Fundamentals of Quantum Technology

Thomas E. Roth

Elmore Family School of Electrical and Computer Engineering
Purdue University

Fall 2023

Preface

This set of lecture notes was prepared for a single-semester junior-level course on the fundamentals of quantum technology, offered for the second time in the Fall 2023 semester in the Elmore Family School of Electrical and Computer Engineering at Purdue University. The course is primarily intended to introduce electrical and computer engineering students with no background in quantum mechanics to many of the concepts and applications of quantum information technology, although other engineering students with a basic course in electrical circuits may also find them useful.

It should be emphasized that this course is on *quantum technology*, with an emphasis on *quantum information technology*, rather than the traditional subject of *quantum mechanics.* We of course must cover many quantum mechanical concepts to build the tools needed to understand the quantum technologies of interest, but we omit various standard topics that would be found in most introductory quantum mechanics courses (e.g., angular momentum) in favor of covering more topics that we envision engineers being more likely to work with in the future (e.g., controlling individual quantum systems with classical ones). We also spend more time covering the main pillars of quantum information technologies – quantum communications, quantum computation, and quantum sensing – than is typical in a standard undergraduate quantum mechanics course. Broadly, our overarching goal was to have the students learn many of the fundamental concepts of quantum mechanics so that any "holes" in their knowledge can be readily filled in through self-study or advanced coursework, but still allow enough time to discuss concrete applications that are naturally of significant interest to many engineers.

In this second offering of the course, we have also moved up our discussion on quantum computing to approximately the middle of the semester. We believe this makes for a more engaging course sequence for the students, rather than leaving all the detailed discussions about quantum technologies to the end of the semester. We also then use this technology to motivate many of the remaining technical topics covered in the course – e.g., why we need to have a detailed understanding of quantum harmonic oscillators, artificial atoms, and the interactions between them. We conclude the course with the students independently researching a quantum technology topic of interest to them that they then present to the class to teach one another. The students have strongly appreciated this component of the course, so we have provided a selection of suitable topics in the Practice Problems section of the final chapter of these lecture notes.

In preparing these lecture notes, we drew heavily upon the textbook *Quantum Mechanics for Scientists and Engineers* by D. A. B. Miller for standard quantum mechanics topics. We also treated this textbook as a required reference for the course so that the students would have an accessible resource for filling in standard quantum mechanics topics we did not cover if they needed them in the future. Outside of “standard topics”, we attempted to keep the discussions of many concepts more concrete by illustrating them within a single framework of quantized circuits that we feel will be more readily understood by electrical and computer engineers. These quantized circuits also benefit from having a natural hardware implementation using superconducting circuits, which is to date one of the leading platforms being pursued to develop quantum information technologies. We have also drawn upon various excellent sets of lecture notes from the field of circuit quantum electrodynamics by S. M. Girvin of Yale University and N. K. Langford of University of Technology Sydney in preparing sections of these notes.

Contents

Chapter 1

Overview

1.1 A (Brief) History of Quantum Physics

Quantum physics traces its origins to the turn of the 20th century, when physicists were increasingly finding that their theories of the world – such as Newtonian mechanics, thermodynamics, and electromagnetics – were beginning to fail in spectacular ways. It took revolutionary thinking from some of histories greatest physicists (e.g., Albert Einstein) to propose bold new ideas that greatly deviated from prior thinking to begin the process of overcoming the obstacles the field of physics was facing at this time. These jumps in logic about how the world works is encapsulated in the transition from the prior theories of *classical physics* to what we now refer to as *modern physics*, which broadly include any physical theories that make use of quantum mechanics or the theory of relativity. Generally, we must begin to use modern physics when phenomena we are concerned about have a close tie to the behavior of systems on a microscopic scale (e.g., atomic structure of a material) or when the systems are moving at very high speeds (i.e., close to the speed of light).

Unfortunately, we find in both of these cases that these systems begin to behave in ways that defy our intuition built on how we experience the world in our day-to-day lives. This is a large reason why such a massive paradigm shift was needed in developing theories to describe these systems, and also contributes to the difficulty in building new intuition for how these theories should work in a given situation. As a result, we often must fall back on some tool to continue forging ahead in our understanding of these perplexing systems. Generally, this tool will turn out to be some kind of mathematical theory, which in modern contexts can become quite sophisticated. Fortunately, we will find that the vital mathematical tool for grasping many of the underlying principles of quantum mechanics is *linear algebra*, which is already a key part of many areas of electrical and computer engineering that we must be familiar with. As a result, we will see that learning quantum mechanics can help reinforce our understanding of other areas of electrical and computer engineering, and vice-versa.

Now, let's briefly discuss some of the problems physics was facing that led to the creation of quantum mechanics. One classic example is that of *black-body radiation* and what was known as the *ultraviolet catastrophe*. Black-body radiation centers around the wavelength and intensity of electromagnetic radiation that should be emitted from a black body (an idealized definition of an object that absorbs *all* incident electromagnetic radiation, and

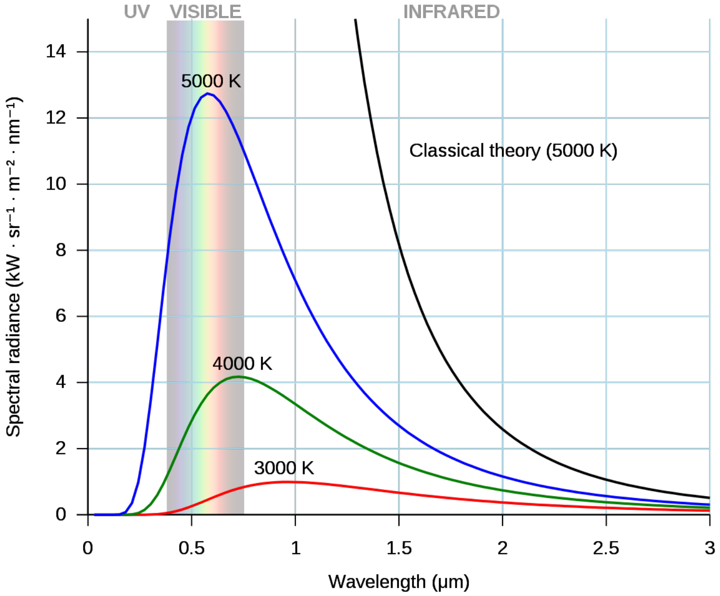


Figure 1.1: Illustration of the spectral intensity distribution for a black body at different temperatures and the (incorrect) Rayleigh-Jeans law (labeled "Classical theory") that was at the center of the ultraviolet catastrophe (image from Wikipedia [1]).

would thus appear perfectly "black" in color) as a function of temperature. You will be familiar with the basic effect of black-body radiation as it relates to the different colors of hot objects, such as the filament in old light bulbs or metals being heated in a furnace. At the time, physics was unable to explain the color that these hot objects should take, and more disastrously predicted that black bodies would emit an ever-increasing/infinite amount of energy as the wavelength moved into the ultraviolet regime (this catastrophe was part of the *Rayleigh-Jeans law* of the time, which is illustrated in Fig. 1.1).

In 1900, Max Planck determined the correct intensity spectral distribution function of radiation for a black body. In order to derive this spectral distribution, Planck had to assume that electromagnetic energy was only emitted or absorbed in discrete packets that are now referred to as *quanta*. (In this context, these quanta are referred to as *photons*.) He determined that these packets of energy would have a value of

$$E = hf \tag{1.1}$$

where $h = 6.626 \times 10^{-34}$ m^{2}kg/s was Planck's constant and f was the frequency of the electromagnetic radiation. At the time, Planck thought this energy quantization was a "formal assumption" that would later be revised in a better way, but it turns out to still hold and often be considered as one of the key steps in the birth of quantum physics. Even today, one can generally determine whether something has its origin in a quantum mechanical effect by seeing whether the formula/equation has Planck's constant in it or not.

Another key observation that contributed to the development of quantum physics was the *photoelectric effect*. This effect is concerned with the emission of electrons from a material when electromagnetic radiation is absorbed by the material. Classical physics would predict that electrons should be able to be emitted due to any frequency of electromagnetic radiation so long as the radiation was applied for a sufficient amount of time to build up the necessary energy. Instead, it was found that electrons would only be emitted once the electromagnetic radiation reached a *particular frequency*. Einstein resolved these issues in 1905 by considering the light to be composed of a discrete set of energy packets (i.e., photons) that were absorbed by the material. *(This was **one** of **four** papers that Einstein published in 1905 that revolutionized physics. He also published an explanation of Brownian motion that helped give credence to the existence of atoms, as well as introduced the theory of special relativity and his famous equation $E = mc^2$. These contributions have made 1905 go down in history as Einstein's aptly named **miracle year** or **wunderjahr**. Einstein was only 26 years old during this miracle year!)*

Although Einstein's formulas were quite simple and explained experimental evidence very well, there was strong resistance to the concepts of discrete packets of electromagnetic energy for various reasons. One central issue was the great success of Maxwell's equations (developed in 1865) in describing electromagnetic physics. Maxwell's equations readily led to the *wave theory of light*, which agreed well with the conclusions of an earlier debate in the 1670's that had Christiaan Huygens wave theory triumph over Isaac Newton's *corpuscular (particle) theory of light*. Due to the successes of the wave theory of light and the many failings of the particle theory of light, physicists were uncomfortable with adopting an approach of photons that "felt" like a return toward a particle view of light. These issues were central to the development of the concept of *wave-particle duality* that is now a key piece of quantum physics. In essence, wave-particle duality is the idea that every quantum "entity" can behave as either a "particle" or a "wave" depending on the circumstances. This notion is another piece of the "weirdness" of quantum physics, and is a good example of how it can become difficult to apply our previous concepts of the "macroscopic world" to the "quantum world".

Jumping forward a bit, another central piece in the development of quantum mechanics was in its role in describing the structure of atoms. Many theories had been developed to explain how electrons "orbited" around the nucleus, but these theories were still incomplete and had certain inconsistencies when attempted to be described using classical physics. For example, classically speaking, if an electron were "orbiting" around a nucleus it would be constantly accelerating, and thus should continually emit radiation. However, it was well known that this constant emission did not occur. In addition to this, there were also further issues about what wavelengths of radiation should be emitted from an atom. As a result, an improved description of atomic structure was needed. After many different approaches from various physicists over the years, Erwin Schrödinger eventually introduced his now famous equation in 1926. In its original form, Schrödinger's equation was a partial differential equation for a "mysterious" wave function. In a simple one-dimensional case, we can write it as

$$i\hbar \frac{\partial}{\partial t}\psi(x,t) = \left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x,t) \right]\psi(x,t), \tag{1.2}$$

where ψ is the wave function, $\hbar = h/2\pi$ is the reduced Planck's constant, m is the mass of the object the wave function is for (e.g., an electron), and V is a potential energy influencing the object of interest. Schrödinger's equation was extremely successful, and its "generalizations" still play a central role in the description of most quantum systems, even in describing the operation of a quantum computer! Schrödinger's equation also played a key role in establishing the essential role that *Hamiltonian mechanics* would play in developing quantized theories of many areas of physics (e.g., that of quantized circuits or quantized electromagnetic fields). Due to this importance, we will spend some time focusing on the basics of Hamiltonian mechanics in this course.

After the introduction of Schrödinger's equation, quantum physics was able to develop at a rapid rate. There have been many significant intellectual achievements along the way, but we will only touch on a few very briefly. One of the next major steps in the development of quantum physics was the introduction of quantum electrodynamics, which describes the quantum interactions between "light and matter" (in the sense of photons and charged particles). Early work in this area was pioneered by Paul Dirac in the 1920s, with many further refinements by a wide range of physicists over the next few decades needed to bring it to its current state that has shown exceptionally good agreement with experimental tests. This represented the first quantum field theory, which has become a vital area of physics that is central to our current understanding of many complex physical systems. Another quantum field theory that has been very successful is that of quantum chromodynamics, which describes the strong force/interaction between quarks mediated by gluons. This was developed in the 1960's and 70's, and plays an important role in the *Standard Model* of particle physics. Although these many theories have been significantly successful, there remain many open questions that are yet to be resolved. As a result, this field is still growing and changing, with many more exciting developments expected still to come!

