin Kets, Bras, and Inner Products

Spatial kets and bras can be thought of as shorthand for functions of space (e.g., wave-functions). Spin states cannot. Spin basis kets are abstract, formless things that obey simple rules. They cannot be represented as functions of space, or anything else. Spin inner products are defined using the presumed (or *defined*) orthonormality of spin basis kets.

In particular, suppose we measure a fermion (spin-1/2 particle) spin along some axis (perhaps using a Stern-Gerlach device). Experiment shows that only two values are possible; thus we assume that any spin state along this axis is a general superposition of exactly two spin eigenstates. Let's call them |z+> and |z->. Subsequent measurements along the same axis always produce the same value of spin, i.e. there is 100% chance of measuring the same spin again, and 0% chance of measuring the opposite spin. This behavior *defines* the inner product of spin states. If we believe that spin states behave similarly to spatial states, we must then say:

$$\langle z+|z+\rangle = 1, \qquad \langle z-|z-\rangle = 1, \qquad \text{and} \qquad \langle z-|z+\rangle = 0\,.$$

Therefore, |*z*+> and |*z*–> form a complete, orthonormal basis for spin states along this axis. Furthermore, subsequent experiments performing measurements along different axes confirm that:

|*z*+> and |*z*–> form a complete, orthonormal basis for a
spin state that defines a particle's spin behavior along *all* axes.

Note that for spatial wave-functions, the kets and bras can be thought of as just "notation," in that we know that the ket |ψ> is "really" a complex valued function of space. In other words, the ket notation is shorthand for a Hilbert space vector which has a known internal structure: a complex function of space. This "internal structure" *does not exist* for spin kets.

The spin kets are abstract vectors with *no further structure*.

All we know about |*z*+> is what we've described above, and we know it from experiment. End of story.

5.4.2 Spinors For Spin ½ Particles

Spinors are kets (and bras) that represent the spin-state of particles with intrinsic angular momentum. We discuss here spin-½ particles (**fermions**). When measured, spin ½ particles produce a component of angular momentum along the measurement axis that is always either +$\hbar$/2 ("up"), or –$\hbar$/2 ("down"). Since a particle's spin can be in a superposition of these two states, a spinor is a 2D vector of complex numbers.

$$|\chi\rangle = \begin{pmatrix} a \\ b \end{pmatrix} \qquad \text{(a spinor)}\,.$$

Every spin-½ spinor is an eigenspinor of *some* direction, so loosely, every spinor "points" in some direction. Be careful, though, that this is absolutely *not* the classical view of angular momentum having a definite direction in space. Spinors, like wave functions, describe the two *probabilities* of measuring the two possible values of angular momentum along any axis. A spinor is a 2D vector of complex numbers, which represents a direction in 3D physical space. From it, you can compute the probability of measuring the spin being up or down along *any* axis.

At first blush, a spinor might seem to have 4 degrees of freedom: 2 real numbers in each of 2 complex components. However, every quantum system of states has an arbitrary overall phase, i.e. you can multiply every ket in your system, at any time, by any unit-magnitude complex number, and you still have exactly the same quantum system. This accounts for one real-number degree of freedom out of the spinor's 4. Secondly, the spinor (like all quantum state kets) must be normalized: the sum of the squares of the component magnitudes must be one. This constraint removes another degree of freedom from the spinor, leaving just two degrees of freedom. It takes two degrees of freedom to represent any direction in 3D physical space (such as with 2 coordinates: θ, ϕ).

A common basis for spinors is the *z*-basis: i.e., we choose our *z*-axis to be along the measurement axis (sometimes called the "quantization axis"). The spinor can be written as the sum of |*z*+> and |*z*–> basis kets:

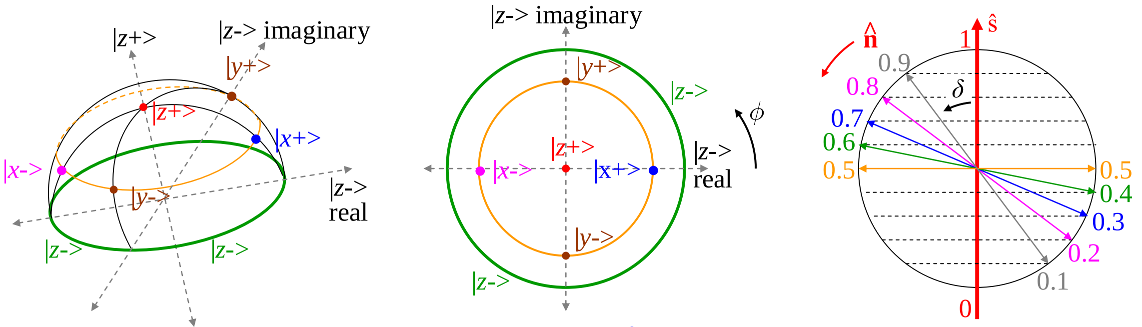
$$|\chi\rangle = a|z+\rangle + b|z-\rangle = \begin{pmatrix} a \\ b \end{pmatrix}, \qquad \textit{where} \quad \begin{matrix} |z+\rangle \text{ component is always on top} \\ |z-\rangle \text{ component is always on bottom}\,. \end{matrix} \tag{5.3}$$

The |*z*+> ket is the eigenket for pointing in the +*z* direction. A particle in the |*z*+> state, when measured along the *z* axis, will always measure spin "up," i.e. positive. Thus, Pr(*z*+) = 1, and Pr(*z*–) = 0. A particle in the |*z*–> state, when measured along the *z* axis, will always measure spin "down," i.e. negative. Thus, Pr(*z*+) = 0, and Pr(*z*−) = 1.

It might seem odd that a spinor can point in *any* direction in 3-space with only the two "*z*-component" basis kets. But recall that each *z*-component is a complex number; they are interpreted as follows:

The magnitudes of the two spinor components define the *z*-axis spin properties of the particle.
The relative phase angle between the two components
defines the spin properties in the *x*-*y* directions.

We can illustrate how spinors map to 3D directions by drawing a spinor map on a hemisphere. The hemisphere is a surface (Figure 5.3) in the ket space of spinor components *a* and *b*, eq. (5.3), which has 4 real dimensions. We reduce the original 4D spinor space to 3D by choosing a phase convention that makes *a*, the |*z*+> component, real and non-negative. We further reduce spinor space to 2D because we map only normalized spinors, so $|a|^2 + |b|^2 = 1$. Then the 2D surface of the "northern" hemisphere is a map of all normalized spinors (Figure 5.3, left and middle).



Map of Spinor Space

Top View of
Map of Spinor Space

Spinor Probability Ball

Figure 5.3 (Left and middle) Spinor space mapped onto a hemisphere.
(Right) Probabilities of measuring spin in various directions, given a particle in the state |*z*+>.

[This is similar to a Bloch sphere; any two-component discrete quantum state can be similarly mapped.]

Every possible spinor maps to exactly one point on the hemisphere, except the |*z*–> eigenspinor.

The north pole is the |*z*+> eigenspinor *a* = 1, *b* = 0:

$$\left|z+\right\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

The whole equator is the |*z*–> eigenspinor *a* = 0, |*b*| = 1. Because the |*z*+> component is zero at the equator, it no longer fixes the phase of the spinor; suddenly the arbitrary complex phase for quantum states is allowed again for |*z*–>. The line of 45° latitude is the set of spinors perpendicular to the *z*-axis in the *x*-*y* plane. Being midway between |*z*+> and |*z*–>, their *z*-axis measurements have 50% probability of |*z*+>, and thus also 50% probability of |*z*–>.

Consider the complex phase of *b* (the |*z*–> component). As the phase of *b* rotates around the complex plane, the spinor direction in 3-space rotates around the *z*-axis by the same angle (Figure 5.3 middle). Recall that the *x* and *y* eigenspinors are [Sak 1.4.17 p28]:

$$\left|x+\right\rangle = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, \quad \left|x-\right\rangle = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}; \qquad \left|y+\right\rangle = \begin{bmatrix} 1/\sqrt{2} \\ i/\sqrt{2} \end{bmatrix}, \quad \left|y-\right\rangle = \begin{bmatrix} 1/\sqrt{2} \\ -i/\sqrt{2} \end{bmatrix}.$$

I.e. (for a real-valued |*z*+> component), the |*x*+> eigenspinor has a |*z*–> component phase of 0; |*y*+> has a |*z*–> phase of 90°, |*x*–> has a |*z*–> phase of 180°, |*y*–> has a |*z*–> phase of 270°. This is a choice of convenience [Sak p27]. (Note that some references use a different convention, at least in some places.) This equality between the complex phase of the |*z*–> component and the real-space angle ϕ is true for all ϕ

(not just the axes). In general, the eigenspinor for a direction in space given by the spherical angles (θ, ϕ), is:

$$\left|\theta,\phi+\right\rangle=\begin{bmatrix}\cos(\theta/2)\\ \sin(\theta/2)e^{i\phi}\end{bmatrix} \qquad \text{[Sak prob 1.9 p61]},$$

where the equality of the real-space ϕ and the |z–> component complex angle ϕ is explicit. Be aware of other phases for the same spinor, such as:

$$\left|\theta,\phi+\right\rangle=\begin{bmatrix}\cos(\theta/2)e^{-i\phi/2}\\ \sin(\theta/2)e^{i\phi/2}\end{bmatrix} \qquad \text{[Sak 3.2.52 p168]},$$

which is the same as above, but multiplied by the unit-magnitude complex number $e^{-i\phi/2}$. In this latter spinor, the phase *difference* between the |z+> component and the |z–> component is still the real-space spherical angle ϕ.

Be careful to distinguish orthogonality (perpendicularity) in *real* space from orthogonality in *spinor* space. For example, in real space, the *x*-axis is perpendicular to the *z*-axis:

$$\mathbf{e}_x \perp \mathbf{e}_z, \text{ i.e.} \qquad \mathbf{e}_x\cdot\mathbf{e}_z=0,$$

but in spinor space, the kets

$$\left|x+\right\rangle \not\perp \left|z+\right\rangle, \qquad \text{i.e.} \qquad \left\langle x+|z+\right\rangle=\begin{bmatrix}1/\sqrt{2} & 1/\sqrt{2}\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix}=\frac{1}{\sqrt{2}}\neq 0.$$

Notice that rotation by a polar angle θ in real space corresponds to rotation by a polar angle $\theta/2$ in spinor space. That means that rotating a |z+> eigenstate by π radians in real space takes the north pole to the equator in spinor space:

$$\left|z+\right\rangle=\begin{pmatrix}\cos 0\\ \sin 0\, e^{i0}\end{pmatrix}=\begin{pmatrix}1\\0\end{pmatrix}, \qquad \hat{R}_y(\pi)\left|z+\right\rangle=\begin{pmatrix}\cos(\pi/2)\\ \sin(\pi/2)\, e^{i0}\end{pmatrix}=\begin{pmatrix}0\\1\end{pmatrix}=\left|z-\right\rangle.$$

Spin ½ Particles Are Really Spin √3/2: The spin of any one component (s_x, s_y, or s_z) of a spin-½ particle is ½ ħ, *but* |**s**| = $\frac{\sqrt{3}}{2}\hbar$. Recall that

$$|\mathbf{s}|=\sqrt{s_x^2+s_y^2+s_z^2},$$

and if each component is ± ½ ħ, the total magnitude is $\frac{\sqrt{3}}{2}\hbar$. This implies that s^2 = (¾) $\hbar^2$. So they're called "spin-½" particles, but the magnitude of their spin bigger than that. We can also see this from the well known eigenvalue of $\hat{s}^2$:

$$\hat{s}^2\begin{pmatrix}a\\b\end{pmatrix}=s(s+1)\hbar^2\begin{pmatrix}a\\b\end{pmatrix} \qquad \Rightarrow \qquad |\mathbf{s}|=\sqrt{\frac{1}{2}\cdot\frac{3}{2}}\hbar=\frac{\sqrt{3}}{2}\hbar.$$

Spinor Probabilities: Given a spinor, and a measurement direction along the unit vector **n**, what is the probability that the particle measures in that direction? We now show it is:

$$\Pr(\mathbf{n}+)=\cos^2\frac{\delta}{2}=\frac{1+\cos\delta}{2}=\frac{1+\mathbf{n}\cdot\mathbf{e_s}}{2}, \tag{5.4}$$

where $\delta=$ angle between spinor direction, $\mathbf{e_s}$ (in 3-space), and measurement direction, $\mathbf{n}$.

The spinor probability ball shows this relationship (Figure 5.3, right). Graphically, the probability equals the height of the tip of **n** projected onto the spin axis, as a fraction of the circle's *diameter*. When measuring "in-line" with the spinor, the probability of measuring "+" is 1 (certainty). When measured perpendicular to the spinor,

$$\Pr(\mathbf{n}+) = 0.5 \qquad \text{(perpendicular measurement)} .$$

We derive Pr(**n**+) in two ways: conceptually, and then from the eigenspinors.

(excerpted)

5.5 Coupling Spin and Position

So far, we have considered wave-functions (spatial states), and separately, spin states. An electron is a spin-1/2 particle, and so has both kinds of states simultaneously. The complete quantum state of an electron, or any spin-1/2 particle, must specify both parts: space and spin. In the simplest case, the space and spin parts are independent of each other. The full quantum state is simply the concatenation of space and spin states. There are no consistent standards for notation. If the spin state has a definite value of "up" or "down", we may write it as:

$$|\psi\rangle|\uparrow\rangle, \quad \text{or} \qquad |\psi\uparrow\rangle .$$

More general spin state can also be written many ways, such as:

$$\psi(x)\begin{pmatrix} a \\ b \end{pmatrix}, \qquad |\psi\rangle|\chi\rangle, \qquad |\psi,\chi\rangle .$$

Note that when the spin is independent of the wave-function, we just "tack on" the spinor to the wave-function to get the complete quantum state. This is a simple "tensor product," or "product state."

However, in many cases, the spin and wave-function depend on each other. E.g., in a Stern-Gerlach experiment (Figure 5.4), the particle approaching the device might be a simple superposition (or even mixture) of up and down. Then after the device, the wave-function would be large for the upper path and spin-up, and also large for the lower path and spin-down. However, the probability of finding the particle spin down in the upper path is (ideally) zero, as is the chance of finding it spin up in the lower path.

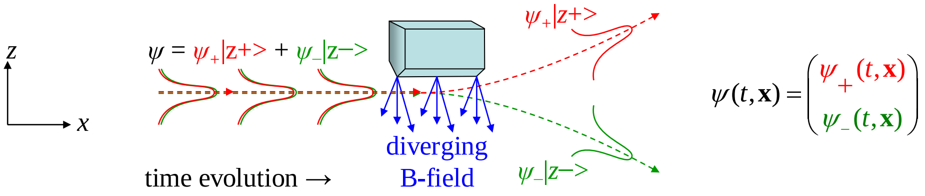


Figure 5.4 Time evolution of a spin-1/2 particle moving through a Stern-Gerlach device. The up and down spatial wave-functions are the same for *t* before the device, but different for *t* after.

(excerpted)

5.6 Angular Momentum for Arbitrary *j*

Angular momentum, which may be a combination of orbit and spin, is quantized by a quantum number *j*, which is allowed values of 0, ½, 1, 3/2, ... For such a *j*, in an angular momentum eigenstate:

$$J \equiv |\mathbf{J}| = \hbar\sqrt{j(j+1)} .$$

As $j \to \infty,\ J \to \hbar\left(j+\frac{1}{2}\right)$, because:

$$\sqrt{j(j+1)} = \sqrt{j^2+j} = j\sqrt{1+\frac{1}{j}} \approx j\left(1+\frac{1}{2j}\right) = j+\frac{1}{2} \qquad \text{(from the binomial theorem).}$$

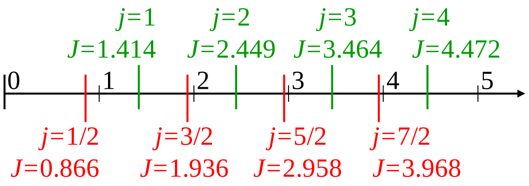


Figure 5.5 The magnitude of angular momentum, J, vs. the quantum number, j, in units of $\hbar$.

(excerpted)

5.7 Addition of Angular Momentum

Terminology: Given a particle of angular momentum quantum number j, its actual angular momentum magnitude is $J=\hbar\sqrt{j(j+1)}$. For brevity, we sometimes colloquially refer to the "angular momentum quantum number" as the "angular momentum," i.e., we may say "a particle of angular momentum j," when we really mean "a particle with angular momentum quantum number j."

Consider two subsystems (or particles) of angular momentum quantum numbers j_1 and j_2. Together, their total angular momentum quantum number lies somewhere between the sum $j_1 + j_2$ and the difference $| j_1 - j_2 |$, because angular momentum is a vector, and the two angular momenta can add constructively, destructively, or somewhere in between, depending on the angle between the two vector angular momenta.

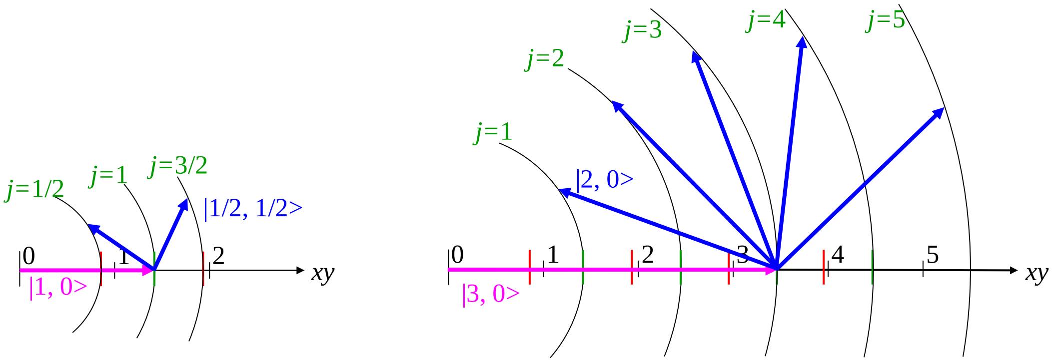


Figure 5.6 (Left) Addition of |1, 0> and |1/2, 1/2>. (Right) Addition of |3, 0> and |2, 0>.

Let's restate the above in terms of actual angular momentum, instead of angular momentum quantum numbers: Consider two particles of angular momentum $\hbar\sqrt{j_1(j_1+1)}$ and $\hbar\sqrt{j_2(j_2+1)}$. Together, their total angular momentum lies somewhere between $\hbar\sqrt{(j_1+j_2)(j_1+j_2+1)}$ and $\hbar\sqrt{(j_1-j_2)(j_1-j_2+1)}$.

For example, suppose $j_1 = ½$ and $j_2 = 1$ (Figure 5.6, left). Then $|J_1| = \sqrt{3}/2\ \hbar = 0.866\ \hbar$, and $|J_2| = \sqrt{2}\ \hbar = 1.414\ \hbar$. At most, the total $j_t = j_1 + j_2 = 3/2$, and the total angular momentum is $\sqrt{15}/2\ \hbar = 1.936\ \hbar$, which is somewhat less than $\sqrt{3}/2\ \hbar + \sqrt{2}\ \hbar = 2.280\ \hbar$. At the least, $j_t = j_2 - j_1 = ½$, and the total angular momentum is $\sqrt{3}/2\ \hbar = 0.866\ \hbar$, which is more than $\sqrt{2}\ \hbar - \sqrt{3}/2\ \hbar = 0.548\ \hbar$. We see that with $j_1 \neq j_2$, the two angular momenta can never fully reinforce, nor fully oppose, each other. This is because total angular momentum J_t is quantized by the same rules as individual particles, and that quantization does not allow for simply adding or subtracting the individual angular momenta. C'est la vie. Figure 5.6, right, shows a similar addition for |3 0> and |2 0>.

In Figure 5.6, notice that the addition of two angular momentum eigenstates does *not* (in general) add to a single total angular momentum eigenstate (however, they *do* add to a J_z eigenstate). Because the *x-y* components are *not* definite, the sum has some chance of being in several different total angular momentum states. In other words:

The sum of two *definite* angular momentum states may or may not be an eigenstate of total angular momentum, i.e., it may be a *superposition* of several *total* angular momentum states.

For example:

$$\begin{aligned} &\left|1,0\right\rangle\left|1/2,1/2\right\rangle = c_1\left|3/2,1/2\right\rangle + c_2\left|1/2,1/2\right\rangle, \\ &\left|3,0\right\rangle\left|2,0\right\rangle = c_5\left|5,0\right\rangle + c_4\left|4,0\right\rangle + c_3\left|3,0\right\rangle + c_2\left|2,0\right\rangle + c_1\left|1,0\right\rangle. \end{aligned} \qquad (5.5)$$

where the c's could potentially be complex. In fact, the c's are the famous Clebsch-Gordon coefficients, and will turn out to be real (described later).

5.7.1 Two Indefinites Can Make a Definite

Quantization of angular momentum leads to some curious consequences. We now show that two completely indefinite vectors can have a definite (real space) dot-product.

Recall that a $|j\ m>$ eigenstate has definite angular momentum, J^2, or equivalently $|J|$, and definite z-component of angular momentum, J_z. Therefore the magnitude of its x-y components in the x-y plane is definite (Figure 5.7):

$$J_x{}^2 + J_y{}^2 + J_z{}^2 = J^2 \qquad \Rightarrow \qquad \left|J_{xy}\right| = \sqrt{J_x{}^2 + J_y{}^2} = \sqrt{J^2 - J_z{}^2} = \hbar\sqrt{j\left(j+1\right) - m^2}\ .$$

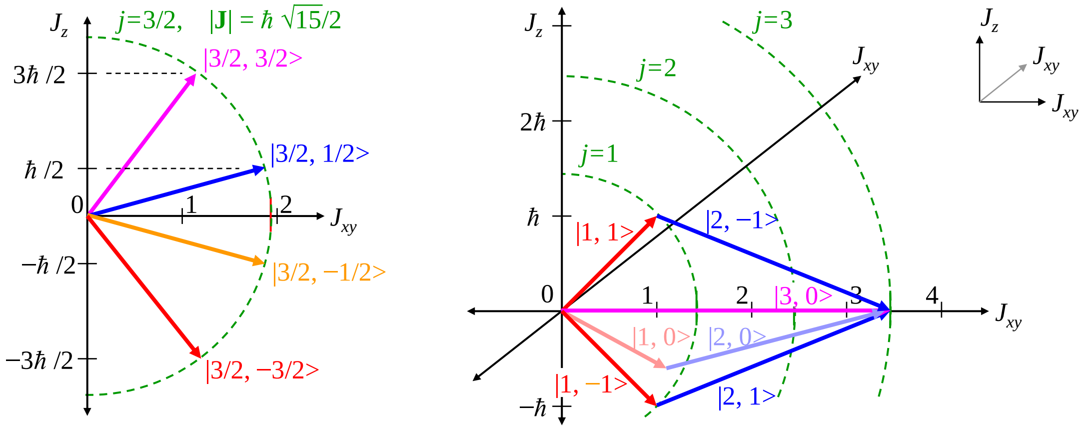
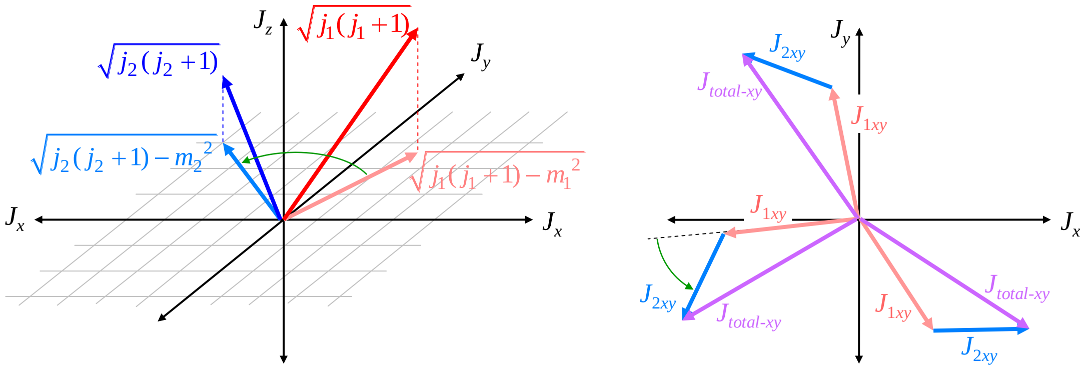


Figure 5.7 (Left) "Side view" of angular momentum states, and (right) perspective view of addition of $|1\ m_1>$ and $|2\ m_2>$ to make $|3\ 0>$. These vectors can all be rotated around the z-axis.

(excerpted)



Perspective view of
1 possible orientation

Top view of
3 possible orientations

Figure 5.8 (Left) Two angular momentum vectors, $\mathbf{j}_1$ and $\mathbf{j}_2$, with one possible orientation of their *x*-*y* components. (Right) Three possible orientations of the *x*-*y* components, with a fixed angle between J_{1xy} and J_{2xy}.

Thus the *x*-*y* orientations of $\mathbf{j}_1$ and $\mathbf{j}_2$ are individually completely uncertain, but the *relative* orientation is known.

The vector space of angular momentum states for the total (combined) system is different from the vector spaces of either system 1 or system 2. What, then, is this new, combined vector space? For that, we must understand tensor products.

5.7.2 Tensor Products

Oddly, you don't need to know anything about tensors to understand the tensor product of 2 vectors (kets). We'll start by describing a tensor product informally, then get more precise. Each of the two vectors exists in a vector space. (Vector spaces are described in 2.4.2 p59.) We'll assume the two original vector spaces are finite dimensional, though the results generalize straightforwardly to infinite dimensions. The two vector spaces are often structurally different, e.g., different dimensions.

We return to the earlier example of two subsystems of angular momentum, where $j_1 = ½$ and $j_2 = 1$. Considered separately, each subsystem has its own vector space of angular momentum:

System 1: dimension 2, |*j*, *m*> basis vectors: |½, +½> |½, –½ >.

System 2: dimension 3, |*j*, *m*> basis vectors: |1, +1> |1, 0> |1, –1> .

The tensor product *space* of these 2 vector spaces, $P = V_1 \otimes V_2$, has 2 x 3 = 6 dimensions, and its basis vectors are all 6 combinations of one basis vector from each of System 1 and System 2. In QM, we often write these pairs as just two kets, side by side:

|½, +½>|1, +1> |½, +½>|1, 0> |½, +½>|1, –1>

|½, –½>|1, +1> |½, –½>|1, 0> |½, –½>|1, –1> .

The scalar fields of V_1, V_2, and P are the complex numbers.

Generalizing: The **tensor product space** of 2 vector spaces, V_1 and V_2, is another vector space. The new vector space has dimension equal to the product of the dimensions of the 2 original vector spaces. The 2 original vector spaces must be over the same field of scalars, typically the complex numbers. The basis vectors of the new vector space are all the pairs of basis vectors you can form by taking one basis vector from V_1 and the other from V_2.

The **tensor product** of two *vectors* is the product of their scalar parts, and the juxtaposition of their vector parts. Sometimes, we write a tensor product explicitly with the ⊗ symbol. Some examples:

$$a|1/2,+1/2\rangle \otimes b|1,0\rangle = ab|1/2,+1/2\rangle|1,0\rangle$$

$$\left(a|1/2,+1/2\rangle + c|1/2,-1/2\rangle\right) \otimes b|1,0\rangle = ab|1/2,+1/2\rangle|1,0\rangle + cb|1/2,-1/2\rangle|1,0\rangle$$

$$\left(a|1/2,+1/2\rangle + c|1/2,-1/2\rangle\right) \otimes \left(b|1,0\rangle + d|1,1\rangle\right)$$

$$= ab|1/2,+1/2\rangle|1,0\rangle + cb|1/2,-1/2\rangle|1,0\rangle + ad|1/2,+1/2\rangle|1,1\rangle + cd|1/2,-1/2\rangle|1,1\rangle .$$

The tensor-product kets represent possible results of measuring the state of each constituent subsystem separately. Thus:

The tensor product of two vectors gives a *single* complex-amplitude
for each possible *set* of experimental results.