These earlier theories of quantum physics are often viewed as being primarily concerned with the quantum description of matter, with the quantum description of other components (such as light) coming about as needed. However, interest also began to grow into the quantum physics of light on its "own". This led to the development of the field of quantum optics, which has seen significant growth even to this day due to its usefulness in harnessing quantum effects for new kinds of applications.

Included in these exciting new areas of application are emerging technologies like quantum computers, quantum communication systems, and quantum sensors. Interest in these applications has exploded in the past few decades as foundational experiments demonstrated that revolutionary technologies leveraging new kinds of quantum "weirdness" not previously harnessed could be within reach. The purpose of this course is to introduce electrical and computer engineers to these revolutionary concepts. Our hope is that more engineers can become engaged in this area, and as a result contribute to the development of these exciting new technologies.

1.2 Overview of Quantum Technologies

As we have just discussed, quantum physics began to be developed at the turn of the 20th century. Since then, it has played a vital role in the explosive development of technologies that have occurred throughout the 20th and 21st centuries. However, it is currently believed that we have only begun to "scratch the surface" of what is possible with engineering technologies based on our understanding of quantum physics. Broadly speaking, we typically classify technologies into two kinds of "quantum revolutions", which we will now briefly discuss.

The *first quantum revolution* generally consists of technologies where understanding quantum physics was essential in their creation, but where the quantum effects are essentially leveraged as they "naturally" occur. For example, quantum physics was central to our increased understanding of chemistry and materials science that has contributed to the development of many technologies.

More "at home" to electrical and computer engineers, quantum mechanics played an essential role in the development of technologies that we use to process information. For instance, transistors and integrated circuits that form the bedrock of information technologies relied on our understanding of the wave nature of electrons for their creation. Similarly, optoelectronic technologies that are central to our ability to generate and process optical signals also have their genesis in our understanding of quantum mechanics. This includes understanding the photoelectric effect for devices like optical detectors, and exploiting quantum mechanical properties of semiconductor crystals to generate other technologies like LEDs or photovoltaic solar cells. Additional key technologies that rely on quantum effects are lasers and atomic clocks, each of which have had incredibly broad impacts on science and our daily lives. These represent only a small sampling of technologies developed during the first quantum revolution. However, it should hopefully be clear how important quantum physics was in leading us to our modern age of technology.

Currently, it is believed that we are just beginning what is sometimes referred to as the *second quantum revolution*. In this revolution, it is expected that we will engineer devices that exploit the "rules" of quantum mechanics in ways that we generally do not observe in a natural setting. Central to this is our recent ability to *isolate* and *control* individual quantum systems. Generally, technologies of the first quantum revolution utilized quantum effects in a setting with many constituent particles – e.g., a transistor with many electrons tunneling through a potential barrier or a laser with many emitters being stimulated to emit coherent radiation. In contrast to this, technologies from the second quantum revolution utilize a "small" number of individual quantum systems that we engineer interactions between to achieve a desired effect. By achieving this level of control over quantum systems, we can begin to think of a broader sense of *quantum information*, i.e., the information that can be encoded into a state of a quantum system. We can then think of ways to process this quantum information to develop exciting new technologies. This is the realm of technologies like quantum computers, quantum communication systems, and quantum sensors; which have been proven to yield higher performance than any classical technology could on problems of practical importance. However, the full scope of problems that these technologies may be able to address is still currently unknown. There is a significant amount of research interest in academia, government, and industry focused on better understanding how to develop these new technologies and on determining where they will be of most use. As a result, the field is evolving at a rapid pace, and represents an incredibly exciting area to become engaged in!

Chapter 2

Lagrangian and Hamiltonian Mechanics

2.1 Introduction to the Calculus of Variations

In quantum mechanics, the unfamiliar concept (to electrical and computer engineers) of the *Hamiltonian* of a system plays an incredibly important role. This concept comes about in a particular formulation of classical mechanics, known as *Hamiltonian mechanics*, that has been found to provide a startling simple link between the classical and quantum theories of a system. Today, the Hamiltonian plays a starring role in Schrödinger's equation, and as a result, finding the Hamiltonian of a system is generally one of the first steps taken in analyzing a new quantum system.

Hence, it is necessary for us to become acquainted with the theory of Hamiltonian mechanics so that we can utilize these concepts in tackling quantum descriptions of various systems of interest. Although it is often possible (once one is familiar with the basic steps) to quickly write down the Hamiltonian of a system, there can be important subtleties that must be handled correctly to successfully analyze a system. These subtleties are most readily addressed by first formulating a related quantity, known as the *Lagrangian* of a system. The Lagrangian comes from an alternate theory of mechanics, known as *Lagrangian mechanics*, which has many similarities to the eventual Hamiltonian description of a system. Once the Lagrangian for a system has been determined, basic rules can be used to determine the Hamiltonian, from which a quantum analysis can then be launched.

Now, to further complicate things, both Lagrangian and Hamiltonian mechanics are based on what is known in the field of mathematics as a *variational principle*. Variational principles play a substantially large role in many areas of mathematics, and although it may sound new, it turns out that the basic idea of a variational principle is something that is relatively familiar. In essence, a variational principle is related to finding some function that optimizes a particular "condition" of interest. Generally, we will set up our "condition" so that the optimum will coincide with finding a minima or maxima. You have already done this same kind of operation in your introductory calculus classes when you set the first derivative of a function equal to 0 to find its minima or maxima. Likewise, if you have ever taken a control theory course or dealt with an optimization problem, you have probably seen other examples of this same logic at play. As we will see, a *single* variational principle will allow us to find a unified theory for deriving the equations of motion for almost any physical system (this is the heart of Lagrangian mechanics). The ubiquitousness and importance of this principle is difficult to overstate. In fact, it was so stunning that its early adopters believed it to be the unifying principle of the universe, to the point that one even tried to prove the existence of God based on it [2]!

As alluded to earlier, working with our variational principle is going to require us to be able to differentiate the "condition" we are trying to minimize or maximize. Since our "condition" is typically going to be a "function of functions", we are going to need to utilize a bit more sophisticated form of calculus than what is covered in introductory calculus classes. In particular, we will need to learn a small amount of what is known as the *calculus of variations*. Although this is likely a new concept to you, it was already starting to be developed in the late 1600's to mid 1700's by many of the same mathematicians that developed the calculus you already know so that they could handle a broader class of problems of practical interest. You will find these concepts very useful in this class, but also will likely encounter concepts of a similar "flavor" in other areas of electrical and computer engineering (if you haven't already).

2.1.1 Functionals

To begin understanding the calculus of variations, we will first need to become acquainted with the kind of "mathematical object" that we will be writing our "conditions" in terms of. We have already colloquially referred to this as a *function of functions*, which mathematically is referred to as a *functional*. To understand the basic idea, recall that a function can be abstractly defined as a "map" that takes in a particular number and returns another number as the result (e.g., $f(x) = x^2$). In similarity to this, a functional will take a function as its input and return a single number as its output.

This may seem like an abstract concept at first, but it turns out that it is something that you are already familiar with. As a simple example, we can consider the *Fourier transform* to be a kind of functional. To see this, we will write our Fourier transform functional at a particular angular frequency as $\mathcal{F}_\omega[y(t)]$, with its definition given as

$$\mathcal{F}_\omega[y(t)] = \int_{-\infty}^{\infty} y(t)e^{-i\omega t}dt. \tag{2.1}$$

This clearly takes in the function $y(t)$ and returns a single number, equal to the value of the Fourier transform of $y(t)$ at angular frequency ω. If we then allow the value of ω to be varied, we find that our functional $\mathcal{F}_\omega[y(t)]$ can be viewed as giving us a new function, which we would recognize as the Fourier transform of $y(t)$.

Clearly, in this example, the use of the terminology of the Fourier transform as being a functional is excessive for most practical purposes. Where the notion of a functional becomes very useful is when we want to find some way to maximize or minimize a particular functional as a *function* of its input function. This will be the operation that we need to do in solving many optimization problems, and will be one of the central steps in a Lagrangian mechanics formulation of a problem. Hence, we need to now determine how to "differentiate" a functional.

2.1.2 Functional Derivatives

For inspiration, let's recall the Taylor series of a regular function $f(x)$ at position $x + \epsilon$, where ϵ is some small offset parameter. We know from our introductory calculus that the Taylor series of $f(x + \epsilon)$ is then

$$f(x+\epsilon) = f(x) + \epsilon f'(x) + \frac{1}{2}\epsilon^2 f''(x) + O(\epsilon^3), \tag{2.2}$$

where $O(\epsilon^3)$ is used to denote that all other missing terms in the series are "of order ϵ^3 or smaller". If we only care about terms of $O(\epsilon)$, we can rewrite this as

$$f(x+\epsilon) - f(x) = \epsilon f'(x) + O(\epsilon^2). \tag{2.3}$$

From this, we can see that we can identify the first derivative of f by taking $f(x+\epsilon) - f(x)$ and then only keeping terms of order $O(\epsilon)$.

Considering this, let's do a similar operation with a relatively generic functional. In particular, we will consider our functional to be of the form

$$J[y] = \int_{x_1}^{x_2} f(x, y, y')dx, \tag{2.4}$$

where f is a generic function of the variables x, y, and y'. As we will see throughout this class, it turns out that this relatively simple generic functional will cover most cases of interest to us. To proceed, we need to determine our analog of the small offset ϵ to the functional setting. Since we now need to perturb the input function, we will need to define a "small function" instead of simply having a constant ϵ. So, let's define our perturbation as

$$y(x) \to y(x) + \epsilon\eta(x), \tag{2.5}$$

where ϵ is still a small parameter to ensure that the perturbation $\epsilon\eta(x)$ is "small". Now, to keep things simpler, we will constrain our perturbation function $\eta(x)$ so that it vanishes at the end points of our region; i.e., $\eta(x_1) = \eta(x_2) = 0$. It is possible to handle problems that don't have *fixed endpoints*, but we won't concern ourselves with this case now (if interested, a simple explanation can be found in [2]).

Now that we know how we will perturb the functional, we can go about determining what our *functional derivative* will be (the analog of $f'(x)$ that we had earlier). To do this, we will calculate $J[y + \epsilon\eta(x)] - J[y]$ as

$$J[y+\epsilon\eta(x)] - J[y] = \int_{x_1}^{x_2} \left\{ f(x, y+\epsilon\eta, y'+\epsilon\eta') - f(x,y,y') \right\} dx. \tag{2.6}$$

To proceed, we need to figure out how to simplify this expression into something we are more familiar with working with. To do this, we can utilize a multivariate extension to the Taylor series from (2.3). In general, these expressions can become quite complex, but when we only keep terms of first-order in ϵ the result is quite simple. In particular, we will find that we primarily just add the first-order Taylor series results of each argument together to get

$$f(x, y+\epsilon\eta, y'+\epsilon\eta') = f(x,y,y') + \epsilon\eta\frac{\partial}{\partial y}f(x,y,y') + \epsilon\eta'\frac{\partial}{\partial y'}f(x,y,y') + O(\epsilon^2). \tag{2.7}$$

Note that when we write something like $\frac{\partial f}{\partial y'}$ in this context, we are treating y' like an independent variable from y. As a result, we don't need to worry about taking complex kinds of chain rules when evaluating the derivative $\frac{\partial f}{\partial y'}$. Finally, using the result of (2.7) in (2.6), we find that the result can be rewritten as

$$J[y+\epsilon\eta(x)]-J[y]=\int_{x_1}^{x_2}\left\{\epsilon\eta\frac{\partial f}{\partial y}+\epsilon\eta'\frac{\partial f}{\partial y'}+O(\epsilon^2)\right\}dx. \tag{2.8}$$

Moving on, we will want to rearrange our expression so that we can "cleanly" separate out our perturbation. In this context, what we want is to be able to write the entire inside of the integral in (2.8) as being $\epsilon\eta$ times "something". In similarity to (2.3), we will be able to identify that "something" as being what we will define as the functional derivative. Now, to do this in (2.8), we will first need to integrate by parts to transfer the derivative off of η'. Doing this, and dropping the $O(\epsilon^2)$ terms that we don't care about, we get

$$J[y+\epsilon\eta(x)]-J[y]\approx\int_{x_1}^{x_2}\left\{\epsilon\eta\frac{\partial f}{\partial y}-\epsilon\eta\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right\}dx+\left[\epsilon\eta\frac{\partial f}{\partial y'}\right]_{x_1}^{x_2}. \tag{2.9}$$

Since by our beginning assumptions we are working with an η that vanishes at x_1 and x_2, we can readily see that the final term in (2.9) will vanish. Hence, we get

$$J[y+\epsilon\eta(x)]-J[y]\approx\int_{x_1}^{x_2}\epsilon\eta\left\{\frac{\partial f}{\partial y}-\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right\}dx. \tag{2.10}$$

It is conventional at this point to rewrite our notation a bit. In particular, if we consolidate $\epsilon\eta(x)\to\delta y(x)$, then we would typically write that

$$\delta J\approx\int_{x_1}^{x_2}\delta y(x)\left(\frac{\delta J}{\delta y(x)}\right)dx, \tag{2.11}$$

where

$$\delta J=J[y+\epsilon\eta(x)]-J[y] \tag{2.12}$$

is referred to as the *variation of* J and

$$\frac{\delta J}{\delta y(x)}=\frac{\partial f}{\partial y}-\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) \tag{2.13}$$

is called the *functional derivative* or *Fréchet derivatve* of J with respect to $y(x)$. As we will see later, taking the variation of a particular functional known as the *action functional* will be central to Lagrangian mechanics.

2.1.3 Stationary Points of a Functional

A *stationary point* from introductory calculus occurs when a function is at a maxima, a minima, or a saddle point. You should recall that we can find these points readily by determining when the first derivative of the function is equal 0. In the higher-dimensional case, we simply expand this to finding the points where *all* the first-order partial derivatives are equal to 0. Written in more abstract language, if we had a function $f(x_1, x_2, \ldots, x_n)$ we would find its stationary points by requiring the *variation of* f to be 0 for all first-order perturbations. Mathematically, this would be written as

$$\delta f = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \delta x_i, \tag{2.14}$$

where δx_i are all the possible small perturbations along the different coordinate directions. Since the δx_i are arbitrary, the only way to ensure that (2.14) is always 0 is to require all the partial derivatives to be 0.

It is this more abstract definition that we will use to find the stationary point of a functional. In particular, we will need to ensure that δJ from (2.11) is always 0 for all first-order perturbations; i.e., that

$$\delta J = \int_{x_1}^{x_2} \delta y(x) \left(\frac{\delta J}{\delta y(x)} \right) dx = 0. \tag{2.15}$$

Now, the logic is more complicated than required for (2.14), but the same conclusion can be reached that the only way to ensure that $\delta J = 0$ will be for the functional derivative with respect to $y(x)$ to be 0 at *all* values of x (for more details on the precise reasons why, the reader may consult [2]). More explicitly, we require that

$$\frac{\delta J}{\delta y(x)} = \frac{\partial f}{\partial y} - \frac{d}{dx} \left(\frac{\partial f}{\partial y'} \right) = 0, \quad x_1 < x < x_2. \tag{2.16}$$

This condition is so important that it has its own name; i.e., the *Euler-Lagrange equation.* We will find later that the equations of motion of a system will correspond to the Euler-Lagrange equation.

2.1.4 Example: Shortest Distance Between Two Points

The derivations up to this point have been quite abstract and have relied on notation that is perhaps confusing and unfamiliar. As a result, it is best to stop now and do a simple example to see how some of this mathematical machinery works for an actual problem. We will also consider further examples more focused on familiar mechanical and electrical problems later.

For now, let's look at a "trivial" example of determining the shortest curve that can connect two points (x_1, y_1) and (x_2, y_2) on a Cartesian plane. Intuitively, you already know that the answer will be a straight line, but let's see how we can use the machinery of the calculus of variations to actually prove this result.

To begin, we will need to define our functional, which we will take to be the arc length of the curve between the two points. If we consider x to be our independent variable, then we can define our arc length functional as

$$A[y] = \int_{x_1}^{x_2} \sqrt{1 + (y')^2} dx. \tag{2.17}$$

To minimize the arc length, we need to take the functional derivative of this and set it equal to 0. Since we already know that the first functional derivative will equal the Euler-Lagrange equation given in (2.16), all we need to do is evaluate the necessary (regular) derivatives to fill out the expressions in (2.16).

In this case, we have that our f will just be $f = \sqrt{1 + (y')^2}$. Hence, we find that

$$\frac{\partial f}{\partial y} = 0 \tag{2.18}$$

and

$$\frac{\partial f}{\partial y'} = \frac{y'}{\sqrt{1 + (y')^2}}. \tag{2.19}$$

Now, typically, we would need to evaluate what $\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)$ is. However, because we have that $\frac{\partial f}{\partial y} = 0$ we can instead immediately recognize from the Euler-Lagrange equation that

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = 0. \tag{2.20}$$

Hence, instead of explicitly evaluating the derivative with respect to x, we can instead integrate with respect to x to find that

$$\frac{\partial f}{\partial y'} = C, \tag{2.21}$$

where C is an integration constant. Filling in from (2.19), we find that

$$\frac{y'}{\sqrt{1 + (y')^2}} = C. \tag{2.22}$$

We can now solve for y' to get

$$y' = \sqrt{\frac{C^2}{1 - C^2}} = m, \tag{2.23}$$

where we have simplified the right-hand side since we can recognize that it is simply a constant. We can solve this "differential equation" quite easily by just integrating with respect to x to get that

$$y = mx + b, \tag{2.24}$$

where b is another integration constant. One can utilize the boundary conditions that $f(x_1) = y_1$ and $f(x_2) = y_2$ to find the explicit formulas for m and b. However, it should not surprise you that you will find they simply match what we already know; e.g., that $m = (y_2 - y_1)/(x_2 - x_1)$ will be the slope of the line connecting the points (x_1, y_1) and (x_2, y_2).

Although this was a simple example, we will find that this same basic set of steps can be used to formulate and solve significantly more complex problems. As a result, it represents a very powerful mathematical tool that can be useful in many circumstances. We will see some more examples of this in the coming sections of this course.

2.2 Introduction to Lagrangian Mechanics

As we briefly discussed previously, one of the most common approaches to developing a quantum theory for a physical system begins with a Lagrangian or Hamiltonian mechanics description of the classical system. From there, the Hamiltonian of the quantized system also plays a fundamental role in the quantum dynamics of the system. Hence, it will be necessary for us to become familiar with the frameworks of Lagrangian and Hamiltonian mechanics. In both cases, these frameworks are built from a suitable *variational principle*, so we will also need to utilize our recently learned tools from the calculus of variations. For a more in-depth introduction to these topics, the reader is referred to [3].

2.2.1 Principle of Least Action

The particular variational principle that Lagrangian mechanics is built from is known as the *principle of least action.* This variational principle begins with the *Lagrangian* of the system, which (using standard notation) is given by

$$L = T - V, \tag{2.25}$$

where L is the *Lagrangian*, T is the *kinetic energy* of the system, and V is the *potential energy* of the system. In the Lagrangian mechanics approach, it is necessary to express both the kinetic and potential energies in terms of *generalized coordinates* and *generalized velocities.* If a system has N *degrees of freedom*, then the generalized coordinates q_1, $q_2 \ldots q_N$ are *any* quantities (of total number N) that fully specify the position of a system. For a system of particles, these could simply be the Cartesian coordinates of each particle, or alternatively, another coordinate system description (if more convenient). The generalized velocities are then simply the time derivative of the generalized coordinates, denoted as $\dot{q}_i$.

Once the generalized coordinates and generalized velocities have been simultaneously specified, its motion at future times can be calculated. That is, the accelerations $\ddot{q}_i$ can all be found, and the mechanical state of the system can continue to be advanced in time. We generally refer to the equations that relate the generalized accelerations, velocities, and coordinates of a system as the *equations of motion.* This terminology is still used even for non-mechanical systems (e.g., for an electrical circuit) when analyzed in terms of Lagrangian or Hamiltonian frameworks.

To determine the equations of motion for our system under study, we utilize the *principle of least action.* As the name suggests, this variational principle involves finding the functions that minimize the *action functional*, which is defined as

$$S[q_1, q_2, \ldots, q_N] = \int_{t_1}^{t_2} L(t, q_1, q_2, \ldots, q_N, \dot{q}_1, \dot{q}_2, \ldots, \dot{q}_N) dt. \tag{2.26}$$

We know that to minimize this functional, we need to take the functional derivative and set it equal to 0. We have already previously studied how to take the functional derivative of a similar functional. The main complication here is that we now have multiple functions q_i that make up our total Lagrangian. However, this is not a significant issue, and we can proceed in much the same way by just taking a functional derivative with respect to each q_i

individually and then force each of the resulting Euler-Lagrange equations to be 0. Hence, we find that our equations of motion for each q_i will be

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0. \tag{2.27}$$

As we will show shortly, these equations of motion are equivalent to Newton's many $\mathbf{F} = m\mathbf{a}$ equations that you are familiar with from Newtonian mechanics.

Generally, Lagrange's formulation of mechanics can be more expedient than the Newtonian approach when dealing with complicated systems. It can also provide a unifying framework for many areas of physics, where this same kind of principle of least action can be used to derive the "equations of motion" for a system. For instance, it is possible to use the principle of least action to derive Maxwell's equations (although, one must be clever about identifying the correct generalized coordinates and velocities). Attempting to do this from a Newtonian approach would be ill-advised.

2.2.2 Lagrangian and Newtonian Mechanics

To see how the Lagrangian and Newtonian approaches can "connect" to one another, let's consider a somewhat generic example. In particular, we will consider a system of particles that can interact with one another but with no other bodies. When we consider all parts of the system in this way, we have what is generally described as a *closed system.* We will primarily be concerned with closed systems in this course. In many practical situations, it is necessary to consider *open systems.* This is a particularly challenging problem to deal with in quantum mechanics, but is of central importance in describing the behavior of realistic systems that we wish to do something "useful" with; e.g., quantum computation. We will touch on this topic in a brief way toward the end of this course.

Now, to begin our description of this system of particles, we will need to determine the Lagrangian. We will assume that a Cartesian-type description of the particles is suitable, so we can readily determine that the kinetic energy will simply be

$$T = \sum_i \frac{1}{2} m_i \, \dot{\mathbf{q}}_i \cdot \dot{\mathbf{q}}_i, \tag{2.28}$$

where m_i is the mass of the ith particle and $\dot{\mathbf{q}}_i \cdot \dot{\mathbf{q}}_i$ is the particle's velocity squared. For now, we will consider the potential energy in a generic form, but will assume that it only depends on the relative positions of the particles. A common example of this kind of potential energy would be due to an electrostatic potential between charged particles, but many other examples also exist (e.g., gravitational potential, masses connected between springs, etc.). Considering this, our total Lagrangian will be

$$L = \sum_{i=1}^{N} \frac{1}{2} m_i \, \dot{\mathbf{q}}_i \cdot \dot{\mathbf{q}}_i - V(\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_N). \tag{2.29}$$

We will now use the Euler-Lagrange equations to find our equations of motion for this system. To begin, we note that

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{q}}_i}\right) = m_i \frac{d}{dt}\dot{\mathbf{q}}_i = m_i \ddot{\mathbf{q}}_i, \tag{2.30}$$

which we can recognize as being the mass of the particle times its acceleration. From this, we can find that our complete Euler-Lagrange equation for the ith particle will be

$$m_i \ddot{\mathbf{q}}_i = -\frac{\partial V}{\partial \mathbf{q}_i}. \tag{2.31}$$

We can identify this as being Newton's law if

$$\mathbf{F} = -\frac{\partial V}{\partial \mathbf{q}_i} \tag{2.32}$$

is the *force* acting on the ith particle.