You can see that the scalar fields of the two vector spaces must be the same, because we multiply the scalar coefficients together. However, the vector parts may be completely unrelated, coming from unrelated vector spaces. Unlike the scalar parts, they cannot be "multiplied" or otherwise combined in any way, so we simply write the basis vector parts of each factor next to each other: $|b_1>|b_2>$, or $|b_1, b_2>$.

There are many notations for tensor products. Often, for example, the j_1 and j_2 are well known and understood. Then it is cumbersome to write them in all the kets and bras (as you can see above). So references frequently omit them, and keep only the *m* values. Thus we define:

$$|1/2, +1/2\rangle \equiv |+1/2\rangle, \qquad |1/2, -1/2\rangle \equiv |-1/2\rangle$$

$$|1, 1\rangle \equiv |1\rangle, \qquad |1, 0\rangle \equiv |0\rangle, \qquad |1, -1\rangle \equiv |-1\rangle .$$

Then:

$$|+1/2\rangle \otimes |1\rangle = |+1/2\rangle|1\rangle \equiv |+1/2, 1\rangle \qquad \text{i.e.,} \qquad |m_1\rangle \otimes |m_2\rangle = |m_1\rangle|m_2\rangle \equiv |m_1, m_2\rangle .$$

Space and spin: We've already seen another common example of a tensor product: the combination of the space and spin parts of a particle. For example, an electron has a spatial wave function, ψ, and a spinor, χ. Their tensor product space spans the complete quantum state space which describes all properties of the electron. If the space part is independent of spin, then we can write the complete description as the tensor product of a spatial state and a spin state ($\psi(x)$ and $\chi = [a \;\; b]^T$). Then:

$$|electron\rangle = \psi(x) \otimes \chi = \psi(x) \otimes \left(a|z+\rangle + b|z-\rangle\right) \equiv \psi(x)\begin{bmatrix} a \\ b \end{bmatrix}.$$

If the space part depends on spin, as in a Stern-Gerlach experiment where the spin determines the path in space, then there are two wave functions: one for spin up (ψ_+), and another for spin down (ψ_-). The resulting state *cannot* be written as a tensor product of two vectors. It can only be written as a sum of such tensor products:

$$|electron\rangle = \psi_+(x) \otimes |z+\rangle + \psi_-(x) \otimes |z-\rangle = \begin{bmatrix} \psi_+(x) \\ \psi_-(x) \end{bmatrix}.$$

5.7.3 Operators on Tensor Products

(excerpted)

5.8 Just a Moment: the Landé g-Factor

The Landé g-factor is essential for understanding atomic spectra, and clarifies LSJ angular momentum coupling, which is essential for atomic electron structure. It is well known that both orbital and spin angular momentum of a charged particle contribute magnetic dipole moments to a system. (The spin-related magnetic moment is sometimes called the "intrinsic" magnetic moment.) However, there is a *quantitative* difference between orbit and spin in the ratio of magnetic dipole moment to angular momentum (sometimes called the **g-factor**). For a $|l\ m_l>$ orbital state, and a $|s\ m_s>$ spin state, experiment shows the magnetic moment is (in gaussian units):

$$\mu_{z,L} = g_L m_l \mu, \qquad\qquad \mu_{z,s} = g_s m_s \mu$$

$$\text{where} \quad \mu \equiv \frac{|e|\hbar}{2(mass)c} \equiv \text{magneton}$$

$$e \equiv \text{particle charge}$$

$$g_L \equiv \text{orbital } g\text{-factor } (= -1 \text{ for electron})$$

$$g_s \equiv \text{spin } g\text{-factor } (\approx -2 \text{ for electron}) .$$

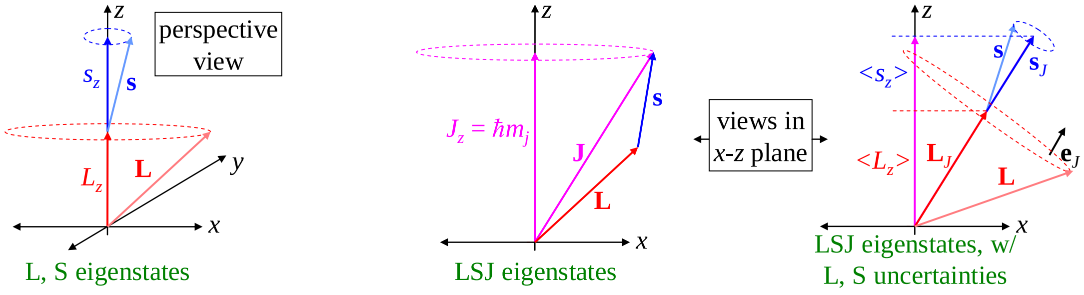
These results are well understood theoretically, as well [Sak2 p79]. The **magneton** is a function of the particle's mass and charge, and is therefore different for different particles. We follow the US National Institute of Standards and Technology (NIST) terminology and sign conventions, which defines the **Bohr magneton** as the magneton for an electron, and the **nuclear magneton** as the magneton for a proton. Both magnetons are defined positive. The magnetic moment vector for an electron is *opposite* its spin and orbital angular momentum, because it is negatively charged, so its g-factors are negative. The intrinsic magnetic moment for the neutron is also opposite its spin (for no good reason), so its g-factor is also negative. Furthermore, the neutron g-factor is calculated using the proton mass and charge, i.e. the afore-mentioned nuclear magneton.

(Note that some references call the nuclear magneton a "Bohr" magneton. Some references define all g-factors as positive, and put minus signs in "by hand," relying on the "understanding" of the correct sign, but that makes the neutron g-factor confusing.)

Because the particle mass appears in the denominator of the magneton formula, the nuclear magneton is about 3 orders of magnitude smaller than the Bohr (electron) magneton. This is why we usually ignore the magnetic moment of atomic nuclei: the mass is so large that μ is negligible.

Since the g-factor for spin is different than that for orbit, we must ask: what if our angular momentum eigenstate is a combination of spin and orbit? There are two important cases to consider: L-s eigenstates, and LSJ eigenstates.

L-s eigenstates: Consider two angular momenta which add, say **L** and **s**. In the *uncoupled* basis, the system is in the state $|l, m_l; s, m_s>$, which is an eigenstate of $\hat{L}^2, \hat{L}_z, \hat{s}^2, \hat{s}_z$, and $\hat{J}_z$, but not $\hat{J}^2$ (as we learned from our study of Clebsch-Gordon coefficients: $j_1 \to l$, $j_2 \to s$). (Note that $m_j = m_l + m_s$, so is not an independent quantum number.) Figure 5.9 (left) shows these angular momenta, and their uncertainties. The magnitudes of the L and s momenta are definite (no uncertainty), as are the z-components. The x-y components are (1) uncertain, (2) uniformly distributed in the x-y plane, and (3) average to zero.



L, S eigenstates

LSJ eigenstates

LSJ eigenstates, w/ L, S uncertainties

Figure 5.9 (Left) L-s angular momenta eigenstates, with definite z-components. (Middle) LSJ angular momenta eigenstates, with definite components parallel to **J**. (Right) Uncertainties and averages for middle diagram.

(excerpted)

6 Multi-Particle Quantum Mechanics

Quantum mechanics covers not just single particles, but systems of particles. Systems of particles exhibit further unexpected behaviors which explain atoms and molecules, enable new technologies (e.g. quantum cryptography), and require dedicated study. This chapter requires understanding single particle QM of waves and spins. We proceed along these topics:

1. Multi-particle wave-functions: distinct particles.
2. Multi-particle wave-functions: identical particles.
3. Multi-particle Schrödinger equation.
4. Multi-particle states with spin.
5. Symmetry of multi-particle wave-functions: fermions and bosons.
6. Atoms: Hund's rules.
7. Multi-particle entanglement

Many new phenomena arise in multi-particle systems. We will examine some of these, such as entanglement creating nonlocal behavior.

6.1.1 Multi-Particle Wave-Functions: Distinct Particles

Suppose we have a system of one proton and one electron. It is described by a multi-particle wave-function, $\psi(\mathbf{r}_p, \mathbf{r}_e)$, which defines the joint probability density of finding a proton at $\mathbf{r}_p$, *and* an electron at $\mathbf{r}_e$:

$$\Pr(\text{proton around } \mathbf{r}_p \text{ and electron around } \mathbf{r}_e) = \left|\psi(\mathbf{r}_p, \mathbf{r}_e)\right|^2 d^3r_p \, d^3r_e \, .$$

Because the proton and electron may interact, their positions depend on each other. Therefore, ψ is a single wave-function which takes account of their interaction, and covers all possible combinations of positions of both the proton and electron. If we take a measurement, we must find the two particles *somewhere*, so our normalization is

$$\iint_\infty \left|\psi(\mathbf{r}_p, \mathbf{r}_e)\right|^2 d^3r_p \, d^3r_e = 1 \qquad \text{(a 6 dimensional integral)} \, .$$

If we have two particles that don't interact, then their joint wave-function can be written simply as the product of the two individual wave-functions:

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2) \, .$$

This is a kind of tensor product. The wave-functions are normalized separately according to:

$$\int_\infty \left|\psi_1(\mathbf{r})\right|^2 d^3r = 1, \qquad \text{and} \qquad \int_\infty \left|\psi_2(\mathbf{r})\right|^2 d^3r = 1 \, ,$$

so that:

$$\iint_\infty \left|\psi(\mathbf{r}_1, \mathbf{r}_2)\right|^2 d^3r_1 \, d^3r_2 = \int_\infty \left|\psi_1(\mathbf{r})\right|^2 d^3r \int_\infty \left|\psi_2(\mathbf{r})\right|^2 d^3r = 1 \, .$$

In other words, when the 2-particle wave-function separates, the individual normalizations insure the joint wave-function is normalized.

6.1.2 Multi-Particle Wave-Functions: Identical Particles

We may have a system of two identical particles, such as two electrons. Like the proton-electron system, this system is also described by a 2-particle wave-function, $\psi(\mathbf{r}_1, \mathbf{r}_2)$. This is the joint amplitude for finding *an* electron at $\mathbf{r}_1$ *and another* electron at $\mathbf{r}_2$. It is a crucial point that:

Identical particles cannot be distinguished from each other. There is no concept of particle #1 and particle #2. There is only the concept of a particle at $\mathbf{r}_1$ and an identical one at $\mathbf{r}_2$.

Since the two particles are identical (indistinguishable), it must be the case that:

$$\Pr(\text{particle at } \mathbf{r}_1 \text{ and another at } \mathbf{r}_2) = \Pr(\text{particle at } \mathbf{r}_2 \text{ and another at } \mathbf{r}_1)$$

$$\Rightarrow \qquad \left|\psi(\mathbf{r}_1,\mathbf{r}_2)\right| = \left|\psi(\mathbf{r}_2,\mathbf{r}_1)\right| .$$

This is a symmetry requirement on the wave-function, and requires that the wave-function can differ by at most a phase under an $\mathbf{r}_1 \leftrightarrow \mathbf{r}_2$ interchange. Furthermore, if we swap particles 1 and 2 a second time, we must get the original 2-particle wave-function back. We can write **particle exchange** as a linear operator, defined by:

$$\hat{P}_{12}\psi(\mathbf{r}_1,\mathbf{r}_2) \equiv \psi(\mathbf{r}_2,\mathbf{r}_1) .$$

By definition, it follows that:

$$\left(\hat{P}_{12}\right)^2 \psi(\mathbf{r}_1,\mathbf{r}_2) = \psi(\mathbf{r}_1,\mathbf{r}_2) .$$

This implies that any eigenvalue of $\hat{P}_{12}$, call it λ, must satisfy $\lambda^2 = 1$. Therefore, the only eigenvalues are $\lambda = \pm 1$. A function is **symmetric** iff $\psi(\mathbf{r}_1, \mathbf{r}_2) = \psi(\mathbf{r}_2, \mathbf{r}_1)$. It is an eigenstate of $\hat{P}_{12}$ with eigenvalue +1. A function is **antisymmetric** iff $\psi(\mathbf{r}_1, \mathbf{r}_2) = -\psi(\mathbf{r}_2, \mathbf{r}_1)$. It is an eigenstate of $\hat{P}_{12}$ with eigenvalue –1. Some functions are neither symmetric nor antisymmetric.

$\hat{P}_{12}$ also operates on spin states, in a similar way: it swaps the states for the particles:

$$\hat{P}_{12}\chi_{12} \equiv \chi_{21}, \qquad \text{e.g.,} \qquad \hat{P}_{12}\left|\uparrow\downarrow\right\rangle = \left|\downarrow\uparrow\right\rangle .$$

We define here the notation where χ_{21} is the spin-state χ_{12} with the particle spins exchanged. E.g.,

$$\chi_{12} = \left|\uparrow_1 \downarrow_2\right\rangle \qquad \Rightarrow \qquad \chi_{21} = \left|\downarrow_1 \uparrow_2\right\rangle .$$

Putting space and spin together, we have:

$$\hat{P}_{12}\left[f(\mathbf{r}_1,\mathbf{r}_2)\chi_{12}\right] = f(\mathbf{r}_2,\mathbf{r}_1)\chi_{21}, \qquad \text{e.g.,} \qquad \hat{P}_{12}\psi(\mathbf{r}_1,\mathbf{r}_2)\left|\uparrow\downarrow\right\rangle = \psi(\mathbf{r}_2,\mathbf{r}_1)\left|\downarrow\uparrow\right\rangle .$$

Note that $\hat{P}_{12}$ operates on *both* the spatial state and the spin state.

6.1.3 Multi-Particle Schrödinger Equation

We extend the Schrödinger equation to multiple-particles in a simple way: just add up the energies of all the particles:

2-particle
$$\underbrace{\frac{\hbar}{i}\frac{\partial}{\partial t}\psi(\mathbf{r}_1,\mathbf{r}_2,t)}_{\text{total energy}} = \underbrace{\sum_{j=1}^{2}\frac{-\hbar^2}{2m_j}\nabla_j^{\,2}\psi(\mathbf{r}_1,\mathbf{r}_2,t)}_{\text{total kinetic energy}} + \underbrace{V(\mathbf{r}_1,\mathbf{r}_2)\psi(\mathbf{r}_1,\mathbf{r}_2,t)}_{\substack{\text{potential energy,}\\ \text{including particle interactions}}} ,$$

n-particle
$$\underbrace{\frac{\hbar}{i}\frac{\partial}{\partial t}\psi(\mathbf{r}_1,\mathbf{r}_2,\ldots\mathbf{r}_n,t)}_{\text{total energy}} = \underbrace{\sum_{j=1}^{n}\frac{-\hbar^2}{2m_j}\nabla_j^{\,2}\psi(\mathbf{r}_1,\mathbf{r}_2,\ldots\mathbf{r}_n,t)}_{\text{total kinetic energy}} + \underbrace{V(\mathbf{r}_1,\mathbf{r}_2,\ldots\mathbf{r}_n)\psi(\mathbf{r}_1,\mathbf{r}_2,\ldots\mathbf{r}_n,t)}_{\substack{\text{potential energy,}\\ \text{including particle interactions}}}$$

where $\mathbf{r}_j \equiv (x_j, y_j, z_j),$

$$\nabla_j^{\,2} \equiv \nabla_j\cdot\nabla_j = \left(\frac{\partial}{\partial x_j}\mathbf{e}_x + \frac{\partial}{\partial y_j}\mathbf{e}_y + \frac{\partial}{\partial z_j}\mathbf{e}_z\right)\cdot\left(\frac{\partial}{\partial x_j}\mathbf{e}_x + \frac{\partial}{\partial y_j}\mathbf{e}_y + \frac{\partial}{\partial z_j}\mathbf{e}_z\right) = \frac{\partial^2}{\partial x_j^{\,2}} + \frac{\partial^2}{\partial y_j^{\,2}} + \frac{\partial^2}{\partial z_j^{\,2}} .$$

6.1.4 Multi-Particle States With Spin

Recall that particles with spin have a state which comprises both a spatial wave-function *and* a spin state. Electrons are spin ½, and can therefore be either spin-up or spin-down (conventionally along the *z*-axis):

$$\left|z+\right\rangle = \begin{bmatrix}1\\0\end{bmatrix} \equiv \left|\uparrow\right\rangle \qquad\qquad \left|z-\right\rangle = \begin{bmatrix}0\\1\end{bmatrix} \equiv \left|\downarrow\right\rangle .$$

The total quantum state is a combination (tensor product) of wave-function and spin-state, written as:

$$\psi(\mathbf{r})\chi, \qquad \textit{where} \quad \chi \text{ is any superposition of } \left|\uparrow\right\rangle \text{ and } \left|\downarrow\right\rangle,$$

or a superposition of such states. For example, an excited hydrogen electron in a 2*p* state, with $n = 2$, $l = 1$, $m = 1$, spin up, is in the state

$$\psi_{211}(\mathbf{r})\left|\uparrow\right\rangle = R_{21}(r)Y_{11}(\theta,\phi)\left|\uparrow\right\rangle .$$

Two-particle states with spin have two-particle wave-functions, and two-particle spin-states, e.g.

$$\psi(\mathbf{r}_1,\mathbf{r}_2)\chi_{12}, \qquad \textit{where} \quad \chi_{12} \text{ is } \left|\uparrow\uparrow\right\rangle, \left|\uparrow\downarrow\right\rangle, \left|\downarrow\uparrow\right\rangle, \left|\downarrow\downarrow\right\rangle,$$

or a superposition of such states. It is easy to show that any quantum state can be written as a sum of a symmetric state and an antisymmetric state.

6.1.5 Symmetry Of Multi-Particle Wave-Functions: Fermions and Bosons

As mentioned above, multi-particle wave-functions are subject to symmetry restrictions. Experimentally, there are two kinds of particles: fermions and bosons. Fermions are half-odd-integer spin, and bosons are integer spin. It turns out that:

> Multi-particle fermion states must be *anti*symmetric under interchange of any two identical fermions, and boson states must be *symmetric* under interchange of any two identical bosons.

[These symmetry requirements are proven by the spin-statistics theorem in QFT.] Non-identical fermions, such as an electron and a muon, are distinguishable and therefore have no symmetry requirement: the probability of finding the electron at $\mathbf{r}_1$ and the muon at $\mathbf{r}_2$ may be different than that of finding the muon at $\mathbf{r}_1$ and the electron at $\mathbf{r}_2$.

However, the symmetry requirement includes both the wave-function (i.e., spatial state) *and* the spin states: it is the *entire* state, wave-function with spin-state, that must be antisymmetric for fermions. *If both particles have the same spin states* (say |↑↑>), or if the spin-states are otherwise symmetric (say |↑↓> + |↓↑>), the symmetry requirement falls only to the wave-function, as follows:

$$\left.\begin{array}{ll}\text{Fermions:} & \psi(\mathbf{r}_1,\mathbf{r}_2)=-\psi(\mathbf{r}_1,\mathbf{r}_2) \\ \text{Bosons:} & \psi(\mathbf{r}_1,\mathbf{r}_2)=\psi(\mathbf{r}_1,\mathbf{r}_2)\end{array}\right\}\ \text{(symmetric spin-states)} .$$

Fermions: The antisymmetry of fermion states includes the **Pauli exclusion principle**. Two electrons cannot be in the same single-particle quantum state, because then (for symmetric spin):

$$\psi(\mathbf{r}_1,\mathbf{r}_2)=\psi(\mathbf{r}_1)\psi(\mathbf{r}_2)=\psi(\mathbf{r}_2,\mathbf{r}_1) \qquad \text{in violation of} \qquad \psi(\mathbf{r}_1,\mathbf{r}_2)=-\psi(\mathbf{r}_2,\mathbf{r}_1) .$$

In fact, for $\mathbf{r}_1 = \mathbf{r}_2 = \mathbf{r}$ and symmetric spins, an antisymmetric state must have:

$$\psi(\mathbf{r},\mathbf{r})=-\psi(\mathbf{r},\mathbf{r}) \qquad \Rightarrow \qquad \psi(\mathbf{r},\mathbf{r})=0, \qquad \forall\, \mathbf{r} .$$

This means that there is zero probability of measuring two identical fermions at the same place. For *n*-particle systems, with fully symmetric spin states, there is zero probability of measuring *any* two identical fermions at the same place.

In contrast to the above case of identical spins, if spin- and space- states separate, this antisymmetry can be either:

$$\begin{array}{llll}\psi(\mathbf{r}_1,\mathbf{r}_2)=-\psi(\mathbf{r}_2,\mathbf{r}_1) & \text{and} & \chi_{12}=\chi_{21}, & \text{or} \\ \psi(\mathbf{r}_1,\mathbf{r}_2)=\psi(\mathbf{r}_2,\mathbf{r}_1) & \text{and} & \chi_{12}=-\chi_{21} . & \end{array}$$

In words, to make the total fermion state antisymmetric (for separable spin/space states), we must have either the wave-function antisymmetric and the spin-state symmetric, or the wave-function symmetric and the spin-states antisymmetric.

But how can a spin-state be antisymmetric under particle exchange, when it is not a function of position? There is no $\mathbf{r}_1$ and $\mathbf{r}_2$. A spin-state can only be antisymmetric by being an antisymmetric superposition of two spin-states, e.g. (ignoring normalization):

$$\chi_{12}=\left|\uparrow\downarrow\right\rangle-\left|\downarrow\uparrow\right\rangle \qquad \Rightarrow \qquad \chi_{12}=-\chi_{21}, \text{ i.e. } \left|\uparrow\downarrow\right\rangle-\left|\downarrow\uparrow\right\rangle=-\left(\left|\downarrow\uparrow\right\rangle-\left|\uparrow\downarrow\right\rangle\right)=\left|\uparrow\downarrow\right\rangle-\left|\downarrow\uparrow\right\rangle ,$$

where on the right-hand-side we have interchanged particles 1 and 2, which interchanges their spins.

Symmetric spin states can be either *z* eigenstates, or symmetric superpositions, such as:

$$\chi_{12}=\left|\uparrow\uparrow\right\rangle, \qquad \chi_{12}=\left|\downarrow\downarrow\right\rangle, \qquad \chi_{12}=\left|\uparrow\uparrow\right\rangle-\left|\downarrow\downarrow\right\rangle, \qquad \chi_{12}=\left|\uparrow\downarrow\right\rangle+\left|\downarrow\uparrow\right\rangle .$$

Note that *any* superposition of symmetric states is symmetric,
even if some of the coefficients are negative or complex.

In the most general case, the spin-states are entangled with the wave-function, and the antisymmetric requirement for fermions requires swapping both spatial positions ($\mathbf{r}_1$ and $\mathbf{r}_2$) and spin state labels. In other words, the total antisymmetric state may be any superposition of antisymmetric states.

The requirement for antisymmetric multi-particle fermion states can make calculating such states difficult. Sometimes, you may have a solution to the multi-particle Schrödinger equation, but which does not have the proper symmetry for the multiple particles. In such a case, you can "symmetrize" or "antisymmetrize" the solution to construct a valid solution with the proper symmetry. Using our particle exchange operator, and given a solution of arbitrary symmetry $f(\mathbf{r}_1, \mathbf{r}_2)\chi_{12}$, we can symmetrize and antisymmetrize that solution as follows:

$$\begin{array}{ll}\text{antisymmetric:} & \psi(\mathbf{r}_1,\mathbf{r}_2)\chi_{12}=f(\mathbf{r}_1,\mathbf{r}_2)\chi_{12}-\hat{P}_{12}\left[f(\mathbf{r}_1,\mathbf{r}_2)\chi_{12}\right]=f(\mathbf{r}_1,\mathbf{r}_2)\chi_{12}-f(\mathbf{r}_2,\mathbf{r}_1)\chi_{21} \\ \text{symmetric:} & \psi(\mathbf{r}_1,\mathbf{r}_2)\chi_{12}=f(\mathbf{r}_1,\mathbf{r}_2)\chi_{12}+\hat{P}_{12}\left[f(\mathbf{r}_1,\mathbf{r}_2)\chi_{12}\right]=f(\mathbf{r}_1,\mathbf{r}_2)\chi_{12}+f(\mathbf{r}_2,\mathbf{r}_1)\chi_{21} .\end{array}$$

Note that $\hat{P}_{12}$ operates on *both* the spatial state and the spin state. For example, suppose $f(\mathbf{r}_1, \mathbf{r}_2)$ and χ_{12} are neither symmetric nor antisymmetric, e.g., $\chi_{12} = |\uparrow\uparrow\rangle + |\uparrow\downarrow\rangle$. Then antisymmetrizing the wave-function gives:

$$\begin{aligned}\psi(\mathbf{r}_1,\mathbf{r}_2)\chi_{12} &= f(\mathbf{r}_1,\mathbf{r}_2)\left|\uparrow\uparrow\right\rangle + f(\mathbf{r}_1,\mathbf{r}_2)\left|\uparrow\downarrow\right\rangle - f(\mathbf{r}_2,\mathbf{r}_1)\left|\uparrow\uparrow\right\rangle - f(\mathbf{r}_2,\mathbf{r}_1)\left|\downarrow\uparrow\right\rangle \\ &= \left[f(\mathbf{r}_1,\mathbf{r}_2) - f(\mathbf{r}_2,\mathbf{r}_1)\right]\left|\uparrow\uparrow\right\rangle + \left[f(\mathbf{r}_1,\mathbf{r}_2)\left|\uparrow\downarrow\right\rangle - f(\mathbf{r}_2,\mathbf{r}_1)\left|\downarrow\uparrow\right\rangle\right].\end{aligned}$$

The first term is spatially antisymmetric and spin symmetric. The 2nd term does not separate into space and spin, but it is still antisymmetric:

$$\begin{aligned} f(\mathbf{r}_1,\mathbf{r}_2)\left|\uparrow\downarrow\right\rangle - f(\mathbf{r}_2,\mathbf{r}_1)\left|\downarrow\uparrow\right\rangle &\overset{?}{=} -\left[f(\mathbf{r}_2,\mathbf{r}_1)\left|\downarrow\uparrow\right\rangle - f(\mathbf{r}_1,\mathbf{r}_2)\left|\uparrow\downarrow\right\rangle\right] \\ &\overset{!}{=} f(\mathbf{r}_1,\mathbf{r}_2)\left|\uparrow\downarrow\right\rangle - f(\mathbf{r}_2,\mathbf{r}_1)\left|\downarrow\uparrow\right\rangle .\end{aligned} \qquad (6.1)$$

Thus we see that superpositions of antisymmetric states are also antisymmetric, regardless of whether space and spin parts separate. Because antisymmetrization removes any symmetric part of a function, antisymmetrized wave-functions generally must be normalized after antisymmetrization.

Any number of fermions can have a fully symmetric spin state, but it is impossible to form a fully antisymmetric spin-state for 3 or more fermions [Bay p457].

Bosons: Bosons have multi-particle states that are *symmetric* under interchange of any two particles, and therefore do not have Pauli exclusion restrictions. Two bosons *can* be in the same quantum state. In fact, *n* bosons can all be in the same quantum state. This is a critical property with many implications, from stimulated radiation (lasers) to superfluids.

We return to anti/symmetrization in the discussion of atoms below, and also extend it to 3 or more identical particles.

6.2 Atoms

Atoms are multi-particle systems comprising a nucleus, and one or more electrons. Understanding the electronic structure of atoms is both practical, and pedagogically important. Many of the concepts involved in atoms are readily applicable to other situations. Also, there is much confusion and many inaccurate statements about atoms in non-peer-reviewed sources; it is important to set the record straight.

This section requires that you understand spherical harmonics, the hydrogen atom, its electronic orbitals, and the *n*, *l*, and *m* quantum numbers. It also requires understanding that electrons can have two different spin states.

Note that the quantum number *l* is often called the "azimuthal" quantum number, even though *m*, not *l*, actually describes the electron's azimuthal motion. This term for *l* seems to be an anachronism from a time before atomic structure was understood [E&R p115]. However, [Bay p156m] uses "azimuthal quantum number" for *m*, the *z*-component of angular momentum, which is consistent with the term "azimuthal," which means "along the horizon." Because of this confusion, we do not use the term "azimuthal quantum number."