To see that this is the case, let's look at the example of an electrostatic potential between a set of charged particles. We know from introductory electromagnetism that the electrostatic potential for a set of point charges with charge e_j at position $\mathbf{q}$ is

$$\Phi(\mathbf{q}) = \frac{1}{4\pi\epsilon_0} \sum_j \frac{e_j}{|\mathbf{q} - \mathbf{q}_j|}, \tag{2.33}$$

where ϵ_0 is the permittivity of free space. For our Lagrangian, we need the *total potential energy* of the system. Hence, we need to add up the potential energy contribution for each particle. Recalling that the *electrostatic potential energy* of a charged particle is the total electrostatic potential times the particle's charge, we can write this as

$$V = \frac{1}{4\pi\epsilon_0} \sum_{i=1}^{N} e_i \sum_{j \neq i} \frac{e_j}{|\mathbf{q}_i - \mathbf{q}_j|}. \tag{2.34}$$

Evaluating the necessary derivative with respect to $\mathbf{q}_i$, we get that the force acting on the ith particle according to (2.32) is

$$\mathbf{F}_i = \frac{1}{4\pi\epsilon_0} e_i \sum_{j \neq i} e_j \frac{\mathbf{q}_i - \mathbf{q}_j}{|\mathbf{q}_i - \mathbf{q}_j|^3}. \tag{2.35}$$

Again, from electrostatics, we can recognize that

$$\mathbf{E}(\mathbf{q}_i) = \frac{1}{4\pi\epsilon_0} \sum_{j \neq i} e_j \frac{\mathbf{q}_i - \mathbf{q}_j}{|\mathbf{q}_i - \mathbf{q}_j|^3} \tag{2.36}$$

is the electric field due to all other particles evaluated at the location of the ith particle. Recalling that the electrostatic force acting on a charged particle is $\mathbf{F} = e_i \mathbf{E}$, we can see that (2.32) does in fact yield the correct force in this case. Hence, we see that (2.31) directly corresponds to Newton's many $\mathbf{F} = m\mathbf{a}$ equations for this case.

The basic idea that Lagrangian mechanics will provide equivalent results to Newtonian mechanics will hold in general. However, We will not always necessarily end up with as direct of a correspondence between the equations of motion that we will work with in the Lagrangian and Newtonian approaches to mechanics. In particular, Lagrange's approach will prove to be more expedient in situations where certain constraints exist between different particles in the system. Lagrange's approach will remove redundant degrees of freedom due to constraint forces from the beginning, while the Newtonian approach will require us to formulate redundant equations that we later eliminate from the overall system.

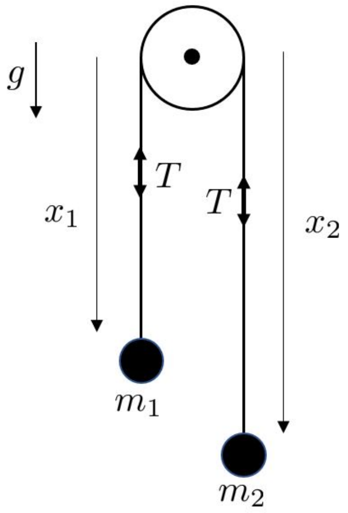


Figure 2.1: Free body diagram of an Atwood machine.

2.2.3 Example: Atwood Machine

Let's now consider a simple example of how to use Lagrangian mechanics for a system with only one degree of freedom. In particular, we will consider the Atwood machine, with it's free body diagram illustrated in Fig. 2.1. We will assume there is no friction and that the string connecting the two masses has a fixed length of ℓ.

The Newtonian mechanics approach to solving this problem would begin by writing an $\mathbf{F} = m\mathbf{a}$ equation for each of the weights in the system. This is simple in this case, and gives us

$$m_1\ddot{x}_1 = m_1 g - T, \tag{2.37}$$

$$m_2\ddot{x}_2 = m_2 g - T, \tag{2.38}$$

where T is the tension of the string and g is the gravitational acceleration. The next step is to recognize that the velocities of the two weights are not independent, but are in fact constrained to be equal and opposite to one another, and hence, so too are the accelerations. As a result, we can eliminate one of our accelerations by noting that $\ddot{x}_2 = -\ddot{x}_1$. Using this, we can rewrite (2.38) to get

$$-m_2\ddot{x}_1 = m_2 g - T. \tag{2.39}$$

We can then eliminate the common tension force between (2.37) and (2.39) to get

$$(m_1 + m_2)\ddot{x}_1 = (m_1 - m_2)g, \tag{2.40}$$

which finishes formulating the equation of motion for the problem.

Let's now follow the Lagrangian mechanics approach to solving this problem. To begin, we need to determine the total kinetic and potential energies in terms of suitably defined generalized coordinates and velocities. For this problem, we recognize that there is only one true degree of freedom in this problem due to the fixed length of the string connecting the masses. As a result, we can set our generalized coordinate to be

$$q = x_1 = \ell - x_2. \tag{2.41}$$

In terms of this generalized coordinate, we can recognize that the total kinetic energy will be

$$T = \frac{1}{2}(m_1 + m_2)\dot{q}^2 \tag{2.42}$$

because the total length ℓ of the string won't influence the generalized velocity. Next, we recognize that the total potential energy can be written as

$$V = -m_1 g x_1 - m_2 g x_2 = -m_1 g q - m_2 g(\ell - q). \tag{2.43}$$

Putting these results together, we find the total Lagrangian is

$$L = \frac{1}{2}(m_1 + m_2)\dot{q}^2 - (m_2 - m_1)gq + m_2 g\ell. \tag{2.44}$$

We can now evaluate our single Euler-Lagrange equation using the usual

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial L}{\partial q} = 0 \tag{2.45}$$

to get

$$(m_1 + m_2)\ddot{q} = (m_1 - m_2)g. \tag{2.46}$$

The first aspect of (2.46) to note is that it matches the Newtonian mechanics result of (2.40), as expected. The next item of interest is to recognize that the potential energy term $m_2 g\ell$ did not factor into the final equation of motion at all. The reason for this is because this term is a constant with respect to the generalized coordinates and velocities of the problem. As a result, the derivatives required in evaluating the Euler-Lagrange equation caused this term to vanish from the equation of motion. *This is a general result: a constant in a Lagrangian does not affect the dynamics of the system at all.* As a result, it is a widely common practice to ignore any constant terms that arise in formulating a Lagrangian of a system and to not even bother to write them down in the first place.

2.2.4 Driven Systems

We will now consider how to incorporate a driving force into our Lagrangian formalism. This is a common occurrence in many practical systems, where it is often useful to consider the case where some "external" force drives the dynamics of our system under study. In many cases, it will be convenient and reasonable to assume that our external drive is "perfectly" applied to our system in the sense that the drive is unaltered by the motion of our underlying system. In the context of electrical circuits, this kind of "perfect" drive could be applied by an ideal voltage or current source.

Now, let's assume that we have a simple mechanical oscillator placed in some external driving field. The *free oscillator* (i.e., the oscillator with no driving force) will have a Lagrangian of

$$L = \frac{1}{2}m\dot{q}^2 - \frac{1}{2}kq^2, \tag{2.47}$$

where k can be thought of as a "spring constant" that occurs in a Hooke's law kind of restoring force. In the presence of the external field, we will gain an additional potential energy $V_e(q,t)$ that needs to be added to the overall Lagrangian.

In many practical cases, we will only be concerned with considering *small oscillations* of the mass, meaning that q will only vary a small distance from its equilibrium value. We can then expand the potential energy contribution in a first-order Taylor series as

$$V_e(q,t) \approx V_e(0,t) + q\frac{\partial V_e(0,t)}{\partial q}, \tag{2.48}$$

where we have assumed the equilibrium position is $q = 0$ for simplicity. We can simplify this result by noting a general property of Lagrangian mechanics: Two Lagrangians differing only by a total derivative with respect to time of some function have the same equations of motion. This occurs because the total time derivative can be integrated out of the action functional quite easily, and with fixed endpoints then vanishes when taking the variation of the functional. Now, because $V_e(0,t)$ only depends on time, we can consider it to be the total time derivative of some other function, and hence, will not modify our final equations of motion. As a result, we can simply omit it from our Lagrangian. We can further recall that the *force* will be given by

$$F(t) = -\frac{\partial V_e(0,t)}{\partial q}. \tag{2.49}$$

Hence, we can write the additional potential energy term in our overall Lagrangian for the driven oscillator case as

$$V_e = -qF(t). \tag{2.50}$$

This basic structure for the additional potential energy will also prove useful when we consider driven electrical circuits.

Considering these points, we find that our complete Lagrangian for the driven system will be

$$L = \frac{1}{2}m\dot{q}^2 - \frac{1}{2}kq^2 + qF(t). \tag{2.51}$$

We can find the resulting equations of motion from the Euler-Lagrange equation

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial L}{\partial q} = 0 \tag{2.52}$$

as

$$m\ddot{q} + kq = F(t). \tag{2.53}$$

It is common to simplify this as

$$\ddot{q} + \omega^2 q = F(t)/m, \tag{2.54}$$

where $\omega = \sqrt{k/m}$ is the *angular frequency* of free oscillations in the system. If there were no driving force, the solution to this differential equation would simply be

$$q(t) = C_1 \cos(\omega t) + C_2 \sin(\omega t), \tag{2.55}$$

which is why ω is referred to as the angular frequency of the oscillator. We will consider the effect of an external driving force in more detail later in the course, as it plays a central role in considering the control of quantum systems.

2.2.5 First Integral and Noether's Theorem

Another advantage of using a variational principle in deriving equations of motion is that it can be used to quickly deduce other important properties of the system under study. We will look at a particular instance of this concept that demonstrates the conservation of energy, and then comment on the broader principle known as *Noether's theorem.*

To begin, we will introduce the concept of a *first integral* of an Euler-Lagrange equation. This concept is of importance when our functional does not have an explicit dependence on the variable being integrated over. Hence, in the case of the action functional we would be working with a Lagrangian of the form $L(q, \dot{q})$ rather than the more general $L(t, q, \dot{q})$ that we considered earlier. In this case, we would have that the total time derivative of the Lagrangian would be

$$\begin{aligned} \frac{dL(q,\dot{q})}{dt} &= \frac{\partial L(q,\dot{q})}{\partial q}\frac{\partial q}{\partial t} + \frac{\partial L(q,\dot{q})}{\partial \dot{q}}\frac{\partial \dot{q}}{\partial t} + \frac{\partial L(q,\dot{q})}{\partial t} \\ &= \dot{q}\frac{\partial L(q,\dot{q})}{\partial q} + \ddot{q}\frac{\partial L(q,\dot{q})}{\partial \dot{q}} \end{aligned} \tag{2.56}$$

We can then determine that in this case

$$\frac{d}{dt}\left(L - \dot{q}\frac{\partial L}{\partial \dot{q}}\right) = \dot{q}\frac{\partial L}{\partial q} + \ddot{q}\frac{\partial L}{\partial \dot{q}} - \ddot{q}\frac{\partial L}{\partial \dot{q}} - \dot{q}\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right), \tag{2.57}$$

which simplifies to

$$\frac{d}{dt}\left(L - \dot{q}\frac{\partial L}{\partial \dot{q}}\right) = \dot{q}\left(\frac{\partial L}{\partial q} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right)\right). \tag{2.58}$$

We can recognize that the interior of the parentheses on the right-hand side of (2.58) is the Euler-Lagrange equation of our system, and so must be 0. Hence, we have that

$$\frac{d}{dt}\left(L - \dot{q}\frac{\partial L}{\partial \dot{q}}\right) = 0, \tag{2.59}$$

and we refer to the quantity

$$I = L - \dot{q}\frac{\partial L}{\partial \dot{q}} \tag{2.60}$$

as the *first integral* of the Euler-Lagrange equation.