6.2.1 Forget the Bohr Model

Bohr was a great physicist, and did well with the information available to him. However, the Bohr model of the atom is completely wrong, and probably shouldn't be taught any more. It does manage to compute the coarse electron energies for hydrogen, but for the wrong reasons. The idea of the electron wave-function of a 3D atom wrapping around in a flat 2D circle is unphysical (though for $l \geq 1$, this idea captures the ϕ dependence of the wave-function). The Bohr model predicts the wrong values for angular momentum, because it excludes $l = 0$. It therefore calls the ground state $l = 1$, instead of $l = 0$. The fact that the coarse energies of hydrogen work out is just (bad) luck.

6.2.2 Why Spherical Harmonics Are *Not* Spherically Symmetric

As prelude to atoms, we consider the hydrogen-like orbitals. Why are spherical harmonics not spherically symmetric? Short answer: Because we're also in an $\hat{L}_z$ eigenstate, and choosing the *z*-axis breaks the spherical symmetry. But let's clarify.

Electron orbital shapes are frequently drawn in physics and chemistry books: the familiar sphere for s-orbitals, the infamous dumb-bell for p-orbitals, the mysterious dumb-bell through a donut for d-orbitals, etc (Figure 6.1).

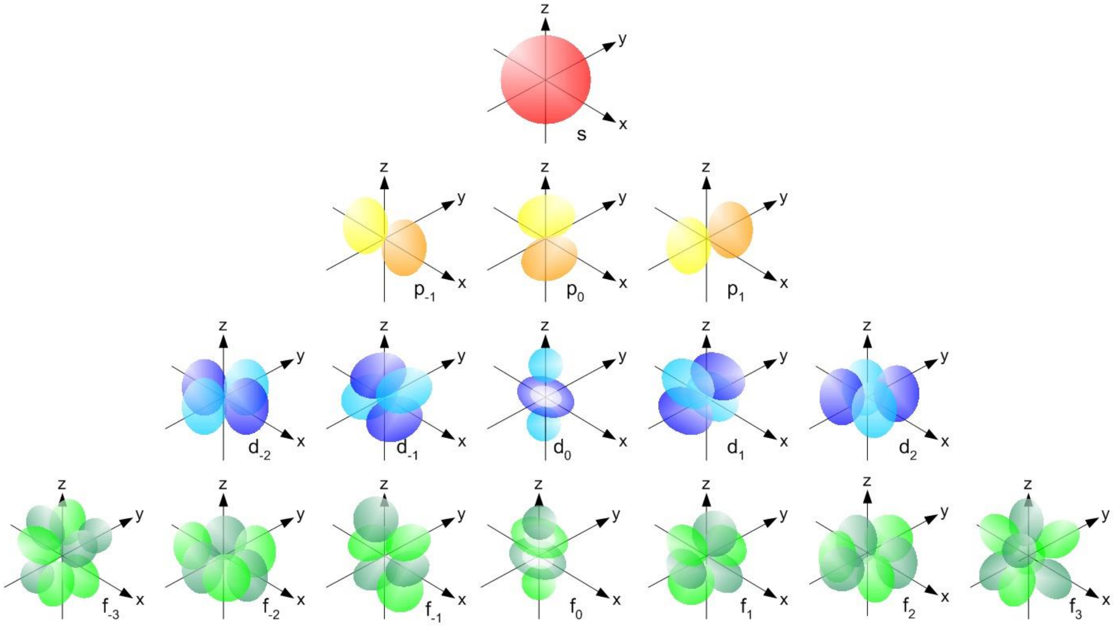


Figure 6.1 The chemist's views of atomic orbitals are linear combinations of the spherical harmonics. [http://upload.wikimedia.org/wikipedia/commons/4/4a/Single_electron_orbitals.jpg]

(excerpted)

6.3 Ground States of Small Atoms, and Hund's Rules

Understanding the electronic structure of small to medium atoms requires understanding antisymmetry of electron (fermion) states, and understanding Hund's rules. These concepts provide good insight into quantum mechanical principles, making atoms worth study for such insight alone. Hund's rules are frequently misunderstood and misapplied, so it is also helpful to review some common *invalid* reasoning, and why it's wrong. Antisymmetry and Hund's rules are sometimes presented as simple, and glossed over. In fact, they are somewhat involved, but manageable. [Bay] and [E&R] are good references. [Bay p452] has a complete periodic table with both electron configurations *and* their spectral term ground states. [Gos p426] and [Gas p307] have the same information in list form. We here discuss the ground states of atoms.

In addition to the prerequisites for the previous section (the hydrogen atom, spherical harmonics), this section requires you understand quantum addition of angular momentum, combined space/spin states, antisymmetry of multi-electron states, and that you have been introduced to the fine structure and spin-orbit energy. We follow these topics:

1. Review of shell structure of atomic electrons, Hartree and Hartree-Fock approximations.
2. Angular momentum coupling of atomic electrons: S, L, and J.
3. Fermion antisymmetry.
4. Hund's rules, by rote.
5. Examples of Hund's rules for interesting atoms.
6. Why do Hund's rules work?

Our entire discussion is done within the Hartree-Fock approximation and perturbation theory. Sometimes, we speak loosely and say something like “The total angular momentum is *J*,” by which we mean “The total angular momentum quantum number is *J*.” Bold symbols indicate the actual angular momentum vector, such as **L**, **S**, or **J**. Squared vectors are defined as magnitude squared, e.g. $\mathbf{L}^2 \equiv \mathbf{L}\cdot\mathbf{L} = |\mathbf{L}|^2$.

In math and physics, the terms “symmetric” and “antisymmetric” have many meanings. In general, a symmetry is some invariant property under a given transformation. You are familiar with a function of one variable being symmetric under reflection about zero (aka an “even” function), or antisymmetric about zero (an “odd” function). In multi-particle wave-functions, we refer to anti/symmetry under particle interchange:

6.3.1 Introduction to Atomic Electron Structure

There are three major factors that determine the coarse and fine structure of atoms (energies and number of Zeeman splittings). They are, in order of significance:

1. The Coulomb energy binding the electron to the nucleus, giving a hydrogen-like shell structure. We take these states as our unperturbed basis.
2. The Coulomb energy between electrons in the outermost subshell, which drives toward antisymmetric spatial states of multi-electron quantum states.
3. The electron spin-orbit coupling to the nuclear charge (responsible for the fine-structure of atoms).

The energy separation of subshells is typically a few eV; that of different multi-electron states within a subshell is of order 0.1 eV, and the spin-orbit energies are roughly 1-10 meV [Bay p455]. The observable properties are determined primarily by the electrons in unfilled subshells, and the nuclear charge. (We do not consider the hyperfine structure, which involves nuclear spin interactions.) The following shell/subshell/orbital structure of atoms is evident in the hydrogen atom, but applies to heavier atoms, as well:

shell	principal quantum number *n*, can hold $2n^2$ electrons.
subshell	a principal and orbital angular momentum quantum number pair, *nl*, can hold $2(2l + 1)$ electrons.
orbital	spatial state with principal, orbital angular momentum, and orbital *z*-component angular momentum (aka magnetic) quantum numbers *nlm*; can hold 2 electrons.

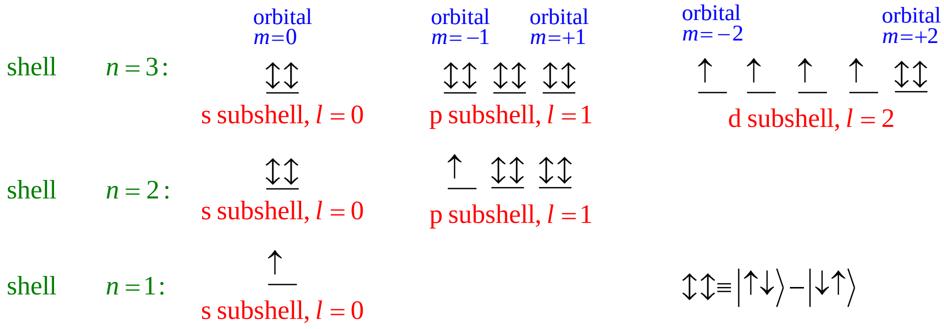


Figure 6.2 Example shells, subshells, and orbitals. The notation “↕↕” is shorthand for “|↑↓> – |↓↑>”.

(Note that [Gos] uses the term “orbital” where most use the term “subshell.” [Bay] and [C&S] use the term “shell” where most use “subshell,” and [Bay] uses “wave-function” where we use “state.” We use **wave-function** to mean the spatial part of a state.)

Shells, subshells, and orbitals appear in the hydrogen atom,
but larger multi-electron atoms also have approximately the same structure.

This shell structure approximation is the **Hartree approximation**, and results from assuming that the electrons in small to medium-sized atoms (Z ≤ ~23) behave almost like independent electrons in a spherically-symmetric net potential [Bay p451]. This allows us to consider the electrons separately, and since the potential is approximately spherically symmetric, the Schrödinger solutions are products of radial functions and spherical harmonics (i.e., shells, subshells, and orbitals). In multi-electron atoms, only the radial functions differ substantially from hydrogen [E&R p323]. The Hartree approximation ignores electron antisymmetry, which makes it very inaccurate. The Hartree-Fock approximation adds the most important electron antisymmetry, which is an essential aspect of atomic structure.

The **electron configuration** of an atom specifies the *n* and *l* quantum numbers of all electrons, but not their *m* quantum numbers (L_z components), or spins. Neither does the configuration specify how the spins and orbital angular momenta combine. It is the purpose of Hund's rules to predict such properties of the atom that depend on these other quantum numbers. While Hund's rules were first determined empirically, we present here the quantum mechanical explanation of them.

Our analysis uses stationary-state degenerate perturbation theory. Our unperturbed basis for each electron is a set of eigenstates of the screened-nucleus Coulomb energy:

$$\hat{H} = \hat{H}_0 + \hat{H}_{so} \qquad where \qquad \hat{H}_0 = -\frac{Z_{eff}\, e^2}{r}; \; Z_{eff} \equiv \text{effective central charge}; \; e \equiv \text{proton charge}; .$$

$$\hat{H}_{so} \equiv \text{spin-orbit energy} \propto \mathbf{L}\bullet\mathbf{S} .$$

The effective central charge is the nuclear charge, minus one unit for each electron in a full inner shell, because full electron shells are approximately spherically symmetric, and the negative electron charge cancels some of the true nuclear charge. This makes the effective central charge one positive unit for each valence electron, i.e. for each electron in the outermost shell.

The hamiltonian above allows our unperturbed basis (for each electron) to be either (l, m_l, s, m_s), or (l, s, j, m_j), since they all have the same energy. For our analysis, it is simpler to choose the (l, m_l, s, m_s) basis. In this basis, each electron has well-defined quantum numbers for n, l, m_l, and m_s. (s = ½ always for an electron). Because these base states are degenerate in energy, we *cannot* use the simple (non-degenerate) perturbation energy $E_\psi^{(1)} = \langle \psi | H_{so} | \psi \rangle$.

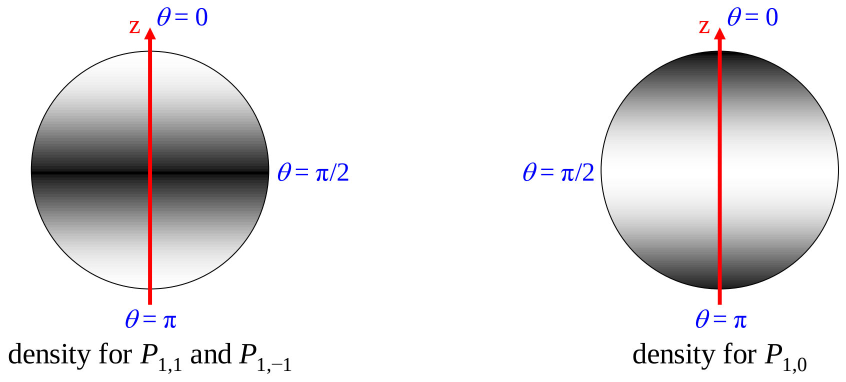
Here is a summary of the symbols we use, which are explained in more detail as we go. In general, lowercase letters refer to a single electron, and capitals to the total over all the subshell electrons:

k	# of electrons in the subshell, $k \le 2(2l + 1)$.
u	is a summation index.
l	orbital angular momentum quantum # of the subshell we're "filling" with electrons.
m	generic *z*-component *orbital* angular momentum quantum #.
m_1, m_2, ...	the *z*-component *orbital* angular momentum quantum # of electron 1, 2, ...
S	the total *spin* angular momentum quantum # of all the electrons in the subshell.
L	the total *orbital* angular momentum quantum # of all the electrons in the subshell.
M_L	the total *z*-component *orbital* L quantum # of all the electrons in the subshell.
M_S	the total *z*-component *spin* S quantum # of all the electrons in the subshell.
J	the total (spin + orbital) angular momentum quantum # of the subshell.
M_J	the total *z*-component (spin + orbital) angular momentum quantum # of the subshell.
Z	atomic number.

Bold letters refer to the actual vectors, still following the lower/upper-case notation for individual/total quantities, e.g. **l**, **s** for individual orbit and spin angular momentum vectors, and **L**, **S**, **J**, etc. for total subshell vectors.

6.3.1.1 Angular momentum coupling of atomic electrons: *S*, *L*, and *J*.

(excerpted)



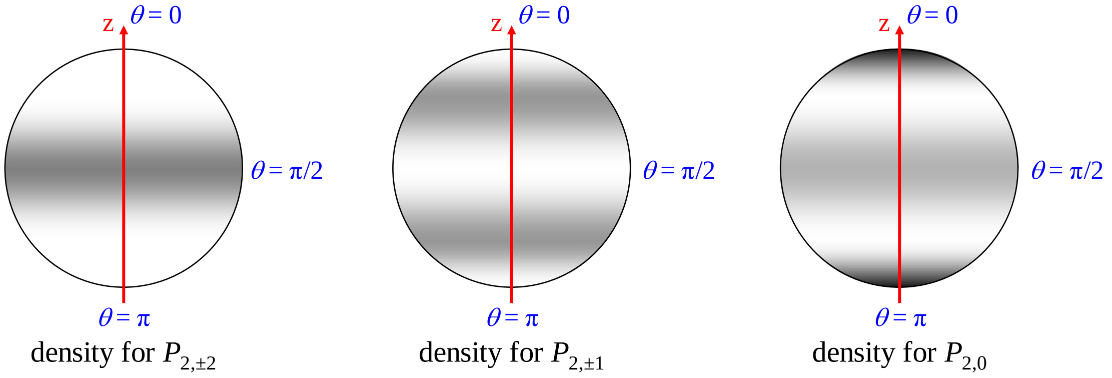
density for $P_{1,1}$ and $P_{1,-1}$

density for $P_{1,0}$

Figure 6.3 Areal particle density for $l = 1$ P_{lm} states (darker is more dense).

When $m = 0$, $|P_{10}|^2 \sim \cos^2\theta$, density is concentrated at the *z*-axis poles (Figure 6.3, right). So two electrons are farthest apart if they are in $m_1 = 1$ or -1, and $m_2 = 0$ states. If they were in $m_1 = 1$ and $m_2 = -1$ states, they'd be nearly "on top of each other" at the equator. Even though antisymmetrization would keep them from being exactly on top of each other, they'd be much closer than in the $m_1 = \pm 1$ and $m_2 = 0$ states. Closer electrons means more Coulomb energy; farther means less energy, and hence the ground state has distinct orbital *m* values of $m_1 = -1$ and $m_2 = 0$. Furthermore, taking the spin as "up," $m_1 = -1$ is lower energy than $m_1 = 1$ because spin-orbit coupling has lower energy when spin opposes orbital angular momentum.

For *l* = 2 (*d*-subshell): When $m = \pm 2$, $|P_{2,\pm 2}|^2 \sim \sin^4\theta$, the density is tightly concentrated at the "equator" (Figure 6.4, left):



density for $P_{2,\pm 2}$

density for $P_{2,\pm 1}$

density for $P_{2,0}$

Figure 6.4 Areal particle density for $l = 2$ P_{lm} states (darker is more dense).

(excerpted)

6.4 Multi-Particle Entanglement

We have seen multi-particle states that are simple tensor products of single-particle states, e.g. two particles where particle 1 is spin up, and particle 2 is spin down:

$$|\chi_{12}\rangle = |\uparrow_1\rangle \otimes |\downarrow_2\rangle \equiv |\uparrow_1\rangle |\downarrow_2\rangle \equiv |\uparrow_1 \downarrow_2\rangle \equiv |\uparrow\downarrow\rangle .$$

Each particle is in a definite state, and measuring either (or both) particles produces the only possible result: particle 1 measures up, and particle 2 measures down.

However, we have also seen that a general 2-particle state need not be a simple (tensor) product of 1-particle states. It could be a superposition of such product states, e.g.:

$$|\chi_{12}\rangle = \frac{1}{\sqrt{2}}|\uparrow\downarrow\rangle + \frac{1}{\sqrt{2}}|\downarrow\uparrow\rangle, \qquad \text{or} \qquad \chi_{12} = \frac{1}{\sqrt{2}}|\uparrow\downarrow\rangle - \frac{1}{\sqrt{2}}|\downarrow\uparrow\rangle . \tag{6.2}$$

Neither of the above states can be written as a tensor product of single particle states. Neither particle has a definite direction of spin. What happens when we measure individual particles in such states? Such states, where neither particle has a definite value of some property, but the values of the two particles are interdependent, are called **entangled states**, and such particles are said to be **entangled**. This means a measurement of one system provides information about the other. [Entangled particles are sometimes called "correlated," but we avoid this term because it conflicts with the statistical term "correlated."]

6.4.1 Partial Measurements

As always, we use here the term "collapse" of the quantum state as a shortcut to the more complicated process of making a measurement, described in section 1.12. Consider the entanglement of two spin-1/2 particles (in the usual *z*-basis):

$$|\chi_{12}\rangle = \frac{1}{\sqrt{2}}|\uparrow_1\rangle|\downarrow_2\rangle + \frac{1}{\sqrt{2}}|\downarrow_1\rangle|\uparrow_2\rangle .$$

What if we measure a property of only one of those particles? How does the state then collapse? Answer:

Partial measurement leads to partial collapse.

In the above state, we easily suppose that there is a 50% chance of measuring particle 1 to be up, but that either way, if we then measure particle 2, it will be opposite to that of particle 1. This is correct, but other superpositions are not so obvious (as shown shortly), so we must develop a mathematical formalism to unambiguously compute the results of measurement on any 2-particle state.

To find the probability of measuring a property of *only one* particle of a two-particle system, we use a "partial inner product" . Recall that in a 1-particle system, an inner product is a scalar. We now extend this idea to a **partial inner product**: Given a ket in the tensor-product space (i.e. a 2-particle ket such as $|\uparrow_1>|\uparrow_2>$), we form a partial inner product with a particle 1 bra; the result is a particle 2 *ket*, i.e. a ket in the particle 2 ket space. We have:

$$\langle\uparrow_1|\chi_{12}\rangle = \frac{1}{\sqrt{2}}\underbrace{\langle\uparrow_1|\uparrow_1\rangle}_{1}|\uparrow_2\rangle + \frac{1}{\sqrt{2}}\underbrace{\langle\uparrow_1|\downarrow_1\rangle}_{0}|\downarrow_2\rangle = \frac{1}{\sqrt{2}}|\uparrow_2\rangle .$$

In other words, the bra from particle one forms an inner product with the particle 1 piece of the 2-particle ket, leaving the particle 2 ket alone. As with our "standard" inner product, the partial inner product is linear in the ket (and anti-linear in the bra), so the partial inner product distributes across the superposition of the ket. (If our quantum states were continuous, our inner product would be an integral, and we would say we "integrate out" particle 1.) With this definition of a partial inner product, we find that the probability of obtaining a given measurement follows essentially the same well-known rule as for a single particle:

$$\Pr(\uparrow_1) = \left|\langle\uparrow_1|\chi_{12}\rangle\right|^2 = \left|\frac{1}{\sqrt{2}}|\uparrow_2\rangle\right|^2 = \frac{1}{2} .$$

The only difference from the single particle case is that for a single particle, the inner product is a scalar, and the probability is its squared magnitude, but in the 2-particle case, the inner product is a *vector* (a particle 2 ket), and the probability is the squared magnitude of this vector.

We now consider a more complicated example. Recall that for a single particle, a state of definite spin pointing in the *x*+ direction is:

$$|x+\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle + |\downarrow\rangle\right).$$

Suppose our 2-particle state is:

$$\left|\chi_{12}\right\rangle = \frac{1}{\sqrt{2}}\left|\uparrow_1\uparrow_2\right\rangle + \frac{1}{\sqrt{2}}\left|\downarrow_1\downarrow_2\right\rangle .$$

By inspection, you might think that measuring either particle 1 or particle 2 along the *x*-axis will give a definite value of |*x*+> (probability = 1). Let's test this by computing the probability for measuring particle 1 in |*x*+>:

$$\begin{aligned}
&\Pr\left(\left|x+_1\right\rangle\right) = \left|\left\langle x+_1 \middle| \chi_{12}\right\rangle\right|^2 \\
&\left\langle x+_1 \middle| \chi_{12}\right\rangle = \frac{1}{\sqrt{2}}\left(\left\langle\uparrow_1\right| + \left\langle\downarrow_1\right|\right)\frac{1}{\sqrt{2}}\left(\left|\uparrow_1\uparrow_2\right\rangle + \left|\downarrow_1\downarrow_2\right\rangle\right) \\
&\qquad = \frac{1}{2}\left[\underbrace{\left\langle\uparrow_1\middle|\uparrow_1\right\rangle}_{1}\left|\uparrow_2\right\rangle + \underbrace{\left\langle\uparrow_1\middle|\downarrow_1\right\rangle}_{0}\left|\downarrow_2\right\rangle + \underbrace{\left\langle\downarrow_1\middle|\uparrow_1\right\rangle}_{0}\left|\uparrow_2\right\rangle + \underbrace{\left\langle\downarrow_1\middle|\downarrow_1\right\rangle}_{1}\left|\downarrow_2\right\rangle\right] \\
&\qquad = \frac{1}{2}\left[\left|\uparrow_2\right\rangle + \left|\downarrow_2\right\rangle\right] = \frac{1}{\sqrt{2}}\left|x+_2\right\rangle .
\end{aligned} \tag{6.3}$$

As expected, the inner product is a particle 2 ket. The probability of $|x+_1>$ is the squared magnitude of this ket, which is only ½! Therefore, despite its look, the state $\frac{1}{\sqrt{2}}\left(\left|\uparrow_1\uparrow_2\right\rangle + \left|\downarrow_1\downarrow_2\right\rangle\right)$ is *not* a state of definite particle 1 spin in the *x*+ direction. Looks can be deceiving. By symmetry, it is also *not* a state of definite particle 2 spin.

We are now ready to address partial collapse. We know that measurements collapse quantum states, so if we measure particle 1 to be $|x+_1>$, what is our resulting system state? The particle 1 state is that consistent with our measurement, i.e. $\left|\chi_1\right\rangle = \left|x+_1\right\rangle$ (just as with single-particle collapse). The particle 2 state is simply that of the inner product above, except that being a quantum state, we must normalize it to unit magnitude:

$$\left|\chi_2\right\rangle = normalize\left(\frac{1}{\sqrt{2}}\left|x+_2\right\rangle\right) = \left|x+_2\right\rangle .$$

Therefore, our 2-particle state after measuring particle 1 is:

$$\chi_{12,after} = \left|x+_1\right\rangle\left|x+_2\right\rangle \quad \text{or} \quad \left|x+_1, x+_2\right\rangle \qquad \text{(after measurement)} .$$

Note that the initial state is entangled, and in this case, measuring particle 1 determines also the state of particle 2.

The probability of measuring $|x+_1, x+_2>$ was only ½. What else might we have measured? And with what probabilities? We can compute these by subtracting this known component state from our initial state:

$$\begin{aligned}
&\left|\chi_{12}\right\rangle = \frac{1}{\sqrt{2}}\left|\uparrow_1\uparrow_2\right\rangle + \frac{1}{\sqrt{2}}\left|\downarrow_1\downarrow_2\right\rangle = \frac{1}{\sqrt{2}}\left|x+_1, x+_2\right\rangle + \left|others\right\rangle \qquad \Rightarrow \\
&\qquad \left|others\right\rangle = \frac{1}{\sqrt{2}}\left|\uparrow_1\uparrow_2\right\rangle + \frac{1}{\sqrt{2}}\left|\downarrow_1\downarrow_2\right\rangle - \frac{1}{\sqrt{2}}\left|x+_1, x+_2\right\rangle .
\end{aligned} \tag{6.4}$$

Now we expand the entangled state $|x+_1, x+_2>$ in our *z* basis:

$$\left|x+_1, x+_2\right\rangle = \underbrace{\frac{1}{\sqrt{2}}\left(\left|\uparrow_1\right\rangle + \left|\downarrow_1\right\rangle\right)}_{\left|x+_1\right\rangle} \otimes \underbrace{\frac{1}{\sqrt{2}}\left(\left|\uparrow_2\right\rangle + \left|\downarrow_2\right\rangle\right)}_{\left|x+_2\right\rangle} = \frac{1}{2}\left(\left|\uparrow_1\uparrow_2\right\rangle + \left|\uparrow_1\downarrow_2\right\rangle + \left|\downarrow_1\uparrow_2\right\rangle + \left|\downarrow_1\downarrow_2\right\rangle\right) . \tag{6.5}$$

Then:

$$\left|others\right\rangle = \frac{1}{\sqrt{2}}\left|\uparrow_1\uparrow_2\right\rangle + \frac{1}{\sqrt{2}}\left|\downarrow_1\downarrow_2\right\rangle - \frac{1}{\sqrt{2}}\frac{1}{2}\left(\uparrow_1\uparrow_2 + \uparrow_1\downarrow_2 + \downarrow_1\uparrow_2 + \downarrow_1\downarrow_2\right)$$

$$= \frac{1}{2\sqrt{2}}\left[\left|\uparrow_1\uparrow_2\right\rangle - \left|\uparrow_1\downarrow_2\right\rangle - \left|\downarrow_1\uparrow_2\right\rangle + \left|\downarrow_1\downarrow_2\right\rangle\right].$$

The ket in square brackets factors into a tensor product of single-particle states:

$$\left|others\right\rangle = \frac{1}{2\sqrt{2}}\left[\underbrace{\left(\left|\uparrow_1\right\rangle - \left|\downarrow_1\right\rangle\right)}_{\sqrt{2}\left|x-_1\right\rangle} \otimes \underbrace{\left(\left|\uparrow_2\right\rangle - \left|\downarrow_2\right\rangle\right)}_{\sqrt{2}\left|x-_2\right\rangle}\right] = \frac{1}{\sqrt{2}}\left|x-_1, x-_2\right\rangle .$$

Plugging in to (6.4) we see that our initial state is a superposition of two simple product states:

$$\left|\chi_{12}\right\rangle = \frac{1}{\sqrt{2}}\left|\uparrow_1\uparrow_2\right\rangle + \frac{1}{\sqrt{2}}\left|\downarrow_1\downarrow_2\right\rangle = \frac{1}{\sqrt{2}}\left|x+_1, x+_2\right\rangle + \frac{1}{\sqrt{2}}\left|x-_1, x-_2\right\rangle .$$

It is interesting that $|\chi_{12}>$ has the same form in both the *z* basis and the *x* basis. We see this by rewriting our shorthand for the *z* basis in the notation we used for the *x* basis:

$$\left|\chi_{12}\right\rangle = \frac{1}{\sqrt{2}}\left|\uparrow_1\uparrow_2\right\rangle + \frac{1}{\sqrt{2}}\left|\downarrow_1\downarrow_2\right\rangle \equiv \frac{1}{\sqrt{2}}\left|z+_1, z+_2\right\rangle + \frac{1}{\sqrt{2}}\left|z-_1, z-_2\right\rangle .$$

In contrast to this superposition of product states, let's consider a simple tensor product state, $|x+_1> \otimes |x+_2>$, which we worked out in (6.5). By construction, it *is* a state of definite particle 1 spin in the *x*+ direction, *and separately* definite particle 2 spin. In expanded form, as on the RHS of (6.5), it is hard to see by inspection that this state has definite spin for both particles.

These principles of partial measurements, partial inner products, probabilities of partial measurements, and partial collapse apply to any tensor product state, including space ⊗ spin. Recall the space ⊗ spin example of the Stern-Gerlach device we considered in section 0: position is entangled with spin, so position is a *proxy* for spin. Measuring the position tells us also the spin, and collapses the state to one of definite position *and* definite spin.

6.4.2 The EPR Paradox (Not)

Multi-particle QM provides a way to decide whether nature is truly probabilistic, or QM is incomplete. The effect we examine is called the "EPR paradox," after Einstein, Podolski, and Rosen, who first brought attention to it with a thought experiment. We show that, though the result may be unexpected and counterintuitive, there is no contradiction in the physics, and nature is indeed probabilistic. This result is quite important: you can now buy commercial quantum encryption systems that use this principle. Such systems are essentially unbreakable (so long as QM is correct). Our development uses only the simple QM we've already developed, and also shows that the axioms of quantization, entanglement, and the correspondence principle together imply that QM must be nonlocal (i.e., must have effects that span distances in a time faster than light can traverse them). Our analysis is inspired by a section of Heinz R. Pagels' book *In Search of Schrödinger's Cat*.

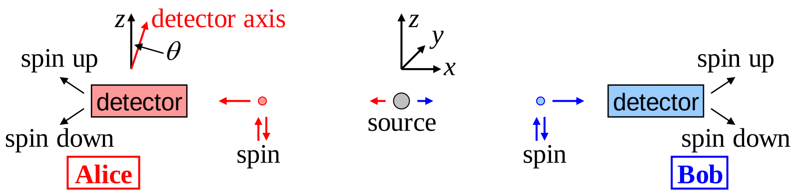


Figure 6.5 Schematic of the EPR experiment. Alice's detector axis is off by θ.

Is QM is incomplete? Are there "hidden variables" which quantum theory is neglecting, that would eliminate the uncertainty of QM? In the EPR experiment, a spin-0 source emits a pair of particles, A and B, in opposite directions. The total spin of the pair is zero, but they are emitted in the entangled state |↑↓>

+ |↓↑> (the first arrow refers to particle A, the 2nd to particle B). The experiment measures the spins of both particles, and is repeated over many pairs, with varying directions of measurement. (In principle, the experiment can use any of a variety of 2-state properties, such as polarization of photons, etc.)

(excerpted)

7 Quantum Electromagnetic Radiation

Electromagnetic interactions are probably the most fundamental and important in our lives. Most of our everyday experience is dominated by EM interactions (and gravity). Quantum electromagnetism is the first Quantum Field Theory, on which all following QFTs are built. Thus quantum electromagnetism not only explains much of the world around us, but is a prerequisite for the QFTs that explain the rest.

We present here an overview of quantized radiation, and matter-radiation interaction. This is a huge field, and an active area of research. [GAF] provides an accessible, but fairly thorough, tutorial on the subject.

In many cases, one can reasonably approximate some features of quantized matter interactions with radiation by using the "semi-classical approximation:" one treats the charged particles as quantized, but the radiation field as a classical potential in which the particle acts. Such an approximation describes EM absorption by matter, and stimulated emission. This can also be thought of as a semi-classical EM field: the photon has a definite energy, and a simple wave-function, just like ordinary quantum mechanics. For a semi-classical EM field, the vector potential $\mathbf{A}(t, \mathbf{r})$ is essentially the wave-function of the photon. $|\mathbf{A}(t, \mathbf{r})|^2$ is proportional to the particle density of the photon at $(t, \mathbf{r})$. We do not address this semi-classical approximation here. We also do not address spin interactions.

However, many phenomena cannot be described by a semi-classical approximation, the most important being spontaneous emission. It requires a fully quantized EM field, which allows for the creation of a photon even in the absence of any prior EM excitation, i.e. even into the vacuum. The quantized EM field is the main topic of this chapter. It also quantitatively describes the Lamb shift of atomic spectra, the gyromagnetic ratio of the electron, the Casimir force, and multiple photon detections.

In general, a quantized field allows for particle creation and annihilation, and is the essence on which all QFTs are built.

We discuss the following aspects of quantum electromagnetic radiation:

1. Quantized EM field: A failed attempt at photons.
2. Quantized EM field: the simple harmonic oscillator (SHO) analogy.
3. Quantized EM fields conceptually introduce a new kind of wave-function (the quantized field), but we find that we never need to explicitly use it.
4. Differences between a 1D harmonic oscillator and the EM field.
5. Combining matter states and photon (radiation) states.
6. Example of a simplified matter-photon operator in action, evaluating an inner product.
7. Example of the complete $\hat{\mathbf{A}}$ operator (vector potential) in action, evaluating an inner product.
8. Spontaneous radiation.
9. Photons have all phases.
10. Photons in other bases (other modes).
11. The wave-function of a photon.
12. Quasi-classical states, average E-field is not quantized, and interference (number operators).

On a first reading, one can skip the more advanced photon and quasi-classical theory.

This section requires a thorough understanding of classical EM propagation, including phasors (the complex representation of sinusoids), wave-vectors, elementary polarization, and the 3-vector potential, $\mathbf{A}(t, \mathbf{r})$. You must also understand Dirac notation, the quantum simple harmonic oscillator (SHO), ladder operators, multi-particle QM, and tensor product states and their inner products.

Before going further, we again caution against stating what is "really" happening at the microscopic, quantum scale. None of quantum mechanics is directly observable. Therefore, we present here the generally accepted model, as simply as we can. The model is based on now-familiar QM principles, and it quantitatively predicts the outcomes of experiments.

7.1 Quantized EM field: A Failed Attempt at Photons

Imagine an essentially classical EM wave of frequency ω in a volume of space, but quantized in amplitude such that its energy is $E = \hbar\omega$. Then:

$$\mathbf{A}(t,\mathbf{r}) = \mathrm{Re}\left\{\mathbf{A}_0 e^{i(\mathbf{k}\bullet\mathbf{r}-\omega t)}\right\} \quad \textit{where} \quad \mathbf{A}_0 \equiv \text{complex vector};$$

$$E = \hbar\omega \propto \int_\infty d^3\mathbf{r}\ \mathbf{A}^2(t,\mathbf{r}) = \frac{1}{2}(Vol)\left|\mathbf{A}_0\right|^2 .$$

We might propose this as a crude model for a "photon," a quantized EM field (though we will see it is not a very good model). We usually separate out the overall magnitude and phase of the wave from its polarization, using a phasor, A_0 (a complex number representing a sinusoid) [GAF p354], such that:

$$\mathbf{A}_0 = A_0\boldsymbol{\varepsilon} .$$

. Then:

$$\mathbf{A}(t,\mathbf{r}) = \mathrm{Re}\left\{A_0\boldsymbol{\varepsilon} e^{i(\mathbf{k}\bullet\mathbf{r}-\omega t)}\right\} \qquad \textit{where} \quad A_0 \equiv \text{phasor};\ \ \boldsymbol{\varepsilon} \equiv \text{complex polarization vector} .$$

$$\left|\boldsymbol{\varepsilon}\right|^2 = 1 \quad \Rightarrow \qquad E = \hbar\omega \propto \frac{1}{2}(Vol)\left|A_0\right|^2 .$$

This simple model of a photon explains the well-known photoelectric effect: it takes radiation of a certain frequency or higher to ionize a substance. No amount of light of lower frequency, no matter how intense, will ionize the substance. We explain this frequency cutoff by noting that for ionization, our incident radiation must provide enough energy to unbind an electron. Since we have supposed that radiation is quantized into bundles of energy called photons, ionization requires a photon of sufficient energy. Finally, since $E = \hbar\omega$, sufficient energy for ionization requires sufficient frequency ω.

However, this simple model does not explain several phenomena, such as spontaneous emission, the Lamb shift of atomic spectra, the gyromagnetic ratio of the electron, and the Casimir force.

7.2 Photon Number States

Despite the failure of our simple model, it does suggest a useful concept for the quantum state of an EM field. We know that some kind of quantized EM field exists, since photons of energy $E = \hbar\omega$ are seen experimentally. Furthermore, the existence of coherent classical EM fields suggests that multiple photons of a single mode (single frequency, phase, and polarization) can exist. Also, multiple modes of excitation can simultaneously exist. We therefore suppose that the quantum state of an EM field can be given by a list of each mode, and the number of photons in that mode. For example, given modes characterized by wave-vectors $\mathbf{k}_j$, and polarization vectors $\boldsymbol{\varepsilon}_1(\mathbf{k}_j)$ and $\boldsymbol{\varepsilon}_2(\mathbf{k}_j)$, we might have EM states such as:

$$\left|N_1(\mathbf{k}_1)\right\rangle; \qquad \left|N_2(\mathbf{k}_1)\right\rangle; \qquad \left|N_1(\mathbf{k}_2)\right\rangle; \quad \ldots \quad \left|N_\lambda(\mathbf{k}_n)\right\rangle,$$

$$\lambda = 1,2; \quad N_\lambda = 0,1,\ldots\infty; \quad n = 1,2,\ldots\infty .$$

Note that the subscript on N and $\boldsymbol{\varepsilon}$ refers to one of two polarization states, whereas the subscript on $\mathbf{k}$ refers to one of an infinite number of wave-vectors.

In the momentum basis, a photon is described by its wave-vector, $\mathbf{k}$ (aka propagation vector), and its polarization vector, $\boldsymbol{\varepsilon}$. Therefore, a photon state is a (polarization-vector, wave-vector) pair, written as $\boldsymbol{\varepsilon}(\mathbf{k})$. The wave-vector tells you the state's EM propagation direction and spatial frequency $|\mathbf{k}| = 2\pi/\lambda$ rad/m, and therefore also temporal frequency $\omega = c|\mathbf{k}|$. The polarization vector tells you how the photon state is polarized. For each $\mathbf{k}$, two independent polarizations exist, $\boldsymbol{\varepsilon}_1(\mathbf{k})$ and $\boldsymbol{\varepsilon}_2(\mathbf{k})$. A general photon state is then written $\boldsymbol{\varepsilon}_\lambda(\mathbf{k})$, where $\lambda = 1$ or 2. [Many references use the alternate notation $\boldsymbol{\varepsilon}_{\mathbf{k}\lambda}$.]

For every wave-vector **k** (aka propagation vector),
there are two independent polarizations that can propagate with that wave-vector.

Here, we do not need to know the details of a polarization vector; we need only that for any given propagation vector **k**, there are two independent polarization modes, say horizontal and vertical, or right-hand circular (RHC) and LHC. Each polarization mode is described by the presence of a distinct photon. Either photon or both may exist in space.

Photons are bosons, so a single state can be occupied by any number of photons, from 0 on up. A complete photon state therefore includes a photon count (aka **occupation number**) for each (**k**, **ε**) pair, written in general as $N_\lambda(\mathbf{k}_j)$.

We write an EM state of multiple mode excitations (multiple photons) as a single ket, e.g.:

$$\left|N_2(\mathbf{k}_1)=3, N_1(\mathbf{k}_2)=2, N_1(\mathbf{k}_3)=1\right\rangle \qquad \text{(multi-mode quantum EM state)}. \qquad (7.1)$$

Of course there are an infinite number of possible modes, and most of them are unoccupied ($N_\lambda(\mathbf{k}) = 0$), so we omit those from the notation: any modes not listed in the ket are defined to be unoccupied. Such a state is said to be written in the **number basis**.

A two-photon state is *not* the vector sum of two one-photon states.

For example:

$$\left|N_1(\mathbf{k}_1)=1, N_1(\mathbf{k}_2)=1\right\rangle \ne \left|N_1(\mathbf{k}_1)=1\right\rangle + \left|N_1(\mathbf{k}_2)=1\right\rangle .$$

For one thing, the RHS is not normalized. Even if we normalized it, though, it would be a superposition of two one-photon states. We might detect either one of the component photons (thus collapsing the wave-function), but not both. In contrast, a two-photon state of the EM field has two whole photons in it, and both can be detected.

The state with no photons at all is called the **vacuum state**, and written:

$$\left|0\right\rangle \equiv \text{vacuum state.} \qquad \text{Note: } \left|0\right\rangle \ne \mathbf{0_v} .$$

NB: the vacuum state is *not* the zero-vector, $\mathbf{0_v}$ (which is also known as the “null ket”).

For brevity, we may write modes of a single excitation (N = 1) as:

$$\left|\boldsymbol{\varepsilon}_1(\mathbf{k})\right\rangle \equiv \left|N_1(\mathbf{k})=1\right\rangle, \qquad \text{or more generally,} \qquad \left|\boldsymbol{\varepsilon}_\lambda(\mathbf{k})\right\rangle \equiv \left|N_\lambda(\mathbf{k})=1\right\rangle .$$

Since the occupation number is 1, we don’t write it explicitly.

In a state of multiple modes, the order doesn’t matter. E.g., a state with two photons, one of wave-vector $\mathbf{k}_1$, and polarization $\boldsymbol{\varepsilon}_1(\mathbf{k}_1)$, and another of $\boldsymbol{\varepsilon}_1(\mathbf{k}_2)$, can be written: $|\boldsymbol{\varepsilon}_1(\mathbf{k}_1), \boldsymbol{\varepsilon}_1(\mathbf{k}_2)>$, or $|\boldsymbol{\varepsilon}_1(\mathbf{k}_2), \boldsymbol{\varepsilon}_1(\mathbf{k}_1)>$. Again, since the occupation numbers are 1, we don’t write them explicitly.

Note that higher excitations, N ≥ 2, describe *independent* photons, each with completely indeterminate phase [Lou p253b].

Aside On Polarization Vectors

The polarization vector completely describes the polarization of the state, given as the polarization of its vector-potential, **A**. Classically, the polarization vector is a vector of three phasors: one for each component of the A-field: A_x, A_y, and A_z. In QM, for a single photon, the polarization vector is a complex valued vector in ordinary 3-space:

$$\boldsymbol{\varepsilon} = (\varepsilon_x, \varepsilon_y, \varepsilon_z) \qquad \textit{where} \quad \varepsilon_x, \varepsilon_y, \varepsilon_z \text{ are complex components},$$

giving the QM amplitude for the *x*, *y*, and *z* components of **A**. Polarization vectors are dimensionless, and normalized to unit magnitude:

$$|\boldsymbol{\varepsilon}|^2 = \boldsymbol{\varepsilon}\bullet\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}^*\boldsymbol{\varepsilon} = |\varepsilon_x|^2 + |\varepsilon_y|^2 + |\varepsilon_z|^2 = \varepsilon_x^*\varepsilon_x + \varepsilon_y^*\varepsilon_y + \varepsilon_z^*\varepsilon_z = 1 .$$

[It is not necessary here to understand the details of what a polarization vector means, however *Funky Electromagnetic Concepts* explains them completely for **E** fields; the transition to **A** fields is straightforward.] All that matters right now is that for every wave-vector **k**, there are two independent polarizations that can propagate with that wave-vector.

As with all quantum states, the polarization state can be a superposition of $\boldsymbol{\varepsilon}_1$ and $\boldsymbol{\varepsilon}_2$. Therefore, for a given propagation direction, $\mathbf{e_k} \equiv \mathbf{k}/|\mathbf{k}|$, any polarization vector (aka polarization "state") can be written as a linear combination of two basis polarization vectors, $\boldsymbol{\varepsilon}_1$ and $\boldsymbol{\varepsilon}_2$. The most common basis polarization vectors are RHC (right hand circular) and LHC (left hand circular), which we could write as $\boldsymbol{\varepsilon}_R(\mathbf{k})$ and $\boldsymbol{\varepsilon}_L(\mathbf{k})$. (These are handy because they are angular momentum eigenstates.) However, for most of our calculations, the basis is irrelevant. Note that the polarization vector $\boldsymbol{\varepsilon}_R(\mathbf{k})$ *is a function of the direction of* **k**, so $\boldsymbol{\varepsilon}_R$(*z*-direction) is a *different* vector than $\boldsymbol{\varepsilon}_R$(*x*-direction). [In fact, $\boldsymbol{\varepsilon}_R(\mathbf{e}_z) = (1, i, 0)/\sqrt{2}$, and $\boldsymbol{\varepsilon}_R(\mathbf{e}_x) = (0, 1, i)/\sqrt{2}$, but we don't need to know that here.]

7.3 Quantized EM Field: The Simple Harmonic Oscillator Analogy

When we describe the interaction of radiation and matter, we must talk about photons being created (radiated), and destroyed (absorbed). Thus we introduce a new concept to quantum mechanics: particle creation and destruction (aka **annihilation**). It is this concept that demands Quantum Field Theory (aka "second quantization"). In this section, we take an approach known as "canonical quantization," where the classical vector-potential (**A**-field) becomes an operator $\hat{\mathbf{A}}$, acting on photon states.

[A different approach uses Feynman Path Integrals (FPIs), where the **A**-field remains a complex number function of space (as in classical E&M when working in Fourier space). FPI have some advantages over canonical quantization, but we do not address that further here.]

In classical EM, we define a plane wave with a definite vector potential **A**(*t*, **r**):

$$\mathbf{A}(t,\mathbf{r}) = \text{Re}\left\{ A_0\, \boldsymbol{\varepsilon}\, e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} \right\} \qquad \textit{where} \qquad A_0 \text{ is a phasor; } \boldsymbol{\varepsilon} \text{ is the polarization vector ,}$$

with no uncertainty. Note that the classical **A**-field oscillate in both space and time.

However, when we quantize the EM field, i.e. when we consider photons as quantum particles, the vector-potential, **A**, is no longer definite. A state of definite photon count, say $|N_\lambda(\mathbf{k}) = 1>$, has uncertain **A**. **A** is a probabilistic function of the photon state $|N_\lambda(\mathbf{k}) = 1>$. This is analogous to a 1D harmonic oscillator: a particle in a state of definite energy, say $|n> \equiv \psi_n(x)$, has uncertain position. The position *x* is a probabilistic function of the state.

Why should the EM field behave like a simple harmonic oscillator (SHO)? The SHO for a particle has a potential energy proportional to x^2. For EM waves, both the E- and B-fields contain energy proportional to E^2 (and B^2). In the SHO, to move a particle to position *x* takes energy proportional to x^2; in EM waves, to create an E-field with magnitude *E* takes energy proportional to E^2. Thus it is reasonable that EM fields quantize similarly to the SHO. Furthermore, the analogy between harmonic oscillator and vector-potential continues. For a given **k** and λ (which selects $\boldsymbol{\varepsilon}_\lambda(\mathbf{k})$):

$$\begin{aligned} \hat{x} \quad &\leftrightarrow \quad \hat{\mathbf{A}}_\lambda(\mathbf{k}) \\ \hat{x} \sim \left(\hat{a} + \hat{a}^\dagger\right) \quad &\leftrightarrow \quad \hat{\mathbf{A}}_\lambda(\mathbf{k}) \sim \left(\hat{a}_\lambda(\mathbf{k}) + \hat{a}_\lambda^\dagger(\mathbf{k})\right) \\ \text{energy level } n \quad &\leftrightarrow \quad \text{photon count } N_\lambda(\mathbf{k}) \end{aligned}$$

Note that like the SHO quantum number *n*, the photon count, $N_\lambda(\mathbf{k})$, for a plane-wave mode is also a measure of the energy level of the EM field for that mode (i.e., for that value of **k** and $\boldsymbol{\varepsilon}_\lambda$).

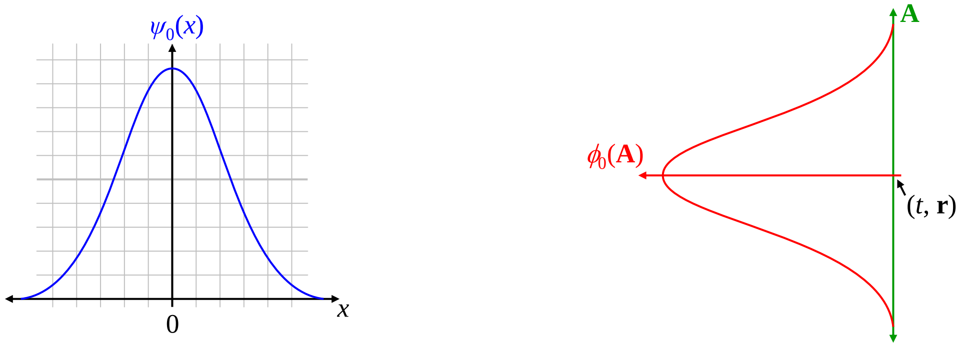


Figure 7.1 Analogy between (left) particle wave-function $\psi_0(x)$ and (right) photon vector-potential wave-function $\phi_0(\mathbf{A})$ at a spacetime point $(t, \mathbf{r})$.

(excerpted)

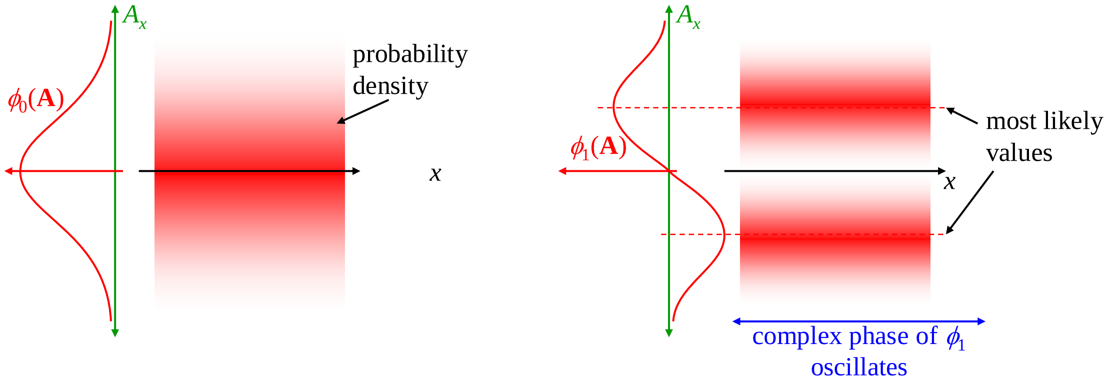


Figure 7.2 Probability densities of values of **A** at a point. (Left) For vacuum, N = 0. (Right) For one photon, N = 1. $\phi(\mathbf{A})$ oscillates in space.

(excerpted)

7.4 Quantum Interaction of Radiation and Matter

We now develop the QM of spontaneous radiation, e.g., an excited atom radiates (creates) a photon, and thus decays to a lower energy state. To compute radiation probabilities, and metastable state lifetimes (or equivalently, decay rates), we must evaluate a matrix element of the form:

$$\left\langle \psi_f, \boldsymbol{\varepsilon}(\mathbf{k}) \middle| \hat{\mathbf{A}} \cdot \hat{\mathbf{p}} \middle| \psi_i, 0 \right\rangle \qquad \textit{where} \quad \left| \psi_i, 0 \right\rangle \text{ is the initial matter state with no photons;}$$

$$\left\langle \psi_f, \boldsymbol{\varepsilon}(\mathbf{k}) \right| \text{ is the final matter state plus one photon .}$$

In other words, we evaluate the complex amplitude to go from an initial matter state with no photons, to a final matter state and one photon. The probability of such a transition is the squared-magnitude of the amplitude, as usual. We now describe the meaning of all the pieces of this inner product, and how to evaluate it, in general.

7.4.1 Multi-particle Matter and Photon States

Quantum states for a system of interacting matter and radiation are combination states (tensor product states): they include a matter piece, and an EM field piece. The matter states are the usual wave-functions (plus spin-states), that we already know and love. For example, the combined matter-EM state:

$$\left|\psi, N_\lambda(\mathbf{k})\right\rangle \equiv \left|\psi\right\rangle \left|N_\lambda(\mathbf{k})\right\rangle \quad \text{has}$$

$$\left|\psi\right\rangle \equiv \text{matter state of a charged particle;}$$

$$\left|N_\lambda(\mathbf{k})\right\rangle \equiv \text{photon state with wave-vector } \mathbf{k}\text{, polarization } \boldsymbol{\varepsilon}_\lambda(\mathbf{k})\text{, and } N_\lambda(\mathbf{k}) \text{ photons present}.$$

A hydrogen atom in the |100> state with a photon in the $\boldsymbol{\varepsilon}_1(\mathbf{k})$ state can be written:

$$\left|100, \boldsymbol{\varepsilon}_1(\mathbf{k})\right\rangle \qquad \text{or} \qquad \left|100\right\rangle \left|\boldsymbol{\varepsilon}_1(\mathbf{k})\right\rangle \qquad \text{(combined matter/photon state)}.$$

[Both of these are shorthand for |100>⊗|$\boldsymbol{\varepsilon}_1(\mathbf{k})$>, which is the tensor-product (aka direct-product) of the two states.] Note that the matter-states exist in one Hilbert space, and the photon states exist in a *different* Hilbert space. Therefore, some operators act on the matter-state alone, some on the photon-state alone, and some act on both the matter state and the photon state.

Because we have chosen that **k** has a definite value for the photons, they are plane-wave eigenstates of momentum. Later, we write them in the position basis, so we can take inner products with $\psi(\mathbf{r})$, the matter state in the position basis.

In this section, our photon counts $N_\lambda(\mathbf{k})$ are always 1 or 0.

7.4.2 Â In Action

(excerpted)

7.5 Photons Have All Phases, Semi-classical Approximation

An individual photon is a superposition of all possible EM field phases [Lou p186-7]. This is completely analogous to a stationary state of a quantum particle harmonic oscillator, which is in a superposition of all possible positions (or "phases" of its oscillation). Therefore, the photon number states we've been discussing, e.g. $|N_1(\mathbf{k}_1), N_1(\mathbf{k}_2), \ldots>$, are called **incoherent states**, since the different modes have no particular phase relationship to each other. (We discuss coherent states later).

The matrix elements we've been computing cannot distinguish between a *superposition* and a *mix* of all phases, so based on experimental results, we could equally well say a single photon state is a *mixed state* of all possible phases [GAL p359t]. However, our model of an A-space "wave-function" (section 7.3) for the A-field is more consistent with a superposition than a mixed state.

A single-photon state is a superposition of all possible phases, and therefore has no definite phase.

The fact that single-photons have all phases leads to an important distinction between the effective A-field of a photon, a complex function of space given by **Error! Reference source not found.**:

$$\left\langle 0\right|\hat{\mathbf{A}}\left|\boldsymbol{\varepsilon}(\mathbf{k})\right\rangle = A_0(\mathbf{k})\boldsymbol{\varepsilon}(\mathbf{k})e^{+i(\mathbf{k}\cdot\mathbf{r}-\omega t)} \equiv \mathbf{A}_{eff}(t,\mathbf{r}) \qquad \text{effective A-field of photon},$$

and the classical EM "analytic" (or phasor) A-field, also a complex function of space, that looks very similar:

$$\mathbf{A}_{analytic}(\mathbf{r}) = A_0\boldsymbol{\varepsilon}e^{+i\mathbf{k}\cdot\mathbf{r}} \qquad \text{classical analytic (phasor) A-field}.$$

The classical *true* A-field is a real-valued function of space, defined by the classical phasor A-field as:

$$\mathbf{A}_{true}(t,\mathbf{r}) = \sqrt{2}\,\mathrm{Re}\left\{\mathbf{A}_{analytic}(\mathbf{r})e^{-i\omega t}\right\} = \sqrt{2}\,\mathrm{Re}\left\{A_0\boldsymbol{\varepsilon}e^{+i(\mathbf{k}\cdot\mathbf{r}-i\omega t)}\right\} = \sqrt{2}A_0\boldsymbol{\varepsilon}\cos\left(\mathbf{k}\bullet\mathbf{r}-\omega t\right).$$

(Other normalizations are often used, as well.) The distinction is that the classical true A-field amplitude varies in space: it has positive peaks, negative troughs, and is zero in places. Where it is zero, there is no chance it interacts with a particle there. In contrast, the effective A-field of the photon is complex, and equal magnitude everywhere. It is equally likely to interact at all points. It has no "zeros." This reflects the fact that the photon is a superposition of all phases, and is equally likely to have any given amplitude at all points.