To see the physical significance of this result, we first recognize that (2.59) is the total time derivative of I, and that this result is 0. We can then conclude that I *must* equal a constant as a function of time, and thus represents a *conserved quantity.* In particular, (by definition) it turns out that $-I$ is the *total energy* of the system being studied, and so its time invariance leads to the conservation of energy.

To see this, let's look at a relatively simple Lagrangian of a harmonic oscillator with

$$L = \frac{1}{2}m\dot{q}^2 - \frac{1}{2}kq^2. \tag{2.61}$$

Now, if we rearrange the first integral (which is acceptable since (2.59) equals 0), we can get

$$\begin{aligned}\dot{q}\frac{\partial L}{\partial \dot{q}} - L &= \dot{q}\bigg(m\dot{q} - 0\bigg) - \bigg(\frac{1}{2}m\dot{q}^2 - \frac{1}{2}kq^2\bigg)\\ &= m\dot{q}^2 - \frac{1}{2}m\dot{q}^2 + \frac{1}{2}kq^2\\ &= \frac{1}{2}m\dot{q}^2 + \frac{1}{2}kq^2\\ &= T + V.\end{aligned} \tag{2.62}$$

We can recognize $T+V$ as being the total energy of the system, which is then conserved due to our earlier conclusions about the first integral. Although we only showed this for a simple example, we will find more broadly that the quantity $\dot{q}\frac{\partial L}{\partial \dot{q}} - L$ corresponds to the total energy of the system. Due to its importance, $\dot{q}\frac{\partial L}{\partial \dot{q}} - L$ goes by another name in mechanics; namely, the *Hamiltonian* of the system. This will be the central element in Hamiltonian mechanics, which we will discuss shortly.

Prior to this, we will briefly comment on a general physical principle known as *Noether's theorem.* This theorem, proved by Emmy Noether, makes an explicit connection between a *symmetry* in an action functional and a resulting *conservation law* of the system. In the case of the conservation of energy, the *symmetry* was that the system was invariant under time translations. Noether's theorem can be used to determine the conservation laws of many systems, and is part of the reason why finding symmetries in systems has become such an important concept in physics. Before moving on, it is worth mentioning that Emmy Noether is a fascinating figure in the history of mathematics. She is consistently considered one of the greatest mathematicians of the 20th century and was labeled as the most important woman in the history of mathematics by many of her contemporaries, such as Albert Einstein, Hermann Weyl, and Norbert Wiener, and is certainly worth reading more about sometime.

2.3 Introduction to Hamiltonian Mechanics

In addition to Lagrangian mechanics, there is another (closely) related framework for analyzing physical systems known as *Hamiltonian mechanics.* As we will see repeatedly throughout this course, the Hamiltonian of a system plays a central role in developing a quantum theory for a physical system. As a result, becoming familiar with the basics of Hamiltonian mechanics is a necessary step in growing acquainted with quantum theory.

2.3.1 Preliminaries

We have already briefly encountered what the Hamiltonian of a system is when we discussed the first integral and energy conservation from the Lagrangian mechanics viewpoint. There, we showed that the first integral of the Euler-Lagrange equations derived from the principle of least action was

$$\dot{q}\frac{\partial L}{\partial \dot{q}} - L. \tag{2.63}$$

We also discussed that this was equal to the total energy of the system, and that the total energy was the *Hamiltonian* of the system.

Although this basic idea does hold true, we will not always refer to (2.63) directly as the Hamiltonian because the Hamiltonian mechanics framework uses differently defined generalized coordinates. In particular, Hamiltonian mechanics is formulated in terms of *canonical position* and *canonical momenta (also, conjugate momenta)* variables, which are closely related to the generalized coordinates and velocities of the Lagrangian approach. In particular, we will have that the canonical position and generalized coordinates will be the same. However, instead of generalized velocities $\dot{q}_i$ we will use canonical momenta, defined as

$$p_i = \frac{\partial L}{\partial \dot{q}}. \tag{2.64}$$

In terms of these variables, we will generally write the Hamiltonian of a system as

$$H(t, q_i, p_i) = \sum_i p_i \dot{q}_i - L(t, q_i, \dot{q}_i), \tag{2.65}$$

which matches the earlier definition of (2.63) when we use both generalized velocities and canonical momenta in our expression. When we go to actually use the Hamiltonian, we must take all of the $\dot{q}_i$'s in (2.65) and write them in terms of the p_i's. From our earlier discussions, we can also recognize that in many cases we will be able to write the Hamiltonian as

$$H = T + V, \tag{2.66}$$

where all energies would be expressed in terms of q_i and p_i. Although it can be tempting to try and write down a Hamiltonian for a system directly using (2.66), it can be risky if the canonical momenta are not correctly identified (this can particularly tricky when dealing with coupled systems, as we will see later). Hence, it is common to follow through the somewhat formal process of determining the Hamiltonian and canonical momenta from a Lagrangian when first working with a new system.

Once we have determined the canonical position, canonical momenta, and Hamiltonian of a system, we can use them to derive the equations of motion. This can be done through taking the variation of the Hamiltonian and Lagrangian and consolidating various terms.

Instead of this more elaborate process, we will follow a simpler approach that gives the same core result. In particular, if we have a Hamiltonian that is not an explicit function of time, then we can determine that the total time derivative of the Hamiltonian will be

$$\frac{dH(q_i(t), p_i(t))}{dt} = \sum_i \left(\frac{\partial H}{\partial p_i} \frac{dp_i}{dt} + \frac{\partial H}{\partial q_i} \frac{dq_i}{dt} \right) = 0. \tag{2.67}$$

In (2.67), we have noted that this time derivative must equal 0 due to energy conservation. One way to ensure that (2.67) does equal 0 is for

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}. \tag{2.68}$$

Typically, we will refer to (2.68) as *Hamilton's equations.* Together, they give us the equations of motion for our system. Intuitively, what (2.67) and (2.68) tell us is that the canonical position and canonical momenta of a system vary in time in such a way to ensure that energy is conserved. It should be noted that even though we followed a somewhat simplistic derivation of (2.68), this result does hold in general when following a Hamiltonian mechanics approach to analyzing a system.

Before looking at an example, a few remarks are in order about the differences between the Hamiltonian and Lagrangian frameworks. First, when we work with a Lagrangian approach, we can often consider the generalized velocities to be a dependent variable of the generalized coordinates in the sense that $\dot{q}_i = dq_i/dt$. We have to be careful with not treating this as a true dependent variable prior to evaluating the Euler-Lagrange equation, but once that has been completed the analysis works out so that when we derive our equations of motion we end up with a *single second-order* differential equation *for each* generalized coordinate in our system. In contrast to this, when we follow a Hamiltonian approach, we *must* consider the canonical position and canonical momenta to be independent variables when we are deriving our equations of motion. This is necessary so that mathematically these can vary independently of one another to ensure that energy is conserved. From a practical perspective, what this means is that we cannot take our Hamiltonian and try and rewrite parts involving the canonical momenta if we "think" we know how the momenta should be related to the position coordinates of our system. Another difference between the two approaches is the equations of motion that we arrive at from our derivation. Considering (2.68), we see that the Hamiltonian approach will give us *two first-order* differential equations *for each* degree of freedom that we have in our system.

2.3.2 Example: Free Oscillator

Let's now consider a concrete example to see how we can transition from a Lagrangian description of a system to a Hamiltonian one. In particular, we will consider a simple harmonic oscillator. This is the same kind of system that we considered when discussing how to include driving forces in a Lagrangian framework, however, we will not consider the driving force in this example. Considering this, we had as our Lagrangian

$$L = \frac{1}{2} m \dot{q}^2 - \frac{1}{2} k q^2. \tag{2.69}$$

We can easily evaluate the Euler-Lagrange equations to find that our equation of motion for this system is

$$\ddot{q} + \omega^2 q = 0, \tag{2.70}$$

where $\omega = \sqrt{k/m}$.

Now, let's go through the process of finding the Hamiltonian for this system and determining the corresponding equations of motion. To begin, we will need to determine what the canonical momentum should be. We do this from its definition to find that

$$p = \frac{\partial L}{\partial \dot{q}} = m\dot{q}. \tag{2.71}$$

In this case, we see that the canonical momentum is simply the *kinetic momentum* of the system (i.e., mass times velocity). *It is essential to understand that this result will not always hold – i.e., the canonical momentum and kinetic momentum of a system will often not be the same in practical situations of interest.*

Our next step is to evaluate what the Hamiltonian will be. By definition, we have that

$$H = p\dot{q} - L. \tag{2.72}$$

To use this definition, we first need to express $\dot{q}$ and L in terms of p. We can readily find from (2.71) that

$$\dot{q} = p/m, \tag{2.73}$$

and we can use this result to write L as

$$L(q, p) = \frac{1}{2m}p^2 - \frac{1}{2}kq^2. \tag{2.74}$$

Using these expressions in (2.72), we find that our Hamiltonian will be

$$\begin{aligned} H &= p\left(\frac{p}{m}\right) - \left(\frac{1}{2m}p^2 - \frac{1}{2}kq^2\right) \\ &= \frac{1}{2m}p^2 + \frac{1}{2}kq^2. \end{aligned} \tag{2.75}$$

We can recognize that this matches what we would expect $H = T + V$ to be from a naive perspective of the canonical momenta matching the kinetic momenta. *It is stressed again that this will not always be the case, so simply writing down these expressions in this way without performing a more formal check can lead to incorrect results!*

Now that we have our Hamiltonian, we can evaluate Hamilton's equations to find our equations of motion. In particular, we find that

$$\frac{dq}{dt} = p/m \tag{2.76}$$

and

$$\frac{dp}{dt} = -kq. \tag{2.77}$$

These first-order ordinary differential equations can be integrated in time together to determine the dynamics of our system. Alternatively, we can take the time derivative of (2.76) to get

$$\frac{d^2q}{dt^2} = \frac{1}{m}\frac{dp}{dt} \tag{2.78}$$

and substitute in our result from (2.77) to find that

$$\ddot{q} + \omega^2 q = 0, \tag{2.79}$$

in agreement with our Lagrangian mechanics result.

2.3.3 Poisson Brackets

In certain situations, it can be useful to derive equations of motion for other functions besides the canonical position and momenta. This can be done with what is known as a *Poisson bracket*. If we have two functions defined in terms of our canonical positions and momenta, then the Poisson bracket is defined as

$$\{f, g\} = \sum_i \left(\frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i}\right). \tag{2.80}$$

We can use this to evaluate the equation of motion for a particular quantity by recognizing that

$$\begin{aligned}\frac{d}{dt}f(q,p,t) &= \sum_i \left(\frac{\partial f}{\partial q_i}\frac{dq_i}{dt} + \frac{\partial f}{\partial p_i}\frac{dp_i}{dt}\right) + \frac{\partial f}{\partial t} \\ &= \sum_i \left(\frac{\partial f}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial H}{\partial q_i}\right) + \frac{\partial f}{\partial t} \\ &= \{f, H\} + \frac{\partial f}{\partial t},\end{aligned} \tag{2.81}$$

where in going from the first to second line we utilized Hamilton's equations from (2.68). In the common situation where f does not depend explicitly on time, we simply get that

$$\frac{d}{dt}f(q,p) = \{f, H\}. \tag{2.82}$$

The Poisson bracket plays a very important role in *quantizing* a classical theory, as we will see in more detail later in the course, and also provides a convenient way to work with some standard problems that appear ubiquitously in analyzing quantum systems.