Note that the factor of $2^{1/2}$ in the formula for $\mathbf{A}_{\text{true}}$ gives the semi-classical "photon" the same energy as a true photon, by making their squared magnitudes, summed over space, equal at every point in time:

$$\int_\infty \mathbf{A}_{true}{}^2 \, d^3r = \int_\infty 2A_0{}^2 \cos^2(\mathbf{r}) \, d^3r = A_0{}^2 (Vol), \qquad \text{and}$$

$$\int_\infty \mathbf{A}_{eff}{}^2 \, d^3r = \int_\infty A_0{}^2(\mathbf{r}) |\boldsymbol{\varepsilon}|^2 \left| e^{i(\mathbf{k}\cdot\mathbf{r})} \right|^2 d^3r = A_0{}^2 (Vol).$$

7.6 Realistic Photons: Multi-Mode Single Photon States

We have taken our basis states for the EM field to be integer excitations of all possible plane-wave EM modes, which for simplicity we take here to be discrete (indexed by l for $\boldsymbol{\varepsilon}_l\mathbf{k}_l$):

$$\begin{array}{ll} |0\rangle, & |n_1=1\rangle, \ |n_1=2\rangle, \ |n_1=3\rangle, \ \ldots \\ & |n_2=1\rangle, \ |n_2=2\rangle, \ |n_2=3\rangle, \ \ldots \\ & \quad \vdots \qquad\qquad\qquad\qquad\qquad \textit{where} \quad l \equiv \boldsymbol{\varepsilon}_l(\mathbf{k}_l). \end{array}$$

However, we know from experiment that atoms radiate single photons, and also that they have some uncertainty in their energy (and therefore in **k**). We must concludes that the EM state for single-photon radiation is a (normalized) superposition of single-photon states of different modes:

$$|\psi_{EM}\rangle = \sum_l c_l \, |\boldsymbol{\varepsilon}_l(\mathbf{k}_l)\rangle = \sum_l c_l \, |N_l = 1\rangle \qquad \textit{where} \quad l \equiv \text{mode of the plane-wave state}. \tag{7.2}$$

When all the $\mathbf{k}_l$ are "nearby," this is a wave-packet. Compared to an infinite plane-wave mode, it is a more realistic single-photon state. This is the EM state after an atom has radiated a photon. It has an envelope of essentially finite size in space.

Such a state is *not* an eigenstate of any single-mode number operator, $\hat{N}_{l'} = \hat{a}^\dagger_{l'}\hat{a}_{l'}$, but *is* an eigenstate of the total number operator, with eigenvalue 1:

$$\hat{N}_{tot} \equiv \sum_{l'=1}^{\infty} \hat{N}_{l'} \qquad \Rightarrow \qquad \hat{N}_{tot}|\psi_{EM}\rangle = \sum_{l'=1}^{\infty}\sum_{l=1}^{n} \hat{N}_{l'} \, c_l \, |N_l = 1\rangle = \sum_{l=1}^{n} 1 \cdot c_l \, |N_l = 1\rangle = 1|\psi_{EM}\rangle.$$

In the double sum, we used the fact that only terms with *l'* = *l* contribute. Since the total photon number observable is 1, it is a single-photon state, even though it is a superposition of many modes. It can be called a "multi-mode single-photon state." Such states could be propagating wave-packets, localized in space, and moving in time.

The inner product for detecting a photon of mode *l'* will be of the form:

$$\langle 0|\hat{a}_{l'}|\psi_{EM}\rangle = \langle 0|\sum_{l=0}^{\infty} c_l \hat{a}_{l'} \, |N_l = 1\rangle = c_{l'}\langle 0|0\rangle = c_{l'},$$

where we have used:

$$\hat{a}_{l'}|N_{l'} = 1\rangle = |0\rangle, \ \text{and} \qquad \hat{a}_{l'}|N_l = 1\rangle = \mathbf{0}_v \quad (l' \ne l), \qquad\qquad \text{and therefore:}$$

$$\begin{aligned} \hat{a}_{l'}|\psi_{EM}\rangle &= \hat{a}_{l'}c_0 \, |N_0 = 1\rangle + \hat{a}_{l'}c_1 \, |N_1 = 1\rangle + \hat{a}_{l'}c_2 \, |N_2 = 1\rangle + \ldots \hat{a}_{l'}c_{l'} \, |N_{l'} = 1\rangle + \ldots \\ &= \mathbf{0_v} + \mathbf{0_v} + \mathbf{0_v} + \ldots c_{l'}|0\rangle + \ldots = c_{l'}|0\rangle. \end{aligned}$$

The multi-mode single-photon state is the EM state most like a single matter particle.

Note that the probability of detecting two photons from such a state is zero, because such a detection requires *two* lowering operators, one for each photon detected [GAF p376]. As just shown, any one lowering operator produces a ket proportional to the vacuum state |0>. Therefore any second lowering operator returns the zero vector **0**$_\mathbf{v}$. Thus:

$$\langle 0 | \hat{a}_{l'} \hat{a}_{l''} | \psi_{EM} \rangle = 0 \qquad \forall\, l, l' ,$$

and the probability of any such double detection is 0.

7.7 Photons in Other Bases

So far, we have chosen plane waves of a given ($\boldsymbol{\varepsilon}$, **k**) mode as our basis states. Such states are eigenstate of energy and momentum, and are single excitations of the number operator in the plane-wave basis. For simplicity, we now describe our plane-wave basis as a discrete basis labeled by modes *l*, where each *l* defines an ($\boldsymbol{\varepsilon}_l$, $\mathbf{k}_l$) pair. Then a single photon in mode *l* is written:

$$\left| N_l = 1 \right\rangle \equiv \left| \boldsymbol{\varepsilon}_l (\mathbf{k}_l) \right\rangle \qquad \text{(single-photon plane-wave state)} .$$

This is appropriate for a resonant cavity, where there exist only discrete modes. However, the same principles apply to both continuous and discrete bases [Lou ch. 6]. We follow the notation of [GAF], and a similar method.

All of the properties of the quantized EM field derive from three facts: (1) the commutator of the annihilation and creation operators for a single mode is the identity operator:

$$\left[\hat{a}_l , a_l^\dagger \right] = \mathbf{1}_{op} ;$$

(2) different modes don't interact, so:

$$\left[\hat{a}_l , \hat{a}_{l'} \right] = \left[\hat{a}_l^\dagger , \hat{a}_{l'}^\dagger \right] = \mathbf{0}_{op} ; \tag{7.3}$$

and (3) photons are bosons, which means the quantized EM field can have from 0 to ∞ photons in a single mode (i.e., in any given single-particle state).

These facts have been used for plane-wave modes, however, we saw above that a photon can exist as a single-photon wave-*packet*: a superposition of many modes. How would we "create" such a state with creation operators? We simply superpose the plane-wave component creation operators to create the wave-packet state (7.2). Define the single-photon wave-packet state as |*b*>. Then:

$$\left| b \right\rangle = \sum_{l=0}^{\infty} c_l \left| N_l = 1 \right\rangle = \left(\sum_l c_l \hat{a}_l^\dagger \right) \left| 0 \right\rangle \quad \Rightarrow \qquad \sum_l c_l \hat{a}_l^\dagger \equiv \hat{b}^\dagger ,$$

where $\hat{b}^\dagger$ is a creation operator for the wave-packet. More generally, we can construct an entire basis of orthogonal wave-packet states, |b_m>, where *m* denotes the wave-packet "mode." Each mode has its own set of coefficients c_l, so we replace the c_l with a notation for a unique set of coefficients for each *m*:

$$\left| b_m \right\rangle = \sum_l U_{ml} \left| N_l = 1 \right\rangle = \left(\sum_l U_{ml} \hat{a}_l^\dagger \right) \left| 0 \right\rangle \qquad \Rightarrow \qquad \sum_l U_{ml} \hat{a}_l^\dagger \equiv \hat{b}_m^\dagger .$$

U_{ml} is a unitary matrix, because it transforms from the *a* to *b* basis: given a column vector of *a*-mode coefficients, multiplying by U_{ml} gives the coefficients in the *b* basis. Similarly, we create *b*-mode annihilation operators:

$$\hat{b}_m \equiv \left(\hat{b}_m^\dagger \right)^\dagger = \sum_l U_{ml}^* \hat{a}_l .$$

Now what are the commutation relations of the *b*-mode creation and annihilation operators? Using the fact that different *a*-modes commute (7.3), that the commutator is linear in both arguments (1.8), and that every row of U_{ml} is a unit-magnitude vector:

$$\left[\hat{b}_m, \hat{b}_m^\dagger\right] = \left[\left(\sum_l U_{ml}^* \hat{a}_l\right), \left(\sum_{l'} U_{ml'} \hat{a}_{l'}^\dagger\right)\right] = \sum_l \left[U_{ml}^* \hat{a}_l, U_{ml} \hat{a}_l^\dagger\right] = \sum_l U_{ml}^* U_{ml} \left[\hat{a}_l, \hat{a}_l^\dagger\right] = \mathbf{1}_{op} .$$

Similarly, because the rows of U_{ml} are orthogonal, if $m \ne m'$:

$$\left[\hat{b}_m, \hat{b}_{m'}^\dagger\right] = \sum_l U_{ml}^* U_{m'l} \left[\hat{a}_l, \hat{a}_l^\dagger\right] = \mathbf{0}_{op} .$$

Thus the *b*-modes, the non-plane-wave modes, satisfy the same commutation relations as the plane-wave *a*-modes. Since all the number-basis properties derive from this commutation relation, there are then excitations in the *b*-number-basis with all the same number-basis properties. These excitations are eigenstates of the *b*-basis number operator $\hat{b}^\dagger \hat{b}$, and are therefore deserving of the name "photons." In fact, the *b*-mode number states are the multi-mode photon states.

When an atom radiates, we usually take the *a*-mode basis states as infinite spherical waves of definite *k*. However, the radiation is in fact a spherical wave-packet (a shell of radiation of finite thickness) radiating outward. Therefore, it may be considered a wave-packet superposition of infinite spherical waves. Equivalently, it may be considered a single excitation in a spherical wave-packet basis. In either basis, it is an eigenstate of the total photon operator, with eigenvalue 1; in other words, in either basis, it is a single photon.

7.8 The Wave-Function of a Photon?

If a photon is a particle, does it have a traditional wave-function? (Not the A-space "wave-function" of section 7.3) This is a very important question, because it provides the link between nonrelativistic quantum mechanics, and the more complete, relativistic quantum field theory (QFT). Because photons are massless, they are easily created and absorbed, unlike (say) electrons, which are more durable throughout interactions. It is the creation and absorption of photons, and their bosonic nature, that demands we use a quantum field theory, rather than ordinary QM. Another complication is that photons, being massless, cannot be at rest; they are always moving. And finally, photons are vector particles, somewhat more complicated than our nonrelativistic view of quantum particles with a scalar wave-function. Nonetheless, it is possible to define a useful "wave-function" for a single photon state, though "with some reservations" [GAF p384t]. Such a photon wave-function is useful in our later discussion of the quantum eraser.

We define our photon wave-function, $\psi_{EM}(\mathbf{r})$, to have the usual properties for a quantum particle, namely that the probability of detecting a photon is proportional to ψ_{EM}^2, therefore it is normalized, and it is complex-valued so that it produces the required interference:

$$\text{pdf}_\psi(\mathbf{r}) = \left|\psi_{EM}(\mathbf{r})\right|^2, \qquad \int_\infty \left|\psi_{EM}(\mathbf{r})\right|^2 d^3r = 1, \qquad \psi_{EM}(\mathbf{r}) \text{ complex} .$$

In fact, if normalized, the classical analytic (phasor) A-field has these properties. So our wave-function can be the *effective* A-field of the photon, but normalized.

A single photon state in a single mode is written |**ε**(**k**)>. One might think that we can simply find the average value of the A-field due to this photon in the usual QM way:

$$\left\langle \mathbf{A} \right\rangle = \left\langle \boldsymbol{\varepsilon}(\mathbf{k}) \right| \hat{\mathbf{A}} \left| \boldsymbol{\varepsilon}(\mathbf{k}) \right\rangle,$$

but we've already seen that this average is zero. The photon is a superposition of all phases, and each component of **A** is equally positive and negative. Instead, we already know the effective A-field, from **Error! Reference source not found.**:

$$\langle 0|\hat{\mathbf{A}}|\boldsymbol{\varepsilon}(\mathbf{k})\rangle = A_0(\mathbf{k})\boldsymbol{\varepsilon}(\mathbf{k})e^{+i(\mathbf{k}\cdot\mathbf{r}-\omega t)} .$$

It is the coupling of the EM state by $\hat{\mathbf{A}}$ *to the vacuum* that gives the effective A-field. We define our wave-function by simply replacing the amplitude $A_0(\mathbf{k})$ and polarization $\boldsymbol{\varepsilon}(\mathbf{k})$ with a normalization constant:

$$\psi_{EM}(\mathbf{r}) = Ne^{+i(\mathbf{k}\cdot\mathbf{r}-\omega t)}$$

Since this is an infinite plane-wave state, we have the same options for normalizing it as we did for incident plane waves in scattering (section 2.5).

A more realistic single photon state is the multi-mode state, (7.2). We construct our wave-function the same way as for a single-mode photon, by coupling the EM state to the vacuum (and assuming all the $\boldsymbol{\varepsilon}_l$ are the same, and dropping them):

$$\psi_{EM}(\mathbf{r}) = N\langle 0|e^{+i(\mathbf{k}_l\cdot\mathbf{r}-\omega_l t)}\hat{a}_l|\psi_{EM}\rangle = N\sum_l c_l\langle 0|e^{+i(\mathbf{k}_l\cdot\mathbf{r}-\omega_l t)}\hat{a}_l|N_l = 1\rangle$$

$$\textit{where}\quad N \text{ is chosen such that } \int_\infty |\psi_{\text{EM}}(\mathbf{r})|^2\, d^3r = 1 .$$

Because the wave-packet is finite in extent, we have used ordinary wave-function normalization.

The effective A-field of a photon, as defined by its “wave-function,” follows Maxwell’s equations.

So all the classical laws of radiation (e.g., reflection, diffraction) apply, even to a single photon.

Like any quantum particle, a single photon can be spread out over an arbitrarily large area. In my graduate research, a single photon (loosely speaking) returning from the moon is spread over many square kilometers. Our telescope’s ~10 m^2 cross section then has only a small probability of detecting such a photon, because ultimately the photon detection is a quantized event: either we detect it, or we don’t. A detailed consideration of “detection” leads us to the interaction of photons (and radiation in general) with matter.

This method of defining a wave-function from a field excitation by coupling it to the vacuum is fairly general, and can be used (with similar reservations) for many other kinds of particles [Sak2 3.376 p147b], [Bay p422], [Wei 2013 p. xvi top], [P&S p24b].

7.9 Quasi-Classical States

The closest quantum analog of a classical field is called a **quasi-classical state**, or a “coherent state,” or a “Glauber state,” after Nobel laureate Roy J. Glauber who first described them in detail [Phys. Rev. 130, 2529–2539 (1963)]. The *average* field of a quasi-classical state is a classical EM field. For example, the long-time limit of a classical oscillating current produces an EM field that is a quasi-classical state, as does a strong-beam laser [Lou p190m].

Let us construct a quasi-classical state. Consider a single-mode classical wave that is essentially infinite in space. In the usual complex notation, the wave is given by:

$$\mathbf{A}(t,\mathbf{r}) = 2\,\mathrm{Re}\left[\alpha A_0\boldsymbol{\varepsilon}e^{+i(\mathbf{k}\bullet\mathbf{r}-\omega t)}\right] \qquad \textit{where}\quad A_0 = \sqrt{\frac{2\pi\hbar c}{(Vol)|\mathbf{k}|}} ,$$

where we have written the phasor for the wave as αA_0 for later convenience, with α a complex number and A_0 taken as real. We seek a normalized quantum state $<\alpha|$ which reproduces the above classical wave as closely as possible. Therefore, its average A-field equals the classical wave:

$$\langle\mathbf{A}(t,\mathbf{r})\rangle = \langle\alpha|\hat{\mathbf{A}}|\alpha\rangle = 2\,\mathrm{Re}\left[\alpha A_0\boldsymbol{\varepsilon}e^{+i(\mathbf{k}\bullet\mathbf{r}-\omega t)}\right] . \tag{7.4}$$

$\hat{\mathbf{A}}$ is hermitian, consisting of two terms proportional to $\hat{a}$ and $\hat{a}^\dagger$. If we make the average of the $\hat{a}$ term equal to the complex-valued signal (the bracketed quantity above), then the hermiticity of $\hat{\mathbf{A}}$ guarantees the classical correspondence (7.4). Thus:

$$\langle\alpha|A_0\boldsymbol{\varepsilon}e^{+i(\mathbf{k}\cdot\mathbf{r}-\omega t)}\hat{a}|\alpha\rangle=\alpha A_0\boldsymbol{\varepsilon}e^{+i(\mathbf{k}\cdot\mathbf{r}-\omega t)} \qquad \textit{where} \quad \hat{a}\equiv\hat{a}_{\boldsymbol{\varepsilon}}(\mathbf{k})\,. \tag{7.5}$$

(We used the fact that only one mode is present to eliminate from $\hat{\mathbf{A}}$ the sum over λ, and the integral over **k**.) Now recall that $\hat{a}$ acts only on the number part of $|\alpha>$, and not on the spacetime part. Therefore (7.5) becomes:

$$\langle\alpha|A_0\boldsymbol{\varepsilon}e^{+i(\mathbf{k}\cdot\mathbf{r}-\omega t)}\hat{a}|\alpha\rangle=A_0\boldsymbol{\varepsilon}e^{+i(\mathbf{k}\cdot\mathbf{r}-\omega t)}\langle\alpha|\hat{a}|\alpha\rangle=\alpha A_0\boldsymbol{\varepsilon}e^{+i(\mathbf{k}\cdot\mathbf{r}-\omega t)} \quad\Rightarrow\quad \langle\alpha|\hat{a}|\alpha\rangle=\alpha\,. \tag{7.6}$$

The simplest way to satisfy the last equality is:

$$\hat{a}|\alpha\rangle=\alpha|\alpha\rangle \qquad \text{(quasi-classical state)}\,.$$

In fact, if different modes $\boldsymbol{\varepsilon}(\mathbf{k})$ are to remain independent (not intermix in $<\alpha|\hat{\mathbf{A}}|\alpha>$), then this eigenvalue equation is the only one allowed. Thus we have the general form for a quasi-classical state: $|\alpha>$ is an eigenstate of the lowering operator $\hat{a}$.

We write $|\alpha>$ in the number basis, with as-yet undetermined coefficients, b_N:

$$|\alpha\rangle\equiv\sum_{N=0}^{\infty}b_N|N\rangle \qquad \text{for a given mode, } \boldsymbol{\varepsilon} \text{ and } \mathbf{k}\,.$$

The eigenvalue α is a given, but arbitrary, complex number, and is used to label the state $|\alpha>$. We now show that α alone determines the state. Recall that α is proportional to the phasor for the average value of the EM field in the quasi-classical state $|\alpha>$.

First, we find a recurrence relation for the coefficients in the superposition, b_N, using the fact that $|\alpha>$ is an eigenstate of $\hat{a}$:

$$\hat{a}b_N|N\rangle=\sqrt{N}b_N|N-1\rangle=\alpha b_{N-1}|N-1\rangle \qquad\Rightarrow\qquad b_N=\frac{\alpha b_{N-1}}{\sqrt{N}}\,.$$

Thus $|\alpha>$ is a superposition of all photon occupation numbers of a quantum state, from 0 to ∞. By induction, the entire state is determined by no more than b_0, the first coefficient, and the eigenvalue α, since (by induction):

$$b_N=\frac{\alpha^N}{\sqrt{N!}}b_0\,.$$

However, this general form determines the normalization, and $|\alpha>$ is normalized. Then:

$$\sum_{N=0}^{\infty}|b_N|^2=\sum_{N=0}^{\infty}\frac{|\alpha|^{2N}}{N!}|b_0|^2=\exp\left(|\alpha|^2\right)|b_0|^2=1 \qquad\Rightarrow\qquad |b_0|=\exp\left(-|\alpha|^2/2\right).$$

Therefore, choosing our phase such that b_0 is real, the entire state is defined by the complex eigenvalue α:

$$|\alpha\rangle\equiv\exp\left(-|\alpha|^2/2\right)\sum_{N=0}^{\infty}\frac{\alpha^N}{\sqrt{N!}}|N\rangle \qquad \text{(quasi-classical state)}\,.$$

As a check on our derivation, we compute the average A-field for $|\alpha>$ in the usual way, with an inner product of $\hat{\mathbf{A}}$, using $\hat{a}|\alpha\rangle=\alpha|\alpha\rangle$ and its adjoint, $\langle\alpha|\hat{a}^\dagger=\alpha^*\langle\alpha|$:

$$\begin{aligned}\langle \mathbf{A}(t,\mathbf{r})\rangle &= \langle\alpha|\hat{\mathbf{A}}|\alpha\rangle = \langle\alpha|\left[A_0\boldsymbol{\varepsilon}e^{+i(\mathbf{k}\bullet\mathbf{r}-\omega t)}\hat{a} + A_0\boldsymbol{\varepsilon}^* e^{-i(\mathbf{k}\bullet\mathbf{r}-\omega t)}\hat{a}^\dagger\right]|\alpha\rangle \\ &= A_0\boldsymbol{\varepsilon}e^{+i(\mathbf{k}\bullet\mathbf{r}-\omega t)}\langle\alpha|\hat{a}|\alpha\rangle + A_0\boldsymbol{\varepsilon}^* e^{-i(\mathbf{k}\bullet\mathbf{r}-\omega t)}\langle\alpha|\hat{a}^\dagger|\alpha\rangle \qquad . \\ &= \alpha A_0\boldsymbol{\varepsilon}e^{+i(\mathbf{k}\bullet\mathbf{r}-\omega t)} + \alpha^* A_0\boldsymbol{\varepsilon}^* e^{-i(\mathbf{k}\bullet\mathbf{r}-\omega t)} = 2\,\mathrm{Re}\left[\alpha A_0\boldsymbol{\varepsilon}e^{+i(\mathbf{k}\bullet\mathbf{r}-\omega t)}\right].\end{aligned}$$

As expected, this reproduces the classical correspondence (7.4) above.

Further properties of the quasi-classical state: The eigenvalue α is arbitrary, so the *average* A-field amplitude is any one of a *continuum* of amplitudes, i.e. it is *not* quantized. Also, $|\alpha>$ is *not* an eigenstate of the photon number operator $\hat{N}$. Therefore, the number of photons in a quasi-classical state is uncertain. This may be surprising, since it is often incorrectly assumed that a general EM state must comprise an integer number of photons.

A general quantum EM state may have an uncertain number of photons in it,
and the average number of photons in a state can be fractional.

The average A-field amplitude is αA_0, where $A_0 \equiv A_0(\mathbf{k})$ is the average A-field amplitude of a single-photon state $|\boldsymbol{\varepsilon}(\mathbf{k})>$. Therefore, α is dimensionless. A quasi-classical state is *not* an eigenstate of the hamiltonian, and therefore has uncertain energy. The energy of a single photon is $\hbar\omega$, and the average energy of the EM state $|\alpha>$ is (relative to the vacuum energy):

$$\hat{H} = \hbar\omega\hat{a}^\dagger\hat{a} \qquad \Rightarrow \qquad \langle E_\alpha\rangle = \hbar\omega\langle\alpha|\hat{a}^\dagger\hat{a}|\alpha\rangle = \hbar\omega|\alpha|^2 .$$

So $|\alpha|^2$ can be thought of as the average number of photons in the state $|\alpha>$, which need not be an integer.

Quasi-classical states are also states of minimum uncertainty: the EM field amplitude is as precisely defined as quantum mechanics allows.

Possibly the most widespread misunderstanding about quantized EM fields is
the belief that they must comprise an integer number of photons.

7.10 Classical EM Superposition vs. Quantum Superposition: Interference of Distinct Photons?

Classical superposition of EM fields is a somewhat different concept than quantum superposition of states and kets. Classical superposition (aka interference) simply says that if two (or more) sources create EM fields at the same point in spacetime, the total field is the vector sum of the component fields:

$$\mathbf{A}_{tot}(t,\mathbf{r}) = \mathbf{A}_1(t,\mathbf{r}) + \mathbf{A}_2(t,\mathbf{r}) , \qquad \text{and similar for } \mathbf{E}(t,\mathbf{r}) \text{ and } \mathbf{B}(t,\mathbf{r}).$$

Since *independent* classical EM fields interfere, and because all classical results are explained by quantum mechanics, it must be true that independent *quantized* EM fields also interfere.

Dirac made a famous, though widely misunderstood, statement that "... each photon only interferes with itself. Interference between different photons never occurs" [Dir p9]. We show here the mechanism by which independent quantum EM fields *do* interfere. (If we say, loosely, that a quantum EM field comprises "photons", then classical EM interference forces us to accept that distinct photons do, in fact, interfere. However, see the earlier discussion on the uncertainty of photon count in a quantum EM state, sec 7.8.)

To study independent photon interference, Pfleegor and Mandel performed a famous experiment in 1967 [P&M] where two independent lasers were attenuated to such a low intensity that the probability of detecting two photons at the same time was negligible. Nonetheless, the two lasers produce a distinct interference pattern, thus confirming that two "photons" that can reasonably be considered independent do, in fact, interfere.

The quantum explanation for both classical superposition, and the Fleegor-Mandel results, is found in the full $\hat{\mathbf{A}}$ operator already developed **Error! Reference source not found.**:

$$\hat{\mathbf{A}}(t,\mathbf{r}) = \sum_{\lambda=1}^{2}\int_{-\infty}^{\infty} d^3\mathbf{k}\; A_0(\mathbf{k}) \left(\underbrace{\boldsymbol{\varepsilon}_\lambda(\mathbf{k}) e^{+i(\mathbf{k}\cdot\mathbf{r}-\omega t)}}_{\text{spacetime}} \underbrace{\hat{a}_\lambda(\mathbf{k})}_{\text{photon}} + \underbrace{\boldsymbol{\varepsilon}^*_\lambda(\mathbf{k}) e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega t)}}_{\text{spacetime}} \underbrace{\hat{a}^\dagger_\lambda(\mathbf{k})}_{\text{photon}} \right).$$

The effective A-field is the coherent sum of all the modes. The resulting A-field becomes part of an inner product, whose magnitude is squared to produce a measurable result. Thus the A-field follows the usual quantum rule: first sum coherently, then square the magnitude.

We see, then, that the Fleegor-Mandel interference is essentially classical interference. As noted in the section on quasi-classical states, the average amplitude of the field is *not* quantized, and is therefore essentially classical. It is the *interaction* of the EM field with matter that is quantized (through the raising and lowering operators in $\hat{\mathbf{A}}$). In the experiment, even though the probability of detecting a given photon is low, the EM field is still excited to some level, and interference occurs.

We can now consider Dirac's statement in a new light. Recall that a single photon is *incoherent* (a superposition of all phases), i.e. it has no definite phase. Therefore, two such photons have no definite phase relationship, and cannot interfere. However, a single photon (even though incoherent) *can* interfere with itself, because each *component* of the superposition *has* a definite phase. In the two-slit experiment, each component reinforces and cancels in the same way, so the complete wave-function (which is the sum of its components) exhibits interference. In contrast:

> Quasi-classical states describe fields with a well-defined sinusoidal phase of small uncertainty, and therefore, independent sources can interfere in essentially the classical way.

This is in contrast to single-photon states, which are a superposition of all phases, so independent single-photon sources cannot interfere.

In summary: A significant point of confusion about Dirac's statement concerns the meaning of the word "photon." The naive view that EM fields comprise a definite number of "photons" is incorrect. EM fields are not quantized like electrons, or marbles. Any single photon state is incoherent, meaning a superposition of all phases. The more complete theory of the quantized EM field defines quasi-classical states, which interfere in an essentially classical way, through the summation over modes in the $\hat{\mathbf{A}}$ operator. However, even in such quasi-classical states, the EM field interactions with matter are quantized in units we call "photons." (We note in passing that the general results of our development apply to all boson fields, not just photons [GAF p384b].)