2.4 Lagrangian and Hamiltonian Mechanics of Simple Circuits

We will now discuss how to apply the Lagrangian and Hamiltonian frameworks to analyzing simple circuits. We will only focus on a simple LC oscillator at this point, but we will

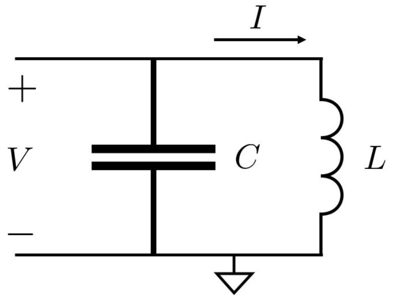


Figure 2.2: Circuit schematic of a parallel LC circuit (also referred to as an *LC tank circuit*).

discuss more complex systems as they naturally arise throughout this course. Examples of more complete discussions about analyzing (quantum) lumped element circuits using a Lagrangian or Hamiltonian mechanics approach can be found in various references, e.g., see [4–6].

To begin, we will consider the parallel LC circuit shown in Fig. 2.2. In principle, we can write down the Hamiltonian of this circuit very easily using simple circuit theory results. In particular, we would say that the total energy is given by the combination of the electrical potential energy stored in the capacitor and the magnetic potential energy stored in the inductor, yielding

$$H = \frac{1}{2}CV^2 + \frac{1}{2}LI^2. \tag{2.83}$$

Although this expression is accurate for this circuit, it turns out that our familiar definitions of voltage and current do not correspond to appropriate variables to work with in either a Lagrangian or a Hamiltonian description of our circuits. As a result, we will need to find alternative descriptions that are more like “position” and “velocity” variables of mechanical systems.

Various choices exist as to which variables in our circuit can be denoted as our “generalized position”, but we will follow the particular choice that is most prevalent due to its ease of use with a popular experimental platform of superconducting quantum circuits [5, 6]. In this approach, we will have our generalized position be the *node flux*, defined as

$$\phi(t) = \int_{-\infty}^{t} V(\tau)d\tau, \tag{2.84}$$

where V is our usual voltage at the node of interest. Considering this, we can find that the energy stored in our capacitor can be given by

$$U_C = \frac{1}{2}C\dot{\phi}^2, \tag{2.85}$$

which now looks similar to the kinetic energy of a particle with “mass” C.

To see what happens to the energy stored in the inductor, we note that the I-V relationship for an inductor can be given as

$$V = \dot{\phi} = L\dot{I}. \tag{2.86}$$

Hence, we can find that in this case

$$I = \frac{\phi}{L}, \tag{2.87}$$

which shows from the definition of an inductance that the node flux *in this case* is also the magnetic flux of the inductor. We can then write the energy stored in the inductor as

$$U_L = \frac{1}{2L}\phi^2, \tag{2.88}$$

which looks like the potential energy of a mass with a spring constant of $1/L$. *It is important to once again stress that the identification of generalized coordinates, velocities, and canonical momenta must be considered on a case by case basis, and so the simple relationships we are seeing here will not always hold in general!*

Considering these points, we can write the Lagrangian for our system as

$$\mathcal{L} = T - V = \frac{1}{2}C\dot{\phi}^2 - \frac{1}{2L}\phi^2, \tag{2.89}$$

where we have identified the capacitive energy as our kinetic energy because it involves the "generalized velocity" $\dot{\phi}$. We can now follow our standard process of determining the equations of motion for ϕ from the Euler-Lagrange equation for (2.89). This will give us

$$\ddot{\phi} + \omega_0^2\phi = 0, \tag{2.90}$$

where $\omega_0 = 1/\sqrt{LC}$ is the *angular resonant frequency* of our LC circuit. We see that this matches the result for the resonant frequency of a LC circuit that we are already familiar with from circuit theory. A similar result can also be obtained from the Hamiltonian approach to analyzing this simple circuit, which we leave as an exercise for the reader.

2.5 Practice Problems

1. We wish to determine the shape of the soap film that will be supported between two coaxial rings, as shown in Fig. 2.3. Physically, the soap film will adopt the shape that *minimizes* its free energy, which for this case will also correspond to minimizing its surface area. From the axial symmetry of the rings, we also expect that the solution will be a surface of revolution. Considering this, the surface area functional for our current case can be found as the integral of the radius of the film times the arc length. Mathematically, this is

   $$S[y] = 2\pi\int_{x_1}^{x_2} y\sqrt{1+(y')^2}dx, \tag{2.91}$$

   where y is a function of x and y' is shorthand for dy/dx.

   (a) Show that the Euler-Lagrange equation that minimizes $S[y]$ is

   $$\frac{1}{\sqrt{1+(y')^2}} - \frac{yy''}{\left(1+(y')^2\right)^{3/2}} = 0. \tag{2.92}$$

Note: you may use the results derived from class, so you do not need to explicitly compute the functional derivative of the functional given in (2.91). Instead, you can directly evaluate the derivatives called for in forming the Euler-Lagrange equation and fill them into the general expression to arrive at (2.92).

(b) Now, multiply the Euler-Lagrange equation by y' and show that it can be rewritten as

$$\frac{d}{dx}\left(\frac{y}{\sqrt{1+(y')^2}}\right) = 0. \tag{2.93}$$

Note: you can show this in the "forward" direction by simply evaluating the expression given in (2.93) and noting that it does equal y' times the Euler-Lagrange equation.

(c) Show that the problem in (b) reduces to the differential equation

$$\frac{dy}{dx} = \sqrt{\frac{y^2}{a^2} - 1}, \tag{2.94}$$

where a is an integration constant.

(d) Solve (2.94) by separating variables as

$$\int dx = \int \frac{dy}{\sqrt{(y/a)^2 - 1}}. \tag{2.95}$$

To aid in the solution, make the substitution $y = a\cosh t$, where cosh is a hyperbolic cosine (if you are unfamiliar with hyperbolic trig functions, which is quite likely, you can find useful identities on the Wikipedia page for "hyperbolic functions" that are *similar* to the "standard" trigonometric identities you are already familiar with). Your final answer should be

$$y = a\cosh\left(\frac{x+b}{a}\right), \tag{2.96}$$

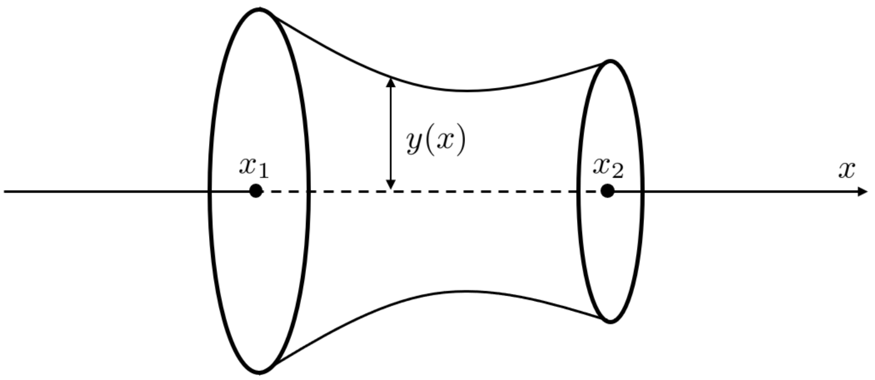


Figure 2.3: Illustration of the soap film problem discussed in Problem 3.

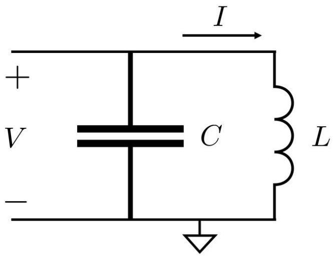


Figure 2.4: Circuit schematic of an LC tank circuit.

where b is another integration constant. If we had specified the radii of the two coaxial rings we could solve for the integration constants, but for simplicity we will leave our answer as is.

Note: we will **not** be solving differential equations like this throughout this course. This is just meant as a concrete example of how some of the machinery of the calculus of variations can be used to approach certain minimization problems.

2. For a Hamiltonian with n degrees of freedom, derive Hamilton's equations

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i} \tag{2.97}$$

using the definition of the Poisson bracket, where i is an index ranging from 1 to n.

3. Prove the following identity for Poisson brackets:

$$\{f_1 f_2, g\} = f_1\{f_2, g\} + f_2\{f_1, g\}. \tag{2.98}$$

4. Answer the following questions for the circuit shown in Fig. 2.4.

   (a) In terms of the nodal flux ϕ, the Lagrangian for this circuit is

   $$\mathcal{L} = \frac{1}{2}C\dot{\phi}^2 - \frac{1}{2L}\phi^2. \tag{2.99}$$

   Considering this, compute the "momenta" that is conjugate to ϕ. It is typical to denote this conjugate momenta as Q in the context of circuits such as this.

   (b) What does Q correspond to physically?
   Hint: look at the units of the expression defining Q.

   (c) Determine the Hamiltonian for this circuit.

   (d) Compute the equations of motion from the Hamiltonian and show that they are equivalent to

   $$\ddot{\phi} + \omega_0^2\phi = 0, \tag{2.100}$$

   where $\omega_0 = 1/\sqrt{LC}$.

5. In this problem, you will analyze the circuit shown in Fig. 2.5 using various methods. You may assume that the voltage $V(t)$ is a known quantity that is driving this circuit.

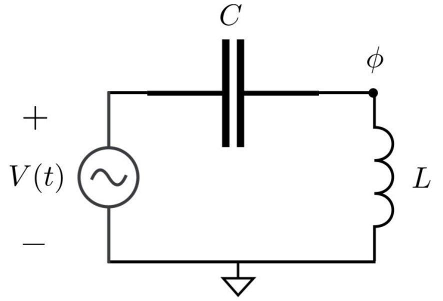


Figure 2.5: Circuit schematic of a voltage-drive series LC resonator.

   (a) What is the Lagrangian for this circuit only considering the energy stored in the capacitor and inductor? Your answer should be in terms of ϕ, $\dot{\phi}$, C, L, and $V(t)$.
   (b) Give an expression for the "conjugate momentum" for this circuit. Denote this quantity as Q.
   (c) What is the Hamiltonian of this circuit? Your answer should be in terms of ϕ, Q, C, L, and $V(t)$.
   (d) Determine the equation of motion for ϕ using the Lagrangian mechanics approach.
   (e) Now, show how a Hamiltonian mechanics approach can be used to derive the result you got for (d).
   (f) Use traditional circuit analysis techniques to derive the equation of motion for ϕ that you found in (d) and (e).

6. A particle of charge e is traveling in the presence of an electric and magnetic field. The Lagrangian for this case can be given as

$$\mathcal{L} = \frac{1}{2} m\dot{\mathbf{q}} \cdot \dot{\mathbf{q}} - e\Phi(\mathbf{q}) + e\dot{\mathbf{q}} \cdot \mathbf{A}(\mathbf{q}), \tag{2.101}$$

where $\mathbf{q}$ is the generalized position vector of the particle, $\Phi(\mathbf{q})$ is the scalar electric potential at position $\mathbf{q}$, and $\mathbf{A}(\mathbf{q})$ is the magnetic vector potential at position $\mathbf{q}$. For this Lagrangian, answer the following questions.

   (a) Verify that each term in the Lagrangian has units of energy. Note that the units of the magnetic vector potential are [Wb/m], where Wb is the abbreviation for Weber.
   (b) For the generalized position $\mathbf{q}$, show that the conjugate momentum $\mathbf{p}$ is

$$\mathbf{p} = m\dot{\mathbf{q}} + e\mathbf{A}. \tag{2.102}$$

   Note: This is *not* equal to the kinetic momentum of the charged particle.

(c) Show that the Hamiltonian for this system is

$$H = \frac{1}{2m}\big(\mathbf{p} - e\mathbf{A}\big) \cdot \big(\mathbf{p} - e\mathbf{A}\big) + e\Phi. \tag{2.103}$$

Note: This is often referred to as the *minimal coupling Hamiltonian*, and is used frequently in quantum mechanics and quantum electrodynamics.