7.11 Don't Be Dense

Density of states is a very important concept for radiation, and other particle creation/annihilation operations. Density of states is often used, and much abused. The term "density of states" comes in many forms, but in quantum mechanics, "density of states" always refers to number density of stationary quantum states per unit of some parameter space. Recall that stationary quantum states are energy eigenstates. We consider these common densities of states:

- 1D harmonic oscillator
- 1D particle in a box
- 1D free particle
- 3D free particle

We consider only the density of spatial states (wave-functions). Fermions, e.g. electrons, would have twice as many states, because each spatial state has 2 spin-states. Similarly, photon states are twice as dense, because each spatial state has 2 polarization states.

Our analysis refers to 3 different "spaces": real physical space (1D, 2D, or 3D), k-space, and phase-space. Since $p = \hbar k$, we take momentum-space to be essentially equivalent to k-space. Phase-space is a "tensor product" of real-space and momentum-space:

$$V_{\text{phase-space}} = (Vol)(V_{\text{momentum-space}}) \quad \Rightarrow \qquad dV_{\text{phase-space}} = d^3r\; d^3p\, .$$

1D harmonic oscillator: Perhaps the simplest density of states is that for a 1D harmonic oscillator. We write the number of states per unit energy as $\rho(E)$. Since the energy spacing is constant, $\hbar\omega$, the number of states per unit energy is (Figure 7.3, left):

$$\rho(E) = \frac{1}{\hbar\omega} \qquad \text{(1D harmonic oscillator)}\, .$$

The units are $[E]^{-1}$, i.e. inverse energy. The momentum (and therefore wave vector, k) of the particle in any stationary state is uncertain, and does not figure into the density of states.

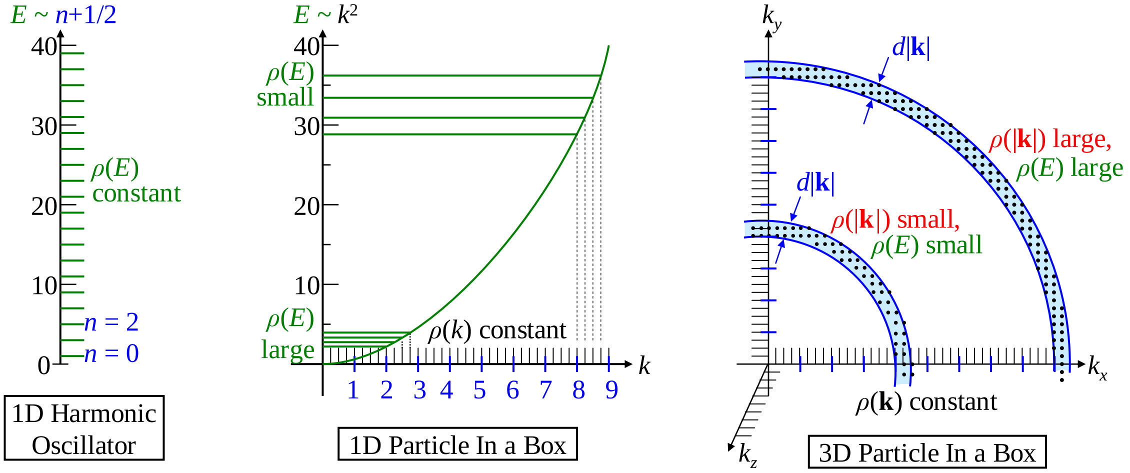


Figure 7.3 Densities of states.

1D particle in a box: Another simple density of states is 1D particle in a box. We will see that the result *per unit energy* is the same as for a free particle, and use this as a warm up for 3D. Each state has a purely real wave-function, which must be 0 at the boundaries. This leads to quantized wave-functions (ignoring normalization):

$$\psi(x) = \sin(kx) = \frac{1}{2i}\left(e^{ikx} - e^{-ikx}\right) \qquad \textit{where} \quad k = n\pi / L, \;\; L = \text{length of box}, \;\; n = \text{quantum number}\, .$$

Each energy eigenstate is a superposition of + and – momentum eigenstates. Because k is easily quantized (an integer number of ½ wavelengths: $kL = n\pi$), we can write the density of states parameterized by k. At first, we let k take only positive values, knowing that each positive value of k describes a state which is a standing-wave superposition of both $+k$ and $–k$ momenta. We wrote k as a function of n above, but for a density of states, we must invert this relation to get number density per unit of positive k:

$$n(+k) = \frac{kL}{\pi} \qquad \Rightarrow \qquad \rho_+(k) = \frac{dn}{dk} = \frac{L}{\pi}, \qquad \text{in units of } [\text{rad/m}]^{-1} = [\text{m}]\, . \tag{7.7}$$

But taking into account that each state consumes a $+k$ and $–k$ value, the final (two-sided) density per unit k (of either sign) is only half the above:

$$\rho(k) = \frac{1}{2}\frac{dn}{dk} = \frac{L}{2\pi} \qquad \text{(1D particle in box)}\, . \tag{7.8}$$

Thus, for a given size box, the density of energy eigenstates per unit k is constant. Note that there is no upper bound to k (or E).

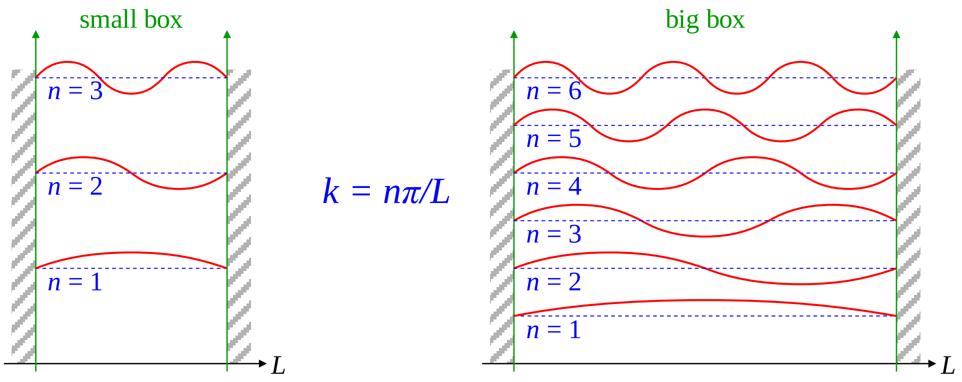


Figure 7.4 Density of states, $\rho(k)$, for particle-in-a-box is proportional to the box size.

(excerpted)

7.12 Perturb Unto Others: Fermi's Golden Rule

Fermi's Golden Rule often first comes up when considering interaction of matter with EM fields. This section assumes you've been through the derivation in a standard text [Sak p327+] [Bay p248+], and want clarification.

For an electromagnetic transition, Fermi's Golden rule allows us to compute the transition rate, in transitions/s. The rule is often confusingly written with a δ-function, such as (for absorption):

$$R = \frac{2\pi}{\hbar}\left|\langle f|\hat{H}_{\text{int}}|i\rangle\right|^2 \delta(E_f - E_i - \hbar\omega) \qquad \text{(not rigorous)}$$

$$\begin{aligned} \textit{where}\quad & |i\rangle \equiv \text{initial state;} \quad \langle f| \equiv \text{final state,} \\ & \hat{H}_{\text{int}} \text{ is the perturbing interaction hamiltonian,} \\ & E_f \equiv \text{final energy of matter,} \quad E_i \equiv \text{initial energy of matter,} \\ & \hbar\omega \equiv \text{energy of incident photon}. \end{aligned}$$

This is clearly requires further description: a "rate" is a number, and cannot have a delta function in it. In fact, the above formula only becomes a rate when it is integrated over a density of states, which removes the δ-function ("...in all practical applications, the δ-function will get integrated over...." [Sha p483b]; "...to get actual numbers from this formula we must sum over a continuous group of ... states..." [Bay p251m]). What the δ-function above means is that the probability is very high that the final state will lie in a narrow range of energies. This narrow range insures energy conservation, and so the δ-function is sometimes called an **energy-conserving δ-function**. Therefore, we should rewrite Fermi's Golden Rule as:

$$\begin{aligned} R &= \frac{2\pi}{\hbar}\left|\langle f|\hat{H}_{\text{int}}|i\rangle\right|^2 \int \delta\left(E_f - E_i - \hbar\omega\right)\rho\left(\textit{something}\right) dE \\ &= \frac{2\pi}{\hbar}\left|\langle f|\hat{H}_{\text{int}}|i\rangle\right|^2 \rho\left(\textit{something}\right), \end{aligned}$$

but we must be careful to describe what is ρ(*something*). In fact, this density of states can appear due to any *one* of the 3 entities in the matrix element $\langle f|\hat{H}_{\text{int}}|i\rangle$: the density of initial states, ρ_i; density of final states, ρ_f; and in the interaction Hamiltonian, H_{int}. Figure 7.5 gives examples of all three cases. The left case is probably the most common: photoionizing an atom moves an electron from a discrete energy state to a continuum of states. The middle case excites one electron from a continuous band of electrons (e.g. in a valence band) into a single state above the band (e.g. due to an impurity atom). The right case excites an electron from a lower to a higher energy state, with radiation having a photon density $\rho(E)$.

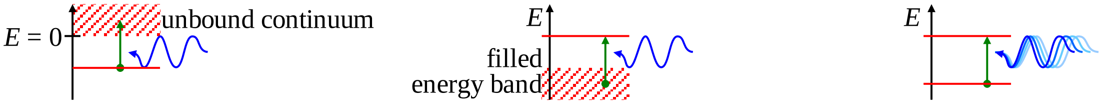


Figure 7.5 (Left) Density of final state. (Middle) Density of initial states. (Right) Density of $\hat{H}_{int}$.

(excerpted)

8 Desultory Topics in Quantum Mechanics

8.1.1 Parity

A parity transformation of a system reflects each point through the origin. It is equivalent to 3 reflections in 3 mirrors: $x \to -x$, $y \to -y$, $z \to -z$, or $\mathbf{r} \to -\mathbf{r}$. You might think that you could achieve that effect with one mirror, if you held it at the proper angle. We see that is not true by supposing such a mirror placement exists, and considering what happens to a vector nearly parallel to the plane of the mirror. Its reflection is barely different than the original vector, and clearly *not* reflected through the origin. You must use 3 mirrors.

All the laws of physics, except the weak force,
are invariant under a parity transformation of the system.

8.1.2 Massive Particle Frequency

Unlike a photon, a massive or charged particle frequency has an *arbitrary* reference point. Therefore, absolute frequency of a massive particle has no physical meaning. Only *differences* in frequency are meaningful. Recall that particle frequency is proportional to *total* energy, kinetic (T) + potential (V):

$$E = T + V = \hbar\omega,$$

but V has an *arbitrary* reference point. In particular, E and therefore ω, can be negative. Negative frequency arises in many areas of physics, engineering, and mathematics, especially when the zero-point of a frequency is arbitrary. Negative frequency is essential to quantum mechanics and QFT, but what is negative frequency?

An ordinary frequency counts oscillations of something, per unit time, which is necessarily non-negative. The oscillating parameter is some one-dimensional quantity. As described in section 1.6.2, to allow for negative frequency, we imagine something rotating about an axis (rather than just oscillating). If it rotates clockwise, physicists call it positive frequency. In this way, energy differences are proportional to frequency differences.

Note that phase velocity $v_p = \omega/k$, and thus v_p also has no physically meaningful absolute zero point. In contrast, the group velocity, $v_g = d\omega/dk$, is independent of an arbitrary frequency (or energy) offset. This is consistent with its interpretation as the physically measurable particle velocity.

The above frequency considerations are in direct contrast to a photon, or other possible massless-chargeless particles, where ω has an absolute value, and $E = \hbar\omega$ is an absolute equation for the total energy of the particle. v_p then also has an absolute value.

Note also that massive particles *do* have an absolute wavelength, which can be measured by diffraction. The absolute wavelength is given by the kinetic momentum, and is gauge invariant. However, the spatial frequency k that appears in the wave-function is gauge dependent. In other words, the wave-function itself is gauge dependent, but all observables are, as in all of physics, gauge invariant.

8.1.3 Uncertainty Isn't From Measurement Clumsiness

The uncertainty principle is often misleadingly described as: "You can't simultaneously know the position and momentum of an object, because when you measure its position, you disturb its momentum, and when you measure its momentum, you disturb its position." This description belies the quantum theory, because it implies that a particle *has* a well-defined position and momentum, but you just can't know what they both are. In fact, a particle has a wave function, which gives probability densities for all possible values of position and momentum. Due to the nature of the position and momentum operators, which act on the wave function to produce the probability density functions (position PDF and momentum PDF), it is impossible to construct a wave function which yields arbitrarily narrow ranges for both position and momentum. Therefore:

Uncertainty arises not because you can't *know* both position and momentum,
but because a particle cannot *have* precise values of both position and momentum.

One significant consequence of this more accurate statement of uncertainty is this: don't bother looking for "gentler" measurement methods, which disturb the particle *less* than existing measurements, in the hopes of beating Heisenberg. The "disturbance" of a previously known value is not some measurement side effect; it is fundamental to the nature of wave functions and measurements.

Now the measurement postulate *does* say that when you measure position, you "disturb" momentum, because you collapse the wave function to one of a precise position. But if you make a precise position measurement, you don't change the momentum to some precise, but unknown, value. When you make a precise position measurement, you change the wave function, making the position PDF narrow, and the momentum PDF wide. The problem is not that now you don't *know* the new momentum; the problem is that now the particle does not *have* a definite momentum. The particle is in a superposition of many different momenta.

8.1.4 Generalized Uncertainty

In general, the uncertainty product of two observables depends on the state of the system. Many people are confused by this, because in the special case of position-momentum uncertainty, the uncertainty product happens to be a constant for any state. However:

In the general case, the uncertainty product depends on the state of the system.

Consider a set of systems in identical quantum states, each of which is the same superposition of eigenstates of an observable. Recall that if we measure this observable for every system, then we will get a variety of measurements, and they will have some standard deviation, σ. It is easy to show that, if $\hat{A}$ and $\hat{B}$ are two non-commuting observables, then in the state $|\psi\rangle$:

Using: $\left\langle \hat{O} \right\rangle = \langle \psi | \hat{O} | \psi \rangle$

$$\sigma_A{}^2 \sigma_B{}^2 \ge \frac{1}{4} \left| \langle \psi | \left[\hat{A}, \hat{B} \right] | \psi \rangle \right|^2 + \frac{1}{4} \left| \langle \psi | \left\{ \hat{A}, \hat{B} \right\} | \psi \rangle \right|^2 \qquad \text{[Sak 1.4.63, p 36]}$$

$$\textit{where} \quad \left[\hat{A}, \hat{B} \right] \equiv \left(\hat{A}\hat{B} - \hat{B}\hat{A} \right); \quad \left\{ \hat{A}, \hat{B} \right\} \equiv \left(\hat{A}\hat{B} + \hat{B}\hat{A} \right).$$

Since the 2nd term in the inequality is ≥ 0, the inequality still holds without it, so sometimes people abbreviate:

$$\sigma_A{}^2 \sigma_B{}^2 \ge \frac{1}{4} \left| \langle \psi | \left[\hat{A}, \hat{B} \right] | \psi \rangle \right|^2 \qquad \text{[Sak 1.4.53, p 35]},$$

or by simply taking square roots:

$$\sigma_A \sigma_B \ge \frac{1}{2} \left| \langle \psi | \left[\hat{A}, \hat{B} \right] | \psi \rangle \right|.$$

For example, the uncertainty product of the *x* and *y* components of angular momentum is zero in the state where the total angular momentum is zero, $|J = 0, M = 0\rangle$. However, the same uncertainty product is non-zero for states with $M \ne 0$. Recalling that $[\hat{J}_x, \hat{J}_y] = i\hbar\hat{J}_z$:

$$\sigma_x{}^2 \sigma_y{}^2 \ge \frac{1}{4} \left| \langle J, M | \left[\hat{J}_x, \hat{J}_y \right] | J, M \rangle \right|^2 = \frac{1}{4} \left| \langle J, M | i\hbar\hat{J}_z | J, M \rangle \right|^2 = \frac{\hbar^2}{4} M^2 \left| \langle J, M | J, M \rangle \right|^2 = \frac{\hbar^4}{4} M^2.$$

The uncertainty product increases as the *z* component of angular momentum increases. The uncertainty in measurements of the system depends on the state of the system.

Back to position-momentum uncertainty: it happens that the commutator of $\hat{x}$ and $\hat{p}$ is a constant for all states, hence the uncertainty product is a constant for any state $|\psi\rangle$:

$$\sigma_x \sigma_p \ge \frac{1}{2} \left| \langle \psi | i\hbar | \psi \rangle \right| = \frac{\hbar}{2} \left| \langle \psi | \psi \rangle \right| = \frac{\hbar}{2}.$$

Finally, there is often talk of "energy-time" uncertainty. However, though energy *is* an observable, time *is not an observable*. Time is a parameter; time marches on; we know what time it is, but we don't measure the "time" of a system (as we do its energy); time has no corresponding hermitian operator. Therefore:

> The energy-time "uncertainty relation" is fundamentally a
> different phenomenon than the uncertainty product of two observables.

According to [Sak p80], "this time-energy uncertainty relation is of a very different nature from the uncertainty relation between two incompatible observables."

8.1.5 Is ∇^2 The Square Or Composition Of Anything?

"∇^2" denotes the Laplacian operator, sometimes called "grad squared." Usually, the superscript ^{2} notation means either the square of a number, or the composition of an operator on itself. For example, x^2 means simply the square of the number 'x'. $\hat{p}^2$ means $\hat{p}\hat{p}$. However, contrary to such uses of the superscript ^{2} notation, ∇^2 is neither the square of a number, nor the composition of some ∇ operator on itself. ∇^2 is actually the composition of two *different* "∇" operators. We consider those operators now.

First, "$\boldsymbol{\nabla}$" is the **gradient** operator. It operates on a scalar field, and produces a vector field. For example, if $\Phi(x, y, z)$ is the electric potential (a scalar field), $\mathbf{E} = -\boldsymbol{\nabla}\Phi(x, y, z)$ is the electric (vector) field. Recall that the gradient $\boldsymbol{\nabla}\Phi$ (pronounced "grad fie" or "del fie") tells how Φ varies when moving in any direction from a point:

$$\Delta\Phi(\text{in direction } \mathbf{s}) \approx \boldsymbol{\nabla}\Phi \cdot \mathbf{s} \qquad \text{(to first order).}$$

The gradient operator is often written in bold, "$\boldsymbol{\nabla}$"**,** indicating its result is a vector.

Second, "$\nabla\cdot$" is the **divergence** operator (pronounced "del dot"). It operates on a vector field, and produces a scalar field. The divergence operator is never written in bold, because its result is a scalar field.

The divergence of the gradient is a composition of operators that arises very frequently. It can be written "$\nabla\cdot\boldsymbol{\nabla}$", e.g.

$\nabla\cdot\boldsymbol{\nabla}\Phi$ = the divergence of the gradient of the scalar field Φ, whose result is another scalar field.

This composition arises so frequently that we use the special notation "∇^2" as shorthand for "$\nabla\cdot\boldsymbol{\nabla}$". But the two "dels" in "$\nabla\cdot\boldsymbol{\nabla}$" are two different things: the right (bold) $\boldsymbol{\nabla}$ is the gradient operator, and the left "$\nabla\cdot$" is the divergence operator. So:

> ∇^2 is not the composition of any operator on itself.
> It's inconsistent notation, but universally used.

Note that (in simple non-tensor mathematics) it is meaningless to square either the gradient or the divergence operators:

$\boldsymbol{\nabla}(\boldsymbol{\nabla}\varphi)$ has the left gradient operator acting on a vector field, which is undefined.

$\nabla\cdot(\nabla\cdot\mathbf{v})$ has the left divergence operator acting on a scalar field, which is undefined.

All that said, there is a sense in which ∇^2 is roughly the square of *something*. It is common to write the square of a vector $\mathbf{p}$ as $p^2 \equiv \mathbf{p}\cdot\mathbf{p}$. Following this idea, we write the definition of the gradient operator $\boldsymbol{\nabla}$ as:

$$\nabla \equiv \left(\frac{\partial}{\partial x}\mathbf{i} + \frac{\partial}{\partial y}\mathbf{j} + \frac{\partial}{\partial z}\mathbf{k}\right) \qquad \Rightarrow$$

$$\nabla^2 \equiv \nabla\cdot\nabla = \left(\frac{\partial}{\partial x}\mathbf{i} + \frac{\partial}{\partial y}\mathbf{j} + \frac{\partial}{\partial z}\mathbf{k}\right)\cdot\left(\frac{\partial}{\partial x}\mathbf{i} + \frac{\partial}{\partial y}\mathbf{j} + \frac{\partial}{\partial z}\mathbf{k}\right) = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\,.$$

In this loose interpretation, ∇^2 could be called the square of the gradient operator, but it is not the composition of the gradient operator with itself (which can only be defined as a rank-2 tensor).

All of these comments on ∇^2 apply directly to the 3D momentum operator:

$$\hat{\mathbf{p}} \equiv \frac{\hbar}{i}\nabla \quad \Rightarrow \qquad \hat{\mathbf{p}}^2 \equiv \hat{\mathbf{p}}\bullet\hat{\mathbf{p}} = -\hbar^2\nabla^2 ,$$

which is a *scalar* operator.

8.1.6 The Only Stationary States Are Eigenstates of Energy

The only stationary states are eigenstates of energy, because every state evolves in time by the operation of exp($i\hat{H}t/\hbar$), and the only states of fixed energy are eigenstates of $\hat{H}$. That is, for $\hat{H}$ independent of time:

$$\hat{U}\left(t,t_0\right) = \exp\left(i\hat{H}\left(t-t_0\right)/\hbar\right).$$

For a stationary state, time evolution must introduce only a (time-dependent) complex phase factor to the wave-function, which means $\hat{H}$ must produce a constant. This is the definition of an energy eigenstate:

$$\hat{H}\left|\psi\right\rangle = E\left|\psi\right\rangle .$$

Thus the only stationary states are energy eigenstates.

8.1.7 g Whiz

The letter "*g*" has at least 3 different meanings in quantum mechanics, the first two of which are especially confusing.

First, when counting quantum states in statistical mechanics, the multiplicity of an energy level, or multiplicity of a spatial state, is called "*g*". For an electron, or any other fermion, its spin is ½, and therefore, in the absence of spin-dependent energy (e.g., no magnetic field), each spatial state has multiplicity $g = 2$ *exactly*: one state spin up, the other spin down.

A second (completely different) use of "*g*" refers to the gyromagnetic ratio: the ratio of the electron's spin-related magnetic dipole moment to its Bohr magneton. [Other disciplines use the term "gyromagnetic ratio" differently.] Recall the Bohr magneton is the magnetic dipole moment created by an electron in an *orbit* of angular momentum $\hbar$. One can easily show that the Bohr magneton is:

$$\mu_B = \frac{e\hbar}{2mc} \quad \text{(gaussian)} \qquad \text{or} \qquad \mu_B = \frac{e\hbar}{2m} \quad \text{(SI)} \quad \textit{where} \quad e \equiv \text{particle charge}, \; m \equiv \text{particle mass} .$$

For an electron, g ≈ –2, but *not exactly*. In fact, $g_e = -2.002\ 319\ 304\ 362\ 2(15)$ [Nis]. And therefore, the actual spin-related dipole moment for an electron is $\mu_e \approx -2.002\mu_B$. For a general particle, $\mu = g\,\mu_{magneton}$.

When referring to the spin-multiplicity of a spatial state, $g = 2$ *exactly*.
When referring to gyromagnetic ratio, $g \approx -2$.

The third use of *g* extends the gyromagnetic ratio to the total angular momentum, **J** = **L** + **S**, which is a combination of orbital and spin angular momentum. This is called the **Landé g-factor** (described in detail in sec. 5.8). Part of **J** is due to **L**, and part due to **S**. Therefore, part of the magnetic dipole moment is due to **L** (with weight 1), and part due to **S** (with weight approximately 2). Thus *g* is a weighted average such that:

$$\mu = g m_j \mu_B \qquad \textit{where} \quad m_j \text{ is the } J_z \text{ quantum number} .$$

8.1.8 Why Are Photons Said To Be Massless?

Photons have energy, and $m = E/c^2$, so why are photons said to be massless? In the old days, to try to retain $F = ma$, physicists said that a moving particle's mass increased by a factor:

$$\gamma \equiv \frac{1}{\sqrt{1-(v/c)^2}} .$$

This “variable mass” turned out to be a bad idea, because the “mass” parallel to the direction of motion is then different than the “mass” perpendicular to the motion. To avoid this, and have only one kind of mass, modern relativity defines mass as a scalar (all observers measure the same mass): “mass” is always the “rest mass” of the particle. These section assumes you understand 4-vectors.

Instead of changing the mass with motion, we use the relativistic laws of motion, which include the γ factor. They are written in 3-vector and 4-vector form as:

$$(3-vector): \qquad \mathbf{p} \equiv \gamma m \mathbf{v}, \qquad\qquad \mathbf{F} = \frac{d\mathbf{p}}{dt}.$$

$$(4-vector): \qquad p^\mu = (E/c, \mathbf{p}) = (\gamma mc, \mathbf{p}), \qquad F^\mu = \frac{dp^\mu}{d\tau}, \qquad where \quad \tau \equiv \text{proper time} .$$

From this, we can write an invariant equation for a particle’s mass from its energy-momentum 4-vector:

$$p^\mu p_\mu = \gamma^2 m^2 c^2 - \gamma^2 m^2 v^2 = m^2 c^2 \gamma^2 \left(1 - \frac{v^2}{c^2}\right) = m^2 c^2 \Rightarrow \qquad m = \sqrt{\frac{p^\mu p_\mu}{c^2}} .$$

For a photon,

$$E = |\mathbf{p}|c \quad \Rightarrow \quad p^\mu = (|\mathbf{p}|, \mathbf{p}) \qquad \Rightarrow \qquad m = \sqrt{\frac{p^\mu p_\mu}{c^2}} = \sqrt{\frac{\mathbf{p}^2 - \mathbf{p}^2}{c^2}} = 0 .$$

Particles that move at the speed of light *must* be massless (to have finite energy), and therefore must satisfy $E = |\mathbf{p}|c$.

8.1.9 The Terrible Electromagnetic Hamiltonian

Some references use a confusing notation for the interaction of an electron with a magnetic field. Before we describe this, recall the general description of a charged particle, with charge q, in a magnetic field. Here, q is the charge of the particle (not a universal physical constant), and is negative for electrons, and positive for positrons and protons:

$$\mathbf{p}_{kin} = m\mathbf{v}, \qquad \mathbf{p} = \mathbf{p}_{kin} + \frac{q}{c}\mathbf{A} \quad \Rightarrow \quad \hat{H} = \frac{\hat{p}_{kin}{}^2}{2m} + V(\mathbf{r}) = \frac{\left(\hat{\mathbf{p}} - \frac{q}{c}\mathbf{A}(\mathbf{r})\right)^2}{2m} + V(\mathbf{r}) \qquad (8.1)$$

$$where \quad \mathbf{p} \equiv \text{canonical momentum} .$$

Note that $\mathbf{p}$ is $\mathbf{p}_{kin}$ *plus* $(q/c)\mathbf{A}$, and therefore the kinetic energy in the hamiltonian *subtracts* $(q/c)\mathbf{A}$ to get $\mathbf{p}_{kin}$.

Some references specialize to the particle being an *electron*, and call its charge “$-e$”, where e is a physical constant (the charge of a proton). This *reverses* the plus and minus signs in the formulas for the canonical momentum and the hamiltonian. This notation is both less general (since it only works for negative particles), and it conflicts with standard notation in classical mechanics (which is not obsessed with electrons).

Don’t do it. Just let charge be charge: q takes the sign of the particle in question, and (8.1) holds.

8.2 Current Events: Probability Current and Electric Current

Many QM references discuss a concept called “probability current,” more accurately called probability current density, measured in particles/s/m^2. This is a mathematical construction, that has limited physical meaning and application. However, comparing it to true particle current illuminates the fundamental postulates and workings of quantum mechanics. Nearly all references give an incomplete equation for probability current. Note that a single-particle *probability* density is actually a *particle* density, for all practical purposes (see "The Meaning of the Wave-Function," section 1.9). However, probability *current* is *different* from the classical currents which generate things like magnetic fields. We proceed as follows:

- the meaning of a general continuity equation
- the quantum continuity equation
- an example of the failure of the "standard" equation
- correcting the failure
- the physical meaning of probability current and other currents, such as electric current and its coupling to magnetics.

This section assumes you are familiar with basic quantum mechanics, such as operators and the Schrödinger equation, with the vector calculus of divergence, and 1D tunneling.

Some popular texts incorrectly say, or suggest, that the EM field
couples to the probability current. This is wrong, as described below.

8.2.1 General Continuity

(excerpted)

8.2.2 Failure of the "Standard" Formula

We now demonstrate the need for **C** with an elementary example of 1D quantum tunneling. When an incident particle impinges on a potential barrier higher than its energy, some of the particle is reflected, and some is transmitted (tunneled), as in Figure 8.1.

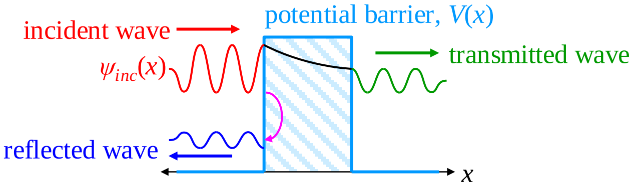


Figure 8.1 1D tunneling: The particle approaches from the left, and interacts with the barrier. The wave-function splits into transmitted and reflected parts. Inside the barrier, $\psi(x)$ varies exponentially.

(excerpted)

8.3 Simple Harmonic Oscillator

8.3.1 Harmonic Oscillator Position PDF: Classical and Quantum

The PDF for a *classical* harmonic oscillator can be found by noting that the probability of being in an interval dx equals the time spent there divided by the period, where the time in dx is the distance divided by the speed:

$$\Pr(\text{being in } dx) = \frac{\text{time in } dx}{T} = \frac{1}{T}\frac{dx}{|v(x)|} \qquad \text{(in each direction)}. \tag{8.2}$$

Therefore, taking a unit amplitude oscillation (without loss of generality):

$$\text{Let} \qquad x(t) = \cos(2\pi t/T) \qquad \Rightarrow$$

$$v(t) = -\frac{2\pi}{T}\sin(2\pi t/T), \qquad \text{and} \qquad v(x) = v(t(x)), \qquad t(x) = \frac{T}{2\pi}\cos^{-1}(x).$$

Each interval dx occurs twice in a cycle (once when the particle moves left, and again when moving right), so Pr(being in dx) for one cycle is double (8.2). Therefore:

$$\text{pdf}_x(x) \equiv \frac{2\,\Pr(\text{being in } dx)}{dx} = \frac{2\,dx}{T|v|\,dx} = \frac{1}{\pi|\sin(2\pi t/T)|} = \frac{1}{\pi}\frac{1}{\left|\sin\left(\cos^{-1}x\right)\right|} = \frac{1}{\pi}\frac{1}{\sqrt{1-x^2}}\,. \qquad (8.3)$$

(excerpted)

8.3.2 Raising and Lowering: How Did Dirac Do It?

The "operator method" of harmonic oscillator analysis is fundamental to all advanced quantum mechanics. It foreshadows a very similar method used for angular momentum, and is the basis of a major part of quantum field theory (including quantum EM radiation), which starts with the fields as an infinite set of quantized harmonic oscillators. The method is far from obvious, so the big question on everybody's mind is "How did Dirac do it?" How did he know how to create the creation and annihilation operators? Most references give them as "Lo! And behold!" (much like they present the Schrödinger equation out of thin air). Here's one way to do it logically, and (who knows?) perhaps is similar to how Dirac figured it out. Our goal here is not to provide the simplest derivation, but to show how such an idea might come about in the first place. In addition, this section further ties together the meaning of operators and Dirac notation, which makes this topic worth understanding. This section assumes you understand how the operator method works, but not how to motivate its development.

(excerpted)

8.4 Make No Mistake With a Quantum Eraser

The quantum eraser is an experimentally confirmed consequence of QM. The name derives from the fact that you can "erase" an interference pattern with a seemingly innocuous change that does not *directly* affect the particles creating the interference. This result is not obvious, and actually highly counter-intuitive. As such, understanding it provides significant insight into quantum physics, which can then be used to help understand more common phenomena. In particular, this effect provides a clear, precise description of what is sometimes called "wave-particle duality" (which is greatly misunderstood, and sometimes discussed in almost mystical terms). However, we show here that the prevention of interference obeys the already-given mathematical rules of QM, and there is no need to invoke duplicities of Nature or other exotic ideas.

Much is often made of so-called "wave-particle duality." It is claimed that a particle behaves like a wave or a particle, but not both. In the double-slit experiment, if we observe the particle going through a slit, it prevents any interference. The explanation is not Nature's defiance, but standard, multi-particle quantum mechanics. It is the entanglement of the measuring device, and not our knowledge of information, which prevents the interference. This same concept applies to the quantum eraser.

In the literature, there are at least 3 different phenomena called a "quantum eraser." However, we believe that one of them is simply a dressed up version of the EPR effect, and not any kind of "eraser." We describe here a recent experiment [Man], which we feel well illustrates the principle. Again, there is no new physics in this experiment, just a surprising consequence of existing multi-particle QM. The result could have been predicted in the 1930s, though parametric down-converters did not then exist.

We proceed as follows:

- Reminder of behavior of 2-particle states.
- The classical and single-photon Mach-Zehnder interferometer.
- The "parametric down-converter."
- The Hong-Ou-Mandel interferometer.
- Preventing ("erasing") the interference pattern.
- Some variations of the experiment, and how they would behave.

- Attempt to debunk some “mystical” interpretations of these QM results.
- What’s not a quantum eraser.

This section assumes you are familiar with superpositions, interference, simple multi-particle states, and entanglement.

Reminder About 2-Particle States

Imagine a system with two distinguishable particles, A and B, each of which can be in one of two states, say $|x>$ and $|y>$. Because of entanglement, a general state of the system can be described using 4 orthonormal basis vectors:

$$\psi = a\left|Ax, Bx\right\rangle + b\left|Ax, By\right\rangle + c\left|Ay, Bx\right\rangle + d\left|Ay, By\right\rangle \qquad \textit{where} \quad a, b, c, d \text{ are complex .}$$

The probability of measuring particle A to be in state $|Ax>$ is $|a|^2 + |b|^2$. Recall that there is no possibility of interference between the a and b coefficients, because the basis vectors $|Ax, Bx>$ and $|Ax, By>$ are orthogonal, even though both vectors describe the state of A as $|Ax>$, i.e. $\left\langle Ax, Bx \middle| Ax, By \right\rangle = 0$.

Interference for one particle can only occur between two quantum state components with a non-zero overlap of the other particle, i.e., a non-zero inner product.

The classical and Single-Photon Mach-Zehnder Interferometer

A Mach-Zehnder interferometer, like all interferometers, starts by splitting a beam of light into two beams (Figure 8.2). (We omit the label t_2 for later consistency with the time-stamps of the Hong-Ou-Mandel interferometer.)

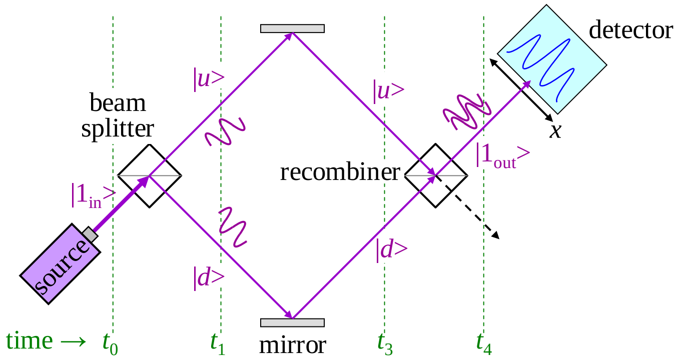


Figure 8.2 The Mach-Zehnder interferometer: Even a single photon at a time through the system produces interference.

(excerpted)

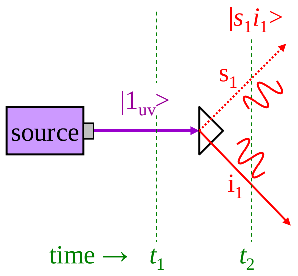


Figure 8.3 A parametric down-converter (PDC) takes a single UV photon (~351 nm) and turns it into two entangled red photons (~702 nm), called “signal” and “idler.”

One of the output photons is arbitrarily called the “signal” photon, and the other is the “idler” photon. (These names have a historical origin which is not relevant to our application.)

The quantum mechanical state of the system evolves in time from a single UV photon at time t_1, to two red photons at time t_2. We denote these spatial states respectively by the kets $|1_{uv}>$, and $|s_1 i_1>$. Note that $|1_{uv}>$ is a single-particle state, and $|s_1 i_1>$ is an entangled two-particle state. It means that if the output photon s_1 is present, then the photon i_1 must also be present. We cannot have one without the other (yet).

Before the PDC, the photon state is time evolving through a series of spatial states, and moving to the right. Similarly, after the PDC, both photons are time evolving through a series of spatial states, and moving in different directions from each other. We have chosen the states $|1_{uv}>$ and $|s_1 i_1>$ as two representative spatial states at the representative times t_1 and t_2.

The conversion efficiency of a PDC is horrible, on the order of one converted photon out of a million, or worse. The vast majority of the incident UV photons go straight through the PDC. Those cases are of no interest to us here, so we ignore them, and consider only the occurrences where down-conversion takes place.

The Hong-Ou-Mandel Interferometer

The Hong-Ou-Mandel interferometer combines a Mach-Zehnder interferometer with two parametric down-converters (Figure 8.4, left). We first summarize the process in words, and then fill in the quantum details.

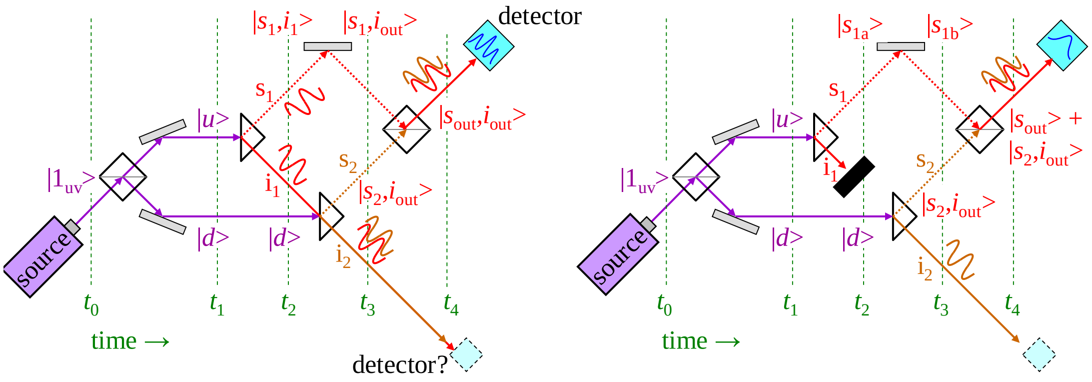


Figure 8.4 The Hong-Ou-Mandel interferometer: (Left) A single photon at a time through the system produces interference. (Right) Blocking the idler eliminates interference between signal photons.

(excerpted)

More Variations of the Quantum Eraser

Other variations of the experiment are also instructive. Return to the setup (Figure 8.4, right) with an absorbing screen. What if the absorber is not complete, but a half intensity filter? In that case, it has a 50% chance of absorbing i_1, and a 50% chance of passing it. That means that half the photons exhibit interference, and half do not. The detector will show some interference, but weaker than the full interferometer (Figure 8.5, middle).

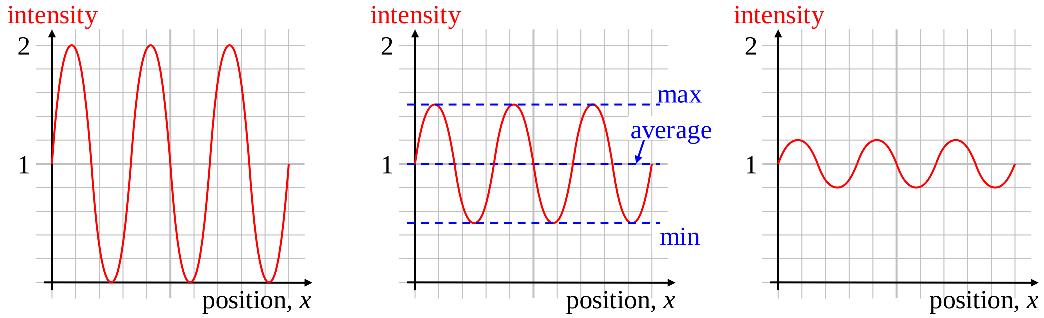


Figure 8.5 The degree of interference is quantified by its "visibility." (Left) Perfect visibility is 100%. (Middle) Visibility = 50%. (Right) Visibility = 20%.

Physicists quantify the degree of interference with its **visibility** (or "fringe visibility"):

$$\text{visibility} \equiv \frac{max - average}{average} = \frac{max - min}{max + min} .$$

(excerpted)

9 Appendices

9.1 Common Quantum Formulas

References are to Goswami's *Quantum Mechanics* [Gos].

General Commutation Relations

$\left[\hat{x},\hat{p}\right]=i\hbar$ [3.5 p56] $\left[\hat{a},\hat{a}^{\dagger}\right]=1$ [7.26 p149] $\left[\hat{\phi},\hat{L}_z\right]=i\hbar$ [9.28 p193]

$\left[\hat{A}\hat{B},\hat{C}\right]=\hat{A}\left[\hat{B},\hat{C}\right]+\left[\hat{A},\hat{C}\right]\hat{B}$ [Q3.5 p72, with other commutation identities]

$\left[\hat{H},\hat{x}\right]=\left[\frac{\hat{p}^2}{2m},\hat{x}\right]=\frac{1}{2m}\left(\hat{p}\left[\hat{p},\hat{x}\right]+\left[\hat{p},\hat{x}\right]\hat{p}\right)=-\frac{i\hbar}{m}\hat{p}$ $\left[\hat{H},\ \hat{L}_i\right]\ =\ 0$ [12.8-9 p247]

$\left[\hat{L}_x,\hat{L}_y\right]=i\hbar\hat{L}_z$ $\left[\hat{L}_y,\hat{L}_z\right]=i\hbar\hat{L}_x$ $\left[\hat{L}_z,\hat{L}_x\right]=i\hbar\hat{L}_y$ [11.6 p221] $\left[\hat{L}^2,\hat{L}_i\right]=0$ [11.10 p222]

$\left[\hat{J}_z,\hat{J}_+\right]=\hbar\hat{J}_+$ $\left[\hat{J}_z,\hat{J}_-\right]=-\hbar\hat{J}_-$ $\left[\hat{J}_+,\hat{J}_-\right]=2\hbar\hat{J}_z$ [11.51 p235]

<u>*x*-basis:</u> free particle: $\psi(t,x)=\exp\left[i(kx-\omega t)\right]$ $\hat{p}=-i\hbar\frac{\partial}{\partial x}$ [3.1 p53]

$\hat{H}\psi=i\hbar\frac{\partial\psi}{\partial t}=-\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2}+V(x)\psi$ [3.10 p58] $\left\langle x|\psi\right\rangle=\psi(\mathrm{x})$ [p 121]

<u>momentum-representation:</u> $\hat{x}=i\hbar\frac{\partial}{\partial p}$ [p 56] <u>Energy representation:</u> $H|E> = E|E>$

ψ concave toward x-axis when $V < E$, away when $V > E$. (correction to Q3.A4 p73)

Time independent Schrödinger equation: $\left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}+V(x)\right]u(x)=Eu(x)$ [3.11 p60]

in standard mathematical form: $\frac{d^2u(x)}{dx^2}+\left(\frac{2m}{\hbar^2}\right)(E-V)u(x)=0$ [4.2 p76]

General wave #: $|k|=\frac{\sqrt{2m(E-V)}}{\hbar}$ Bound: $\xi = k'a$, $\eta = \beta a$, $\xi^2+\eta^2=\frac{2ma^2V_0}{\hbar^2}$

<u>even:</u> $k' \tan k'a = \beta$ $\eta = \xi \tan \xi$ <u>odd:</u> $k' \cot k'a = -\beta$ $\eta = -\xi \cot \xi$

Dirac Notation (Chapter 6)

$<x|\psi> = \psi(x)$ [p 121] $\hat{x}\left|x\right\rangle=x\left|x\right\rangle$ [p 125] $\left\langle x\right|\hat{x}\left|x'\right\rangle=x'\delta(x-x')$ [p 125]

For basis, $\left|\phi_i\right\rangle$: $\left|\psi\right\rangle=\sum_{\substack{i=complete\\ set}}\left|\phi_i\right\rangle\left\langle\phi_i|\psi\right\rangle \quad\Rightarrow\quad \sum_i\left|\phi_i\right\rangle\left\langle\phi_i\right|=\mathbf{1}_{op}$ completeness operator $\left[6.3\text{-}4\ \text{p119}\right]$

$\left\langle j\right|\hat{A}\left|\psi\right\rangle=\sum_i\left\langle j\right|\hat{A}\left|i\right\rangle\left\langle i|\psi\right\rangle$ [6.16, & top p124] $\left\langle\phi\right|\hat{A}^{\dagger}\left|\psi\right\rangle=\left\langle\psi\right|\hat{A}\left|\phi\right\rangle^{*}$ [6.18 p124]

Harmonic Oscillator (Chapter 7)

$H = \frac{p^2}{2m} + \frac{1}{2}kx^2 = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2$ [7.1 p137], $x_0 = \sqrt{\frac{\hbar}{m\omega}}$ (units of x)

$\xi = \sqrt{\frac{m\omega}{\hbar}}x = \frac{x}{x_0}, \quad \omega = \sqrt{\frac{k}{m}}$ [p 138], $E = \frac{1}{2}\hbar\omega\varepsilon = \hbar\omega\left(n + \frac{1}{2}\right)$ [7.13 p142]

Approximate spring constant for general potential minimum: $k = \left.\frac{d^2V}{dx^2}\right|_{x-\min}$ [7.2 p137]

lowering: $\hat{a} = \sqrt{\frac{m\omega}{2\hbar}}\hat{x} + \frac{i\hat{p}}{\sqrt{2m\hbar\omega}}$, raising: $\hat{a}^\dagger = \sqrt{\frac{m\omega}{2\hbar}}\hat{x} - \frac{i\hat{p}}{\sqrt{2m\hbar\omega}}$ [7.25 p149]

$\hat{x} = \sqrt{\frac{\hbar}{2m\omega}}\left(\hat{a}^\dagger + \hat{a}\right)$ $\hat{p} = i\sqrt{\frac{m\hbar\omega}{2}}\left(\hat{a}^\dagger - \hat{a}\right)$ [7.43 p156]

$\hat{N} = \hat{a}^\dagger\hat{a}$ [p152t], $\hat{H} = \hbar\omega\left(\hat{N} + \frac{1}{2}\right) = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right)$ [7.27 p149]

$\left[\hat{H}, \hat{a}\right] = -\hbar\omega\hat{a}$ [7.28 p150] $\left[\hat{H}, \hat{a}^\dagger\right] = +\hbar\omega\hat{a}^\dagger$ [7.29 p150]

$\hat{a}^\dagger\left|u_n\right\rangle = \sqrt{n+1}\left|u_{n+1}\right\rangle$ [7.40 p155] $\left|u_{n+1}\right\rangle = \frac{1}{\sqrt{n+1}}\hat{a}^\dagger\left|u_n\right\rangle$ [p154b]

$\hat{a}\left|u_n\right\rangle = \sqrt{n}\left|u_{n-1}\right\rangle$ [7.42 p155] $\left|u_{n-1}\right\rangle = \frac{1}{\sqrt{n}}\hat{a}\left|u_n\right\rangle$, $\hat{a}\left|u_0\right\rangle = \mathbf{0_v}$ [7.32 p151]

Equations of Motion (Chapter 8)

$i\hbar\frac{\partial}{\partial t}U\left(t, t_0\right) = \hat{H}U\left(t, t_0\right)$ [6.27 p132] $U(t, t_0) = \exp\left[\frac{-i\hat{H}(t - t_0)}{\hbar}\right]$ [8.1, p163]

$\left|\psi\right\rangle(t) = \exp\left[\frac{-i\hat{H}t}{\hbar}\right]\left|\psi\right\rangle(0)$ [8.2 p163] $\hat{A}(t) = e^{i\hat{H}t/\hbar}\hat{A}e^{-i\hat{H}t/\hbar}$ [8.4, p163]

Heisenberg eq. of motion: $\frac{d}{dt}\hat{A}(t) = \frac{i}{\hbar}\left[\hat{H}, \hat{A}(t)\right]$ [8.5 p164]

$\frac{d}{dt}\hat{a}(t) = -i\omega\hat{a}(t) \Rightarrow \hat{a}(t) = e^{-i\omega t}\hat{a}(0)$, $\frac{d}{dt}\hat{a}^\dagger(t) = i\omega\hat{a}^\dagger(t) \Rightarrow \hat{a}^\dagger(t) = e^{i\omega t}\hat{a}^\dagger(0)$ [p165]

$\frac{d\left\langle\hat{A}\right\rangle}{dt} = \frac{i}{\hbar}\left\langle\left[\hat{H}, \hat{A}\right]\right\rangle$ [2nd 8.7 p166] $\frac{d\left\langle x\right\rangle}{dt} = \left\langle\frac{p}{m}\right\rangle$ [8.8 p167] $\frac{d\left\langle p\right\rangle}{dt} = -\left\langle\frac{dV(x)}{dx}\right\rangle$ [8.10 p168]

Particle in Two Dimensions (Chapter 9)

$\hat{p}_x = -i\hbar\frac{\partial}{\partial x}$ $\hat{p}_y = -i\hbar\frac{\partial}{\partial y}$ [9.3 p184] $\hat{\mathbf{p}} = \hat{p}_x\mathbf{e}_x + \hat{p}_y\mathbf{e}_y = -i\hbar\nabla$ [9.4, p186]

Time independent Schrödinger eq: $\nabla^2\psi + \frac{2m}{\hbar^2}\left(E - V\right)\psi = 0$ [9.8 p187]

Motion on a ring: $I = m\rho^2$ [p191] $\psi(\phi) = \frac{1}{\sqrt{2\pi}} e^{im\phi}$ [9.27 p193]

$\hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x = -i\hbar \frac{\partial}{\partial \phi}$ [9.24-5 p192]

Two particle system reduction to 1: $x = x_1 - x_2$ $X = \frac{m_1 x_1 + m_2 x_2}{m_1 + m_2}$ [9.32 p194]

$\frac{1}{\mu} = \frac{1}{m_1} + \frac{1}{m_2}$ $M = m_1 + m_2$ [9.33, 9.34+ p194]

Time dependent Schrödinger eq: $i\hbar \frac{\partial \psi(t,\mathbf{r})}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2 \psi(t,\mathbf{r}) + V(\mathbf{r})\psi(t,\mathbf{r})$ [9.45 p198]

Angular Momentum (Chapter 11)

$\hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x = -i\hbar \frac{\partial}{\partial \phi}$ [11.4 p220, 11.19 p225] $L^2 = r^2 p^2$ [11.2 p219] $H = L^2 / 2I$ [11.3 p219]

$$\hat{L}^2 = -\hbar\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right] \quad [11.20\ p226]$$

$\hat{L}_z \left| l\ m \right\rangle = m\hbar \left| l\ m \right\rangle$, $\hat{L}^2 \left| l\ m \right\rangle = l(l+1)\hbar^2 \left| l\ m \right\rangle$, $l = 0,1,2,...$ [p229]

$\left\langle \theta,\phi \middle| l\ m \right\rangle = Y_{lm}(\theta,\phi)$, $l = 0,1,2,...;$ $m = -1, ... , +l$ [p230]

Legendre polynomials:

$\zeta = \cos\theta$ (11.27 p226) $P_0(\zeta) = 1$, $P_1(\zeta) = \zeta$, $P_2(\zeta) = ½(3\zeta^2 - 1)$ [11.34 p228]

Recursion:

$$(l+1)P_{l+1}(\zeta) = (2l+1)\zeta P_l(\zeta) - lP_{l-1}(\zeta) \qquad \frac{d}{d\zeta}\left(1-\zeta^2\right)\frac{dP_l}{d\zeta} = -l\zeta P_l + lP_{l-1}(\zeta)$$

$$P_{lm}(\zeta) = \left(1-\zeta^2\right)^{m/2}\frac{d^m}{d\zeta^m}P_l(\zeta) = \frac{1}{2^l \cdot l!}\left(1-\zeta^2\right)^{m/2}\frac{d^{l+m}}{d\zeta^{l+m}}\left(\zeta^2 - 1\right)^l \quad [11.35\ p228]$$

$m > 0$: $Y_{lm}(\theta,\phi) = (-1)^m \sqrt{\frac{2l+1}{4\pi}\cdot\frac{(l-m)!}{(l+m)!}}e^{im\phi}P_{lm}(\cos\theta)$ [11.41 p231] $Y_{l,-m} = (-1)^m Y_{lm}^*$ [11.40 p231]

$\hat{J}_+ \equiv \hat{J}_x + i\hat{J}_y$ $\hat{J}_+ \equiv \hat{J}_x + i\hat{J}_y$ [11.50 p234]

$\hat{J}_+\hat{J}_- = \hat{J}^2 - J_z{}^2 + \hbar\hat{J}_z$ $\hat{J}_-\hat{J}_+ = \hat{J}^2 - J_z{}^2 - \hbar\hat{J}_z$ [11.52-3 p235]

$\hat{J}_+ \left| j,m \right\rangle = \hbar\sqrt{j(j+1) - m(m+1)}\left| j,m+1 \right\rangle$ $\hat{J}_- \left| j,m \right\rangle = \hbar\sqrt{j(j+1) - m(m-1)}\left| j,m-1 \right\rangle$ [11.62 p238]

Motion In A Central Potential (Chapter 12)

$H = \frac{P^2}{2M} + \frac{p^2}{2\mu} + V(|r|)$ [12.2 p245] $\hat{H} = \frac{p^2}{2\mu} + V(|r|) = -\frac{\hbar^2}{2\mu}\nabla^2 + V(r)$ [12.3 p245]

$\hat{p}^2 = -\hbar^2\nabla^2$ [12.12] $\underbrace{\hat{p}_r = -i\hbar\frac{\partial}{\partial r}}_{not\ hermitian}$ [12.15 p249] $V_{eff}(r) = V(r) + \hbar^2\frac{l(l+1)}{2\mu r^2}$ [12.25 p252]

Hydrogen Atom (Chapter 13)

$$V(r) = -\frac{Ze^2}{r} \quad [\text{p268}] \qquad E_n = \frac{-\mu Z^2 e^4}{2\hbar^2 n^2} \quad [\text{13.12 p271}] \qquad n = n_r + l + 1, \ \ l < n \quad [\text{p271}]$$

Spin and Matrices (Chapter 15)

|z+> = [1, 0]^T |z–> = [0, 1]^T [15.12 p309]

$$\text{(z-basis)}\ |x+\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix} \qquad |x-\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix} \quad [\text{15.16 p312, \& p316}]$$

$$|y+\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\i\end{bmatrix} \qquad |y-\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-i\end{bmatrix} \quad [\text{p314}]$$

$$\text{Pauli matrices:}\quad \sigma_x = \begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}, \quad \sigma_y = \begin{bmatrix}0 & -i\\i & 0\end{bmatrix}, \quad \sigma_z = \begin{bmatrix}1 & 0\\0 & -1\end{bmatrix} \qquad [\text{15.27 p316}]$$

$$\sigma_\text{i}^{\ 2} = \mathbf{1}_2 \qquad \{\sigma_i, \sigma_j\} \equiv \sigma_i\sigma_j + \sigma_j\sigma_i = 0 \ \ (i \ne j), \qquad \underset{(x,y,z\text{ cyclic})}{[\sigma_x, \sigma_y] = 2i\sigma_z} \qquad [\text{15.29 p317}]$$