Point of interest: Although the calculus required to do this is more complicated than we wish to consider in this course, the Euler-Lagrange equation for the Lagrangian given in (2.101) can be shown to be

$$m\ddot{\mathbf{q}} = e\big(\mathbf{E} + \dot{\mathbf{q}} \times \mathbf{B}\big), \tag{2.104}$$

where $\mathbf{E} = -\nabla\phi - \dot{\mathbf{A}}$ and $\mathbf{B} = \nabla \times \mathbf{A}$. This is referred to as the *Lorentz force law* in electromagnetism.

Chapter 3

Time-Independent Schrödinger Equation

3.1 Introduction to the Time-Independent Schrödinger Equation

Although we will mainly focus on analyzing quantum electrical circuits in this course, an introduction to quantum mechanical principles is not complete without spending some time discussing a common form of the Schrödinger equation for electrons. This introduces many important concepts in a (slightly) less abstract setting, and also highlights the wave properties of matter (such as electrons). These wave properties are a foundational principle that played a significant role in the development of quantum mechanics, and continue to be of vital importance in areas of electrical and computer engineering concerned with designing semiconductor electronics.

To motivate the development of the time-independent Schrödinger equation, we will consider a famous experiment (the *Davisson-Germer experiment*) that looked at the *diffraction* of electrons off a crystal. In this experiment, a *beam* of electrons was shined on a crystal sample and the intensity of the scattered electrons were registered as a function of position. The basic experimental setup is shown in Fig. 3.1.

If electrons *only* behaved as particles, one would perhaps expect that the resulting pattern of scattered electrons would follow a *ballistic path*, i.e., similar to how a ball might bounce off a wall. Instead, it was found that the scattered electrons followed a *diffraction pattern* that looked similar to what was familiar from other experiments with waves (e.g., with light) scattering from a *diffraction grating*. Since electrons were already known to behave as particles in other contexts, this was found to be another example of *wave-particle duality* that would become a common concept in quantum mechanics as it continued to be developed.

A diffraction grating is a kind of *periodic structure* that is composed of “small” scattering features that are spaced an equal length from each other. If this spacing has a period that is comparable to the wavelength of the wave that is incident on the grating, then we get some very interesting wave behavior evident in the scattered wave. In particular, we find that *different wavelengths* get scattered primarily in *different directions*. This occurs due to wave interference between all the “secondary waves” emanating out from each of the different

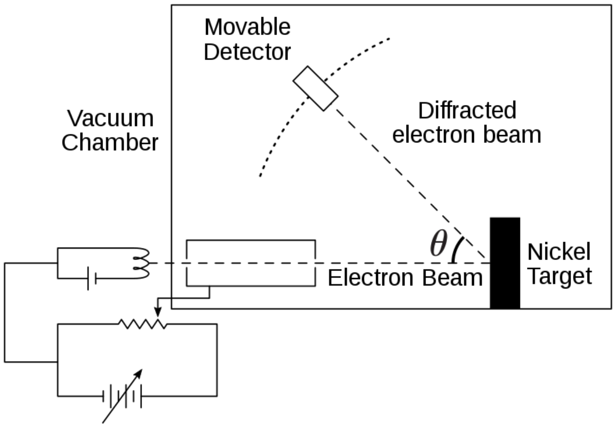
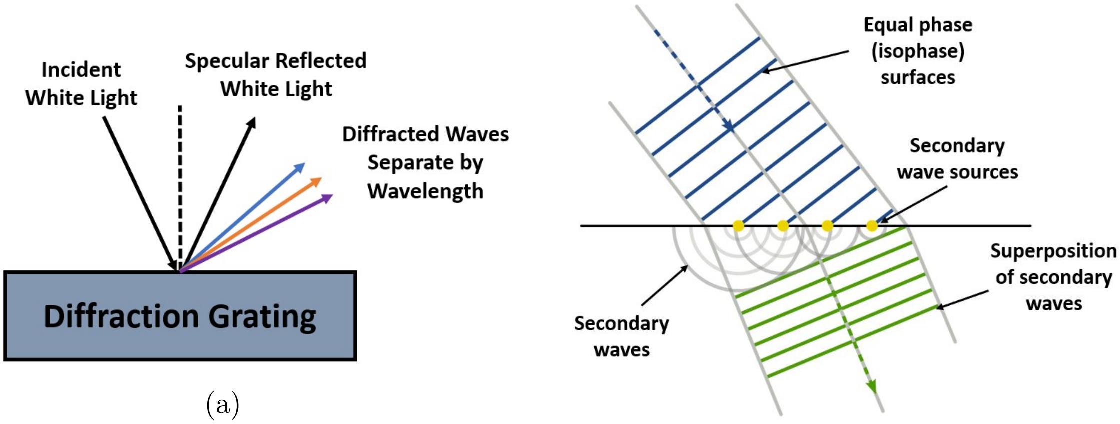


Figure 3.1: Schematic of the Davisson-Germer experiment (image from Wikipedia [7]).



(a)

(b)

Figure 3.2: (a) Simple schematic of light scattering from a diffraction grating and (b) illustration of Huygens' principle to understand scattering in terms of *secondary waves*.

scattering features in the grating (this is known as *Huygens' principle* in the general study of wave physics). This basic effect is shown schematically in Fig. 3.2. In the case of an electron scattering from a crystal, the "scattering features" of the crystal are the different atoms that are spaced along the *crystal lattice*. The highly ordered nature of the crystal is what leads to it being able to serve as a diffraction grating for an electron.

Now, it was eventually found that the diffraction patterns seen in these electron scattering experiments behaved like they had a wavelength given by

$$\lambda = h/p, \tag{3.1}$$

where p is the momentum of the electron and h is Planck's constant. The relation given in (3.1) is known as the *de Broglie hypothesis* due to it's role in Louis de Broglie's 1924 proposal that matter behaved like a wave. Using this wavelength, we can expect that in simple situations the electron could be described as a *plane wave* (a simple form of wave that has equal amplitude and phase over a planar surface in space). Then, we could propose a *wavefunction* for the electron of the form

$$\psi \propto e^{ikx} \tag{3.2}$$

if it were propagating along the x-axis, where k is the *wavenumber* given by

$$k = 2\pi/\lambda. \tag{3.3}$$

Note that this wavenumber has the same role as the $\omega = 2\pi/T$ that you are familiar with for describing the periodicity of sinusoidal signals in your other courses. It is also common to see de Broglie's hypothesis written in terms of the wavenumber, which we can see would give us

$$k = p/\hbar, \tag{3.4}$$

where $\hbar = h/(2\pi)$.

From other experiments, it was considered reasonable to treat the electron wavefunction as a *scalar wave* (i.e., it has a scalar amplitude instead of a vector one like an electric field). For this kind of wave, the simplest (and very common) wave equation for a monochromatic wave is the *scalar Helmholtz wave equation.* For our one-dimensional example, this wave equation is

$$\frac{d^2}{dx^2}\psi + k^2\psi = 0. \tag{3.5}$$

We can easily check that (3.2) is a solution of this equation by simple substitution. Now, using (3.4), we can rewrite our wave equation as

$$-\hbar^2\frac{d^2}{dx^2}\psi = p^2\psi. \tag{3.6}$$

If we divide both sides by $2m$, where m is the mass of our particle, then we would get that

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi = \frac{p^2}{2m}\psi. \tag{3.7}$$

We can recognize that the right-hand side of this equation is simply the *kinetic energy* of the electron from a simple Newtonian mechanics perspective *times* the wavefunction.

We now need to consider one of the "big jumps" in logic provided by Schrödinger in postulating his equation. In particular, he suggested that if the electron were traveling in some potential that depended on the position of the electron then we could account for this with a term in the form of $V(x)\psi$, i.e., the *potential energy* $V(x)$ *times* the wavefunction. We could then express the *total energy* in the system by adding our kinetic energy, expressed by the left-hand side of (3.7), and our potential energy together. This would give us

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\right)\psi = E\psi, \tag{3.8}$$

where E is the total energy. This is the famous *time-independent Schrödinger equation* for a one-dimensional system. If we were to extrapolate it to three-dimensions, we would have

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right)\psi = E\psi, \tag{3.9}$$

where $\nabla^2 = \partial_x^2 + \partial_y^2 + \partial_z^2$ is the *Laplacian operator* (in Cartesian coordinates) from vector calculus.

From our discussion of Hamiltonian mechanics, you may recognize that the left-hand side of (3.9) can be thought of as *one particular instance* of an "abstract" way of writing the Hamiltonian for our system (i.e., it has the form of $T + V$ we are familiar with). In a purely mathematical context, we may simplify our notation by defining a *Hamiltonian operator*

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r}) \tag{3.10}$$

that is essentially just a symbol to represent some set of mathematical steps we should take when we *apply* this operator to a function (much like why we called ∇^2 an *operator*). Note that it is very common in quantum mechanics to use the "hat" over a symbol to denote that it is an operator.

In this abstract notation, we would write (3.9) as

$$\boxed{\hat{H}\psi = E\psi.} \tag{3.11}$$

The *amazing* thing is that we will find that Schrödinger's equation written in the abstract form of (3.11) is generally applicable to a *huge* number of quantum systems. One of the main difficulties comes in determining the correct form of the Hamiltonian operator (and wavefunction) for a particular system. However, as we will see, this can be guided from the intuition of a Hamiltonian mechanics analysis of a classical system, which is why familiarity with the Hamiltonian mechanics framework is so vital to understanding most modern quantum analyses of next-generation quantum technologies.

We will also find that the broad concept of a *quantum mechanical operator* is a central component of working with quantum systems. We will return to this abstract concept in more depth when we discuss the mathematical framework of quantum mechanics, which is another necessary tool to help guide ourselves in developing a quantum theory for a particular system of interest. However, before taking this more abstract mathematical approach, it is useful to consider more examples of how to work with the Schrödinger equation in its simple differential form given in (3.8) or (3.9).

3.2 Wavefunctions and Probability

Although we have motivated the time-independent Schrödinger equation from a combination of basic wave physics and Hamiltonian mechanics, we find that the final form of our partial differential equation still involves solving for the wavefunction ψ. As a result, we need to have a better understanding of what exactly this wavefunction actually represents.

For our particular case of

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}+V(x)\right)\psi(x)=E\psi(x), \tag{3.12}$$

we found that our wavefunction may have a form similar to

$$\psi \propto e^{ikx}. \tag{3.13}$$

(This may be the case even when there is a potential $V(x)$ present for certain simple situations.) Considering this, we see that in general $\psi(x)$ may be a *complex-valued number*. As electrical and computer engineers, we are familiar with working with complex numbers to help simplify problems that only involve real-valued quantities; e.g., using phasors to express sinusoidal time domain waveforms. One fascinating detail about quantum mechanics is that the wavefunction *is* complex-valued in a general situation, and is not contingent on us simply electing to make it complex-valued to simplify a problem. *Rather, the wavefunction is complex-valued by its nature.*

Now, for what the wavefunction actually represents, it has been found that $\psi(x)$ has a *probabilistic interpretation.* More explicitly, we have that the probability of finding the electron near a particular location x is given by

$$P(x)=|\psi(x)|^2. \tag{3.14}$$

Due to this probabilistic interpretation of the wavefunction, it is sometimes common to refer to the value of the wavefunction at a particular point (i.e., $\psi(x)$) as the *probability amplitude.* It is important to remember that the actual probability is related to $|\psi(x)|^2$, which is then often referred to as the *probability density.* That this probability needs to be related to $|\psi|^2$ rather than just ψ is relatively obvious since we do not have a way to interpret what a complex-valued probability would mean. However, there are also significant implications to this result.