Addition of Angular Momentum (Chapter 17)

J = **L** + **S** **L**·**S** = $L_zS_z + ½(L_+S_- + L_-S_+)$ [17.13 p355]

General Clebsh-Gordon coefficient:

$$\langle j_1\ m_1; j_2\ m_2 | J\ M\rangle = \sqrt{\frac{(J + j_1 - j_2)!(J - j_1 + j_2)!(j_1 + j_2 - J)!(J + M)!(J - M)!(2J + 1)}{(J + j_1 + j_2 + 1)!(j_1 - m_1)!(j_1 + m_1)!(j_2 - m_2)!(j_2 + m_2)!}}$$

$$\times\delta(m_1 + m_2, M)\sum_k \frac{(-1)^{k + j_2 + m_2}(J + j_2 + m_1 - k)!(j_1 - m_1 + k)!}{(J - j_1 + j_2 - k)!(J + M - k)!(k!)(k + j_1 + j_2 - M)!} \qquad [\text{C\&S } 14^3 5 \text{ p75}]$$

Perturbation Theory (Chapter 18)

$$E_n^{(1)} = \langle\varphi_n|\hat{H}_1|\varphi_n\rangle \quad [\text{18.10 p380}] \qquad E_n^{(2)} = \sum_{k\ne n}\frac{\left|\langle\varphi_k|\hat{H}_1|\varphi_n\rangle\right|^2}{E_n^{(0)} - E_k^{(0)}} \quad [\text{18.14 p381}]$$

$$\left|\psi_n^{\ (1)}\right\rangle = \sum_{k\ne n}\left[\frac{\langle\varphi_k|\hat{H}_1|\varphi_n\rangle}{E_n^{(0)} - E_k^{(0)}}\right]|\varphi_k\rangle \quad [\text{18.12 p380}] \qquad \Gamma = \frac{2\pi}{\hbar}\left|\langle f|V|i\rangle\right|^2\rho(E_f) \quad [\text{Bay 12-20 p251}]$$

Spherical Harmonics and Their Friends

All spherical harmonics are functions of θ and ϕ. They can be considered functions of all space by ignoring the radial coordinate: $Y_{lm}(r, \theta, \phi) \equiv Y_{lm}(\theta, \phi)$, and can thus be written in rectangular coordinates as $Y_{lm}(x, y, z)$. The transformations require only the conversions in the right column, here:

$$x = r\sin\theta\cos\phi \qquad \cos\theta = z/r$$

$$y = r\sin\theta\sin\phi \qquad e^{i\phi}\sin\theta = \frac{x + iy}{r}$$

$$z = r\cos\theta \qquad e^{-i\phi}\sin\theta = \frac{x - iy}{r}$$

$$Y_{00} = \sqrt{\frac{1}{4\pi}} \qquad\qquad Y_{l0}(\theta) = \sqrt{\frac{2l+1}{4\pi}} P_l \cos\theta$$

Plane wave expansion: $$e^{ikz} = \sum_{l=0}^{\infty} i^l \left(2l+1\right) j_l(kr) P_l(\cos\theta) = \sum_{l=0}^{\infty} i^l \sqrt{4\pi\left(2l+1\right)}\, j_l(kr) Y_{l0}(\theta)$$

The functions below use the Condon-Shortley phase (the $(–1)^m$ factor) [C&S $4^3$17 p52]:

$$Y_{lm}(\theta,\phi) \equiv \begin{cases} (-1)^m \sqrt{\dfrac{(2l+1)}{2}\dfrac{(l-m)!}{(l+m)!}} P_{lm}(\cos\theta)\dfrac{e^{im\phi}}{\sqrt{2\pi}}, & m \ge 0, \\ \sqrt{\dfrac{(2l+1)}{2}\dfrac{(l-|m|)!}{(l+|m|)!}} P_{l|m|}(\cos\theta)\dfrac{e^{im\phi}}{\sqrt{2\pi}}, & m < 0, \end{cases} \qquad \text{[Wyl 3.6.5 p96]}$$

$P_{lm}(x)$ is the associated Legendre function, $l = 0, 1, 2 ...,$ $m = -l, -l+1, ... \, l-1, l$.

$Y_{lm}(\theta,\phi) = \left(-1\right)^m Y_{l,-m}^{*}(\theta,\phi),$ $\forall$ m (positive and negative).

$$Y_{11} = -\sqrt{\frac{3}{8\pi}}\sin\theta e^{i\phi} = \sqrt{\frac{3}{8\pi}}\frac{x+iy}{r} \qquad Y_{10} = \sqrt{\frac{3}{4\pi}}\cos\theta = \sqrt{\frac{3}{4\pi}}\frac{z}{r} \qquad Y_{1,-1} = \sqrt{\frac{3}{8\pi}}\sin\theta e^{-i\phi} = \sqrt{\frac{3}{8\pi}}\frac{x-iy}{r}$$

$$Y_{22} = \sqrt{\frac{15}{32\pi}}\sin^2\theta e^{2i\phi} = \sqrt{\frac{15}{32\pi}}\frac{\left(x+iy\right)^2}{r^2} \qquad Y_{21} = -\sqrt{\frac{15}{8\pi}}\sin\theta\cos\theta e^{i\phi} = -\sqrt{\frac{15}{8\pi}}\frac{\left(x+iy\right)z}{r^2}$$

$$Y_{20} = \sqrt{\frac{5}{16\pi}}\left(3\cos^2\theta - 1\right) = \sqrt{\frac{5}{16\pi}}\frac{-x^2-y^2+2z^2}{r^2} = \sqrt{\frac{5}{16\pi}}\left(3\frac{z^2}{r^2}-1\right)$$

$$Y_{2,-1} = \sqrt{\frac{15}{8\pi}}\sin\theta\cos\theta e^{-i\phi} = \sqrt{\frac{15}{8\pi}}\frac{\left(x-iy\right)z}{r^2} \qquad Y_{2,-2} = \sqrt{\frac{15}{32\pi}}\sin^2\theta e^{-2i\phi} = \sqrt{\frac{15}{32\pi}}\frac{\left(x-iy\right)^2}{r^2}$$

Spherical Hankel Functions

[h_l and h_l^* are sometimes written as $h_l^{(1)}$ and $h_l^{(2)}$.]

$h_l(\rho) = j_l(\rho) + i\eta_l(\rho)$ (outgoing) $\qquad$ $h_l^*(\rho) = j_l(\rho) - i\eta_l(\rho)$ (ingoing)

$$h_l(\rho\to\infty) \to -i\frac{e^{i(\rho - l\pi/2)}}{\rho} \qquad h_l^*(\rho\to\infty) \to +i\frac{e^{-i(\rho - l\pi/2)}}{\rho} \qquad \text{[Wyl 6.2.23-24 p183]}$$

$$j_l(\rho\to 0) = \frac{2^l\, l!}{(2l+1)!}\rho^l = \frac{\rho^l}{(2l+1)!!} \qquad \text{[Wyl 6.2.21 p183, Bay 6-82 p166]}$$

$$\eta_l(\rho\to 0) = -\frac{(2l)!}{2^l\, l!}\frac{1}{\rho^{l+1}} = -\frac{(2l-1)!!}{\rho^{l+1}} \qquad \text{[Wyl 6.2.22 p183, Bay 6-82 p166]}$$

Integrals (See also en.wikipedia.org/wiki/Lists_of_integrals)

$$\int_{-\infty}^{\infty} dx\, e^{-ax^2} = \sqrt{\frac{\pi}{a}} \qquad \int_{-\infty}^{\infty} dx\, x^2 e^{-ax^2} = \frac{1}{2}\sqrt{\frac{\pi}{a^3}}$$

$$\int_0^{\infty} dr\, r^3 e^{-ar^2} = \frac{1}{2a^2}$$

More generally: $$\int_0^{\infty} dr\, r^n e^{-ar^p} = \frac{\Gamma\left(\left(n+1\right)/p\right)}{pa^{(n+1)/p}} \qquad \text{[Tal 370 p235]}$$

General Mathematical Formulas

completing the square: $$ax^2 + bx = \left(\sqrt{a}\,x + \frac{b}{2\sqrt{a}}\right)^2 - \frac{b^2}{4a} \qquad (\text{x-shift} = -b/2a)$$

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \qquad \nabla^2 = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2}{\partial \phi^2} + \frac{\partial^2}{\partial z^2}$$

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta\frac{\partial}{\partial \theta}\right) + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2}{\partial \phi^2}, \qquad \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) = \frac{1}{r}\frac{\partial^2}{\partial r^2}(r\,\bullet)$$

$\cos 2a = \cos^2 a - \sin^2 a = 2\cos^2 a - 1 = 1 - 2\sin^2 a$ $\sin 2a = 2\sin a \cos a$

$$\cos\frac{a}{2} = \sqrt{\frac{1+\cos a}{2}} \qquad \sin\frac{a}{2} = \sqrt{\frac{1-\cos a}{2}}$$

$\cos(a \pm b) = \cos a \cos b \pm \sin a \sin b$ $\sin(a \pm b) = \sin a \cos b \pm \cos a \sin b$

$\cos^2 a = ½ (1 + \cos 2a)$ $\sin^2 a = ½ (1 - \cos 2a)$

$\cos a \cos b = ½ [\cos(a + b) + \cos(a - b)]$ $\sin a \sin b = ½ [\cos(a + b) - \cos(a - b)]$

$\cosh^2 x - \sinh^2 x = 1$

$\cosh 2x = \cosh^2 x + \sinh^2 x = 2\cosh^2 x - 1 = 1 + 2\sinh^2 x$ $\sinh 2x = 2\sinh x \cosh x$

$$e^{ikz} = \sum_{l=0}^{\infty} i^l (2l+1)\, j_l(kr) P_l(\cos\theta) \qquad \text{[Gos 23.34 p500b]}$$

9.2 Glossary

Definitions of common Quantum Mechanics terms:

$<x>$	the average (sometimes called "expectation") value of 'x'
abelian group	a commutative group: $a + b = b + a$.
adjoint	The adjoint of an operator produces a bra from a bra in the same way the original operator produces a ket from a ket: $\hat{O}\lvert\psi\rangle = \lvert\phi\rangle \;\Rightarrow\; \langle\psi\rvert\hat{O}^\dagger = \langle\phi\rvert, \;\forall\lvert\psi\rangle$. Some mathematics references use "adjoint" differently: the "adjoint matrix" ≡ transpose of the cofactor matrix.
aka	also known as.
amplitude	a complex number specifying the magnitude and phase of a quantum value. This is quite different from most applications, where "amplitude" is a *real* number giving the maximum value of a wave. In QM, an "amplitude" can be considered a phasor.
arg(z)	the angle of the complex number, z, in polar form, measured counter-clockwise from the positive real axis
azimuthal quantum number	This is an anachronism for l, the *orbital* angular momentum quantum number. However, [Bay p156m] defines it as m, the z-component of angular momentum, which is consistent with the term "azimuthal."
baryons	3-quark particles, including protons, neutrons, and others. All baryons are hadrons.
basis	(plural: bases) a set of vectors used to construct arbitrary vectors in the vector space. .

bra	a vector in the Hilbert space dual to kets: this means we can take an inner product of a bra with a ket to get a complex number (a scalar).
by definition	in the very nature of the definition itself, without requiring any logical steps. To be distinguished from "by implication."
by implication	combining the definition, and other true statements, a conclusion can be shown by implication.
c-number	"commuting number," or more realistically, "complex number". In general, it would be a "scalar" in the mathematical vector-space of kets.
c'est la vie	French for "that's life."
canonical momentum	In Lagrangian mechanics, canonical momentum is a generalize momentum conjugate to a generalized coordinate, defined as the derivative of the lagrangian with respect to the coordinate: $p_{canonical} \equiv \partial L / \partial q$. (Conjugate here is *not* complex conjugate.)
cf	"compare to." Abbreviation of Latin "confer."
CGS	centimeter-gram-second: a system of measuring distance, mass, and time. There are two different CGS systems, with different units of charge.
closed interval	between *c* and *d* is written "[*c*, *d*]", and means the range of numbers from *c* to *d including c* and *d*.
complex	having a real an imaginary component: the sum of a real and imaginary number.
complex plane	a 2-D graph of complex numbers, with the real part on the abscissa (horizontal axis), and imaginary part on the ordinate (vertical axis).
complicated	involved or intricate.
component	one vector, usually a basis vector, in a superposition of vectors. Each vector in the superposition has a complex coefficient, so the superposition is called "coherent."
comprise	to include. E.g., "An insect comprises 3 parts: a head, thorax, and abdomen." We could say "An insect is composed of 3 parts," but there is no "comprised of".
conjugate bilinear	a function of two mathematical objects which scales as the conjugate of the amplitude of the first, and linearly with the amplitude of the 2nd. For example, see "inner product."
constituent	as used here: one of the states, $\|\psi_k>$, of a mixture (distinct from "component").
continuous	having the property that between any two elements there are an infinite number of other elements, e.g., real numbers are continuous.
contrapositive	The contrapositive of the statement "If A then B" is "If not B then not A." The contrapositive is equivalent to the statement: if the statement is true (or false), the contrapositive is true (or false). If the contrapositive is true (or false), the statement is true (or false).
converse	The converse of the statement "If A then B" is "If B then A". In general, if a statement is true, its converse may be either true or false. The converse is the contrapositive of the inverse, and hence the converse and inverse are equivalent statements.
correlated	in QM, correlated is used "colloquially" (i.e., incorrectly) to mean "dependent." A better word is "entangled." See "entangled." Properly, two sequences of numbers are correlated if there is any component of a linear relationship between the sequences, i.e. if the covariance is nonzero. See *Funky Mathematical Physics Concepts* for more.
density	some quantity per unit volume, e.g. probability density. Outside QM, density can also be per unit mass.

dependent	two quantum systems are dependent (aka entangled) if information about one of them provides information about the other.
dimension	(1) the number of bases in a complete set; (2) units of measure, e.g., the dimensions of voltage in the MKSA system are J/C, or kg-m^2/(C-s^2).
discrete	having the property that between two adjacent elements, there are no other elements, e.g., integers are discrete.
dot product	inner product.
dual vector	In this context, a "dual vector" is one which can be combined with a "vector" to produce a scalar (via the inner product). (Mathematics defines at least two other meanings for "dual.")
ensemble	a hypothetical set of identically prepared systems.
entangled	the quantum states of two systems are entangled when neither system has a definite value of some property, but the values of the two systems' properties are interdependent. This means a measurement of one system provides information about the other.
esu	elector-static unit: a unit of charge which repels a like unit with a force of 1 dyne when spaced by 1 meter. AKA stat-coulomb.
expectation value	not used in this book. Usually, the average value.
fact	a small piece of information backed by solid evidence (in hard science, usually repeatable evidence). If someone disputes a fact, it is still a fact. "If a thousand people say a foolish thing, it is still a foolish thing." See also "speculation," and "theory."
field	1. In mathematics: a set of elements and two operators in which simultaneous linear equations can be solved. Infinite fields have an infinite number of elements. 2. In physics, a (scalar or vector) function of space.
flux	particles per second (distinguish from "flux-density"). Some references use "flux" to mean "flux density."
flux density	flux per unit area, i.e. particles per second per area (distinguish from "flux"). Some references use "flux" to mean "flux density."
forbidden	forbidden *in first order perturbation theory*. See "strictly forbidden."
hadrons	particles that interact via the strong force. E.g., includes pions and baryons, but not leptons. 3-quark hadrons are also baryons. 2-quark hadrons (a quark and anti-quark) are also mesons.
hermitian	her-mish'-un: an operator whose eigenvalues are real; equivalently, a self-adjoint operator.
Hilbert space	Physics: a vector space, of finite or infinite dimension, with a metric (dot product). Mathematicians require that a Hilbert space be infinite dimensional.
idempotent	an operator is idempotent if $\hat{O}^2 = \hat{O}$, such as a projection operator, $\|\psi><\psi\|$. By extension, $\hat{O}^n = \hat{O}$ for all positive integers n.
idler photon	one of the two photons produced by a parametric down-converter. See also signal photon.
iff	if and only if; both necessary and sufficient. Used in definitions.
implies	guarantees. In conversation, "implies" means "suggests." But in math and science, "implies" is stronger.
inflection	a change in curvature (from up to down, or down to up).

inner product	a conjugate bilinear function of two vectors producing a scalar, i.e. the inner product scales as the conjugate of the amplitude of the first vector, and linearly with the amplitude of the second vector.
instantaneous amplitude	the magnitude of a wave at a given point in space at a given point in time.
inverse	The inverse of the statement "If A then B" is "If not A then not B." In general, if a statement is true, its inverse may be either true or false. The inverse is the contrapositive of the converse, and hence the converse and inverse are equivalent statements.
ket	a vector in the Hilbert space of quantum states and of the space of vectors resulting from operators acting on states. In wave mechanics, a ket is a function of space. In matrix mechanics, a ket is a vector with discrete components. The ket space is dual to the bra space, which means we can take an inner product of a bra and a ket to get a complex number (a scalar).
logarithmic derivative	the ratio of a derivative to the function, i.e. $f'(x)/f(x) = d/dx\,(\ln f(x))$. This is the *fractional* rate of change with x, with units of $[x]^{-1}$.
MKSA	meter-kilogram-second-ampere: a subset of the SI system for measuring mechanical and electromagnetic phenomena.
momentum representation	a wave function expressed as a function of momentum, often written $a(p)$.
NIST	National Institute of Standards and Technology: the US government body which establishes US standards and units of measurement. Works closely with ISO.
number current	particle number current density, synonym: probability current.
observation	a measurement.
occupation number	the number of particles in a single-particle quantum state, e.g. photons in a given mode.
old quantum mechanics	aka "Wilson-Sommerfeld quantization": the incorrect notion that the classical action of a periodic or quasi-periodic system is quantized to multiples of $\hbar$. The hydrogen atom disproves this, but for higher energies, W-S quantization corresponds to the WKB approximation.
open interval	between c and d is written (c, d), and means the range of numbers from c to d *exclusive* of c and d.
orthogonal	having a dot product of zero.
PDF	probability distribution function (or probability density function): e.g., pdf(x) = probability per unit interval of x, for differentially small intervals, 'dx'. Mathematically, pdf(x) dx = Pr(value being in the region $[x, x+dx]$).
phasor	a complex number that represents the amplitude and phase of a sinusoid. The sinusoid frequency is *not* part of the phasor, and must be known from other sources.
polar form	a complex number expressed as a magnitude and angle, $z = (r, \theta) = re^{i\theta}$.
position representation	a wave function expressed as a function of positions in space.
Pr(event)	the probability of 'event'.
probability amplitude	a complex number whose magnitude-squared is a probability. Probability amplitudes generally add as complex numbers, then the magnitude-squared of the sum is a probability.
QED	quod erat demonstrandum. Latin for "which was to be demonstrated."
Quantum Field Theory	the physics of relativistic particles and fields (such as EM fields). Massive particles are not conserved, since they may be created or destroyed in interactions.

quantum mechanics	the physics of nonrelativistic particles and systems at microscopic scales. Massive particles are conserved, i.e. they are not created or destroyed.
quantum state	a complex valued function of all space, combined with a spin state, that defines all of the properties of a particle; in particular, the state defines the probability amplitudes, and thus density functions, for all values of all measurements of all observables of the particle.
probability current	particle number current density, synonym: number current.
positive definite	a matrix or operator which is > 0 for all *non-zero* operands. It may be 0 when acting on a “zero” operand, such as the zero vector. This implies that all eigenvalues > 0. E.g., L^2-hat is a positive definite operator when acting on eigenstates $\vert l\ m>$, when $l \geq 1$.
positive semidefinite	a matrix or operator which is ≥ 0 for all *non-zero* operands. It may be 0 when acting on a non-zero operands. This implies that all eigenvalues ≥ 0. E.g., L^2-hat is a positive semidefinite operator when acting on all eigenstates $\vert l\ m>$, for all l.
probability amplitude	a quantum amplitude (a complex number) whose squared-magnitude is the probability of something [Bay p8t].
rectangular form	a complex number expressed as the sum of a real and imaginary component, $z = x + iy$
RHS	right hand side.
signal photon	one of the two photons produced by a parametric down-converter. See also idler photon.
spatial state	the wave-function of a particle or system, which represents all its properties except spin-related ones.
speculation	A guess, possibly hinted at by evidence, but not well supported. Every scientific fact and theory started as a speculation. See also “fact,” and “theory.”
spherical wave	a wave which can be written as $(e^{ikr}/r)f(\theta, \phi)$. A spherical wave is *not*, in general, spherically symmetric, as its amplitude and phase may vary with (θ, ϕ).
spin	the intrinsic angular momentum of a particle. Most particles cannot avoid having spin (electrons, protons, neutrons). It is a property of the particle.
spin-state	the quantum state of a particle’s spin. It can be represented as a spinor.
spinor	a vector in a 2-dimensional Hilbert space representing a particle’s spin. It can be written as a column of 2 complex numbers. In relativistic QM, a spinor has 4 components.
square-integrable	a function whose squared magnitude integrates over all space to a finite value: $\int_{-\infty}^{\infty} dx\,[f(x)]^2 = C$.
stat-coulomb	an esu, or electro-static unit: a unit of charge in the CGS (aka Gaussian) electromagnetic unit system.
static	not moving, compare to “stationary.” A uniformly rotating sphere is stationary, but not static.
stationary	properties constant in time. compare to “static.” A uniformly rotating sphere is stationary, but not static.
strictly forbidden	can never happen at any order, such as one which violates conservation of angular momentum. See “forbidden.”
superposition	a linear combination of vectors, each with a complex coefficient: $a_1 \vert b_1\rangle + a_2 \vert b_2\rangle + \dots$.
symmetry	in general, an invariant property under a given transformation.
theory	the highest level of scientific achievement: a quantitative, predictive, testable model which unifies and relates a body of facts. A theory becomes accepted science only after

	being supported by overwhelming evidence. A theory is not a speculation, e.g. Maxwell's electromagnetic theory. See also "fact," and "speculation."
thought experiment	a way to test the logical consistency of a theory against itself, and other trusted theories, by following the theory to some end. E.g., a thought experiment might compare the predictions of QM against those of Special Relativity. If there is a contradiction, then one of the theories must have an error. (No such contradiction is known.)
trace	the trace of a square matrix is the sum of its diagonal elements.
Twister™	the game that ties you up in knots (a stocking feet game).
unit	the base measure of a measurement, e.g., the unit of distance is the meter.
vectors	mathematically, abstract entities that meet the requirements of vector elements of a vector space. In QM, vectors are usually wave-functions or discrete vectors, aka "kets."
vector space	A mathematical set (often infinite) of a "field" of scalars and a "group" of vectors, with algebraic rules that allow solving linear equations. See text for complete definition.
wave-function	the spatial part of a state: a complex-valued function of space that defines everything there is to know about a particle, except spin, including the probabilities of measuring every value of every observable property of the particle. The complex value at each point is the *probability amplitude*.
wave-number	the spatial frequency of a sinusoid, typically in radians per meter.
wave-vector	a vector describing a propagating sinusoid, whose magnitude is the wave-number and direction is the direction of propagation.
welter Weg	"which way" in German.
Wilson-Sommerfeld quantization	see "old quantum mechanics".
WKB	Wentzel-Kramers-Brillouin: an approximation method for solving the time-independent Schrödinger equation.
WLOG	without loss of generality.
x-representation	sometimes used as a synonym for "position representation"

9.3 References

Books:

[Bay]	Baym, Gordon, *Lectures on Quantum Mechanics*, Westview Press, 1990. ISBN 0-805-30667-6.
[Blo]	D.I. Blokhintsev, *Quantum Mechanics*, Springer; 1 edition (July 31, 1964), English language, ISBN-10: 9027701040, ISBN-13: 978-9027701046.
[C&S]	Condon, E. U., and G. H. Shortley, *The Theory of Atomic Spectra*, Cambridge at the University Press, 1957.
[Dir]	Dirac, P. A. M., *Quantum Mechanics*, 4th ed., Oxford University Press, London, 1958.
[E&R]	Eisberg, Robert and Robert Resnick, *Quantum Physics of Atoms, Molecules, Solids, Nuclei, and Particles*, 2nd Ed., John Wiley and Sons, Inc., 1985. ISBN 0-471-87373-X.
[GAF]	Grynberg, Gilbert, Alain Aspect, and Claude Fabre, *Introduction to Quantum Optics: From the Semi-classical Approach to Quantized Light*, Cambridge University Press (October 18, 2010), ISBN-13: 978-0521551120.
[Gas]	Gasiorowicz, Stephen, *Quantum Physics*, John Wiley & Sons, 1974. ISBN: 047129280X (ISBN13: 9780471292807).

[Gos] Goswami, Amit, *Quantum Mechanics*, 2nd Ed., Wm. C. Brown Publishers, 1997. ISBN 0-697-15797-0.

[Gri] Griffiths, David J., *Introduction to Quantum Mechanics* (2nd Edition), Benjamin Cummings; 2nd edition (April 10, 2004), ISBN-13 978-0131118928.

[Lou] Loudon, Rodney, *The Quantum Theory of Light*, 3rd Ed., Oxford University Press, USA, November 23, 2000, ISBN-13: 978-0198501763.

[P&C] Powell, John I. and Bernd Craseman, *Quantum Mechanics*, Addison-Wesley Publishing Company, Inc, 1961.

[Paa] Paar, Hans P., *An Introduction to Advanced Quantum Physics*, Wiley; 1st edition (May 17, 2010), ISBN-13: 978-0470686751.

[S&W] Benjamin Schumacher, Michael Westmoreland, *Quantum Processes, Systems, and Information*, Cambridge University Press (April 26, 2010), ISBN-13: 978-0521875349.

[Sak] Sakurai, John Jun, *Modern Quantum Mechanics*, revised edition, Addison-Wesley Publishing Co., 1994. ISBN 0-201-53929-2.

[Sak2] Sakurai, John Jun, *Advanced Quantum Mechanics*, Addison-Wesley Publishing Co., 1967, Reading, MA. ISBN 0201067102.

[Sch] Schiff, Leonard Isaac, *Quantum Mechanics*, Mcgraw-Hill College; 3 edition (June 1968), ISBN 13: 9780070856431.

[Sha] Shankar, Ramamurti, *Principles of Quantum Mechanics*, 2nd ed., Plenum Press, New York, 1994. ISBN 0-306-44790-8.

[Tow] Townsend, John S., *A Modern Approach to Quantum Mechanics*, McGraw-Hill, Inc., 1992. ISBN 0-07-065119-1.

[Wyl] Wyld, H. W., *Mathematical Methods for Physics*, 1999, Perseus Books Publishing, LLC, ISBN 0-7382-0125-1.

[Zee] Zee, Anthony, *Quantum Field Theory in a Nutshell*, 2003, Princeton University Press, ISBN 9780691010199.

Papers:

[Jen] Jensen, William B., "*The Origin of the s, p, d, f Orbital Labels*," *J. Chem. Educ.*, 2007, 84, 757-758.

[Man] Mandel, L., *Quantum effects in one-photon and two-photon interference*, Reviews of Modern Physics, Vol. 71, No. 2, Centenary 1999.

[P&M] Pfleeger, R. L. and L. Mandel, *Physical Review* 159, p1084 (1967).

9.4 Index

Here are some suggested index keywords. Some of them might be split into a multi-level term. For example, "complex amplitude" might be listed as-is under "C", or might be listed as "amplitude, complex" under "A". Also, many index keywords may appear in slightly different forms in the actual text. For example, "coherence" may appear as-is, or as "coherent," "coherently," etc.

aaa
aaa
abelian group
absorption
addition of angular momentum
adjoint operator
Airy function
algebraic operator