As an example of an important implication, we can readily see that if $\psi_1(x)$ and $\psi_2(x)$ are two solutions to (3.12) then their summation $\psi_1(x)+\psi_2(x)$ is *also* a solution to (3.12). We can then find that the resulting probability of finding an electron at location x would be

$$\begin{aligned}P(x)&=|\psi_1(x)+\psi_2(x)|^2\\&=(\psi_1+\psi_2)^*(\psi_1+\psi_2)\\&=\psi_1^*\psi_1+\psi_1^*\psi_2+\psi_2^*\psi_1+\psi_2^*\psi_2\\&=|\psi_1(x)|^2+2\mathrm{Re}\{\psi_1^*(x)\psi_2(x)\}+|\psi_2(x)|^2\end{aligned} \tag{3.15}$$

This is a very important result, as it shows that the *interference* between two wavefunctions can modify the probability of a particular "measurement" (e.g., in this case, determining the position of the electron). This kind of interference is behind the diffraction pattern seen from the Davisson-Germer experiment – in that case, we had to add all the wavefunctions associated with the "secondary waves" emanating from each scattering feature of the diffraction grating (again, this secondary wave concept is usually referred to as *Huygens' principle* in the physics of waves).

To see the impact of this interference more visually, let's consider a famous example of a *double slit experiment.* This is a historically important kind of experiment that has traditionally been used to demonstrate the wave nature of various physical theories. This was initially done for light by Thomas Young, and later for electrons and other quantum mechanical "particles" (atoms, molecules, etc.). In this experiment, we have a uniform wave incident on an opaque screen (meaning that the wave cannot pass through it) that has two very small slits in it where the wave can "leak" through the screen. Far away from this opaque screen at a position y_0 we have placed a "detecting screen" where we can measure what comes through the opaque screen. According to Huygens' principle, we can consider the net wave at the screen as a superposition of a "cylindrical wave" originating from each of the slits in the opaque screen. If we consider the first slit to be located at $x = x_0$ and the second slit to be located at $x = -x_0$, then the cylindrical waves emanating from the two slits evaluated at the detecting screen will be proportional to

$$\psi_1(x) = \frac{\exp\left[ik\sqrt{(x-x_0)^2+y_0^2}\right]}{(x-x_0)^2+y_0^2}, \tag{3.16}$$

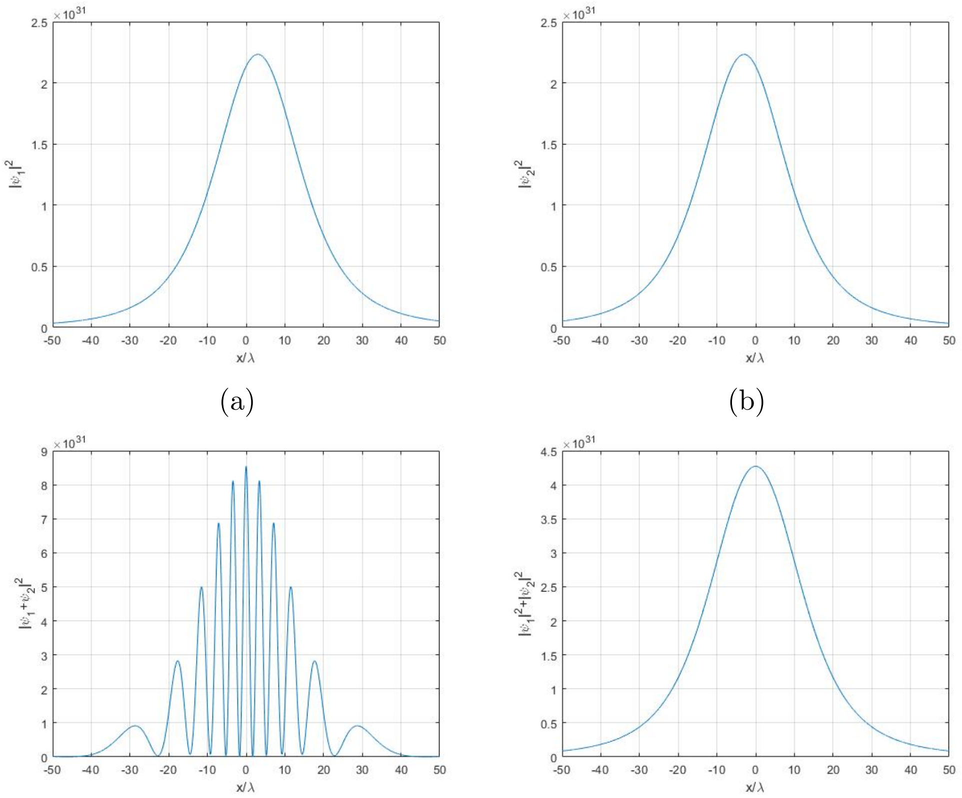
$$\psi_2(x) = \frac{\exp\left[ik\sqrt{(x+x_0)^2+y_0^2}\right]}{(x+x_0)^2+y_0^2}. \tag{3.17}$$

We will now consider a few different scenarios assuming that we have an incident electron speed of $10^6\,\mathrm{m/s}$ with the detecting screen placed 20λ from the opaque screen and the two slits separated by a distance of 6λ.

The first case we will consider is what happens when "slit 2" is covered so that no part of the wave can leak through it. In this case, the probability density that we measure will be purely due to ψ_1. Hence, if we evaluate $|\psi_1|^2$ over the region $-50\lambda \leq x \leq 50\lambda$ on the detecting screen we will get the "probability density" given in Fig. 3.3(a) (Note: we say "probability density" because we are using un-normalized wavefunctions, so the plotted result is only proportional to probability. We will discuss normalization of wavefunctions in more detail shortly). If we repeat this scenario but with "slit 1" covered instead of "slit 2" we will then get a result that only depends on ψ_2, which is shown in Fig. 3.3(b).

The more interesting case happens if we open both slits simultaneously. In this case, the total wavefunction is given by $\psi_1 + \psi_2$ and so we can now see interference between the two wavefunctions when we compute the "probability density". This is shown in Fig. 3.3(c), where we see that the interference leads to a number of "lobes" to the "probability density" where we are more likely to find the electron on the detecting screen. We also see that there are some locations where we will never observe the electron to land! This is a classic example of wave interference, and was a stunning result to see with something that was traditionally considered to be purely a particle (such as an electron).

To further emphasize the importance of the interference terms in evaluating the "probability density" we also plot the *incorrect* result that would occur if we just added the "probability densities" of ψ_1 and ψ_2 together. This would be given by $|\psi_1|^2 + |\psi_2|^2$ and is



(a) (b)

(c) (d)

Figure 3.3: Illustration of the result of various double slit experiments performed with an electron. The cases considered are for (a) "slit 2" is covered, (b) "slit 1" is covered, (c) both slits are open, and (d) the incorrect calculation of what happens with both slits open by simply adding the probability densities of the individual wavefunctions. Interference effects like those observed in (c) are crucially important in many quantum mechanical systems, including in many quantum algorithms that run on quantum computers!

shown in Fig. 3.3(d). Here, we clearly see that *no* interference occurs. As a result, it is essential to remember the correct calculations of probability densities to observe the quantum mechanical interference effects that are central to many quantum technologies!

The property that $\psi_1(x)+\psi_2(x)$ is a solution to Schrödinger's equation if $\psi_1(x)$ and $\psi_2(x)$ are independently solutions is a consequence of the *linearity* of Schrödinger's equation. You are likely familiar with the concept of linearity from many of your other engineering and mathematics courses. In many engineering applications, the linearity of a system is actually just an approximation that can break down under certain conditions (e.g., at high powers in an amplifier). However, the linearity of quantum mechanics is to the best of our knowledge a very *general and exact* property of quantum mechanics. In particular, it is the linearity with respect to the probability amplitude that is being solved for that seems to be an exact property of quantum mechanics. As a result, the tools of linear algebra play a very important role in our conceptualizing of how to analyze quantum mechanical systems. We will touch on this in much more depth when we discuss the mathematical framework of quantum mechanics.

Now, determining the probability of the location of the electron in a particular setup is of course of interest, however, there are also many other properties of our system that we may be interested in having a probabilistic description of. It turns out that the same wavefunction will be able to be used to describe the expected value of *all* of the possible measurable properties of our system. For instance, we will be able to use it to determine the expected value of the electron's momentum. Understanding *how* to use the wavefunction for these more general purposes is something that we will return to after building up some more tools for working with quantum systems.

It should be mentioned that we will not always be finding ourselves concerned with a wavefunction expressed as a function of position for an electron. However, the general concepts we have discussed will still hold true. In particular, we will still have probability amplitudes associated with our chosen description of our system, we will still have linearity, and we will have interference between possible probability amplitudes. For instance, later in this course we will touch on how quantum computers work by performing a sequence of operations to interfere probability amplitudes to arrive at a higher probability of a desired output from the computer!

Before moving on, it is useful to pause and emphasize some of the "weird" concepts we have just introduced. First and foremost, when we perform a quantum mechanical calculation we are only solving for the statistics of what will happen. Due to the underlying probabilistic nature of quantum mechanical systems, if we prepare two systems identically we are not guaranteed to get the same measured results each time. This goes beyond simple ideas of noise or a lack of control over our system, there is a profound and apparently fundamental random nature to quantum mechanical systems. Although this is weird, it has been verified experimentally that this is apparently how these systems work. We will return to discussing some of the interpretations of this odd result toward the end of this course. For now, we will mention that it is fortunate that these systems do behave in this perhaps unsettling statistical way. If they did not, many of the revolutionary quantum technologies being built would not be able to function in a superior manner to classical systems!

Now, considering that the probability amplitude will play such an important role in determining the probabilistic outcome of a quantum system, we will need to require an additional property of our probability amplitudes. In particular, we will require them to be *normalized*. In the context of the electron wavefunction, we have said that the probability of finding the electron at a position x was given by

$$P(x) = |\psi(x)|^2. \tag{3.18}$$

For $P(x)$ to be associated with a well-defined probability, we must have that

$$\int P(x)dx = \int |\psi(x)|^2 dx = 1. \tag{3.19}$$

In mathematics, we would refer to $P(x)$ as a *probability density*, since (rigorously) speaking it describes the probability of a particular result over some infinitesimal length.

Typically, when we first solve a differential equation like Schrödinger's equation we will find that our initial answer will lead to a wavefunction that *is not* normalized. In general, we would expect to have something that we can generically write as

$$\int |\psi(x)|^2 dx = |a|^2. \tag{3.20}$$

We then must take an additional step of *normalizing* our solution before using it in any quantum mechanical calculation. This can be done very easily by just defining our normalized wavefunction ψ_N as

$$\psi_N(x) = \frac{1}{a}\psi(x). \tag{3.21}$$

Note that because we will always need to work with normalized wavefunctions, we will not generally use the subscript N to denote this unless we are needing to make this explicitly clear in the particular context.

As a short aside, we typically refer to a function for which

$$\int |f(x)|^2 < \infty \tag{3.22}$$

as being *square integrable.* These kinds of functions play a very important role in many areas of physics and engineering. This generally occurs because we often find that the total energy of a system is proportional to the square of some value (for example, the kinetic energy of a particle or the energy stored in an electric field). As a result, square integrable functions generally also represent functions with finite energy, which are necessarily the kinds of functions we typically want to consider as being "valid" solutions to the kinds of differential equations we are solving. In mathematics, these are referred to as L^2 functions (pronounced as "ell-two", where the L here is in reference to the mathematician Henri Lebesgue). These are a special case of a more general set of functions called L^p functions, which find use in various areas of engineering (especially in modern signal processing). In general, the set of all L^2 functions is referred to as a *function space*, in similarity to the concept of a *vector space* from linear algebra. We will find that function spaces play a very important role in the mathematical framework of quantum mechanics, and so, we will often hear this kind of terminology used in various quantum mechanical contexts.

3.3 Solving the 1D Time-Independent Schrödinger Equation

We will now turn our attention to solving the one-dimensional time-independent Schrödinger equation for a few simple (but important) cases. This essentially reduces to a standard process for solving differential equations that describe waves, which is also applicable in many areas of physics. As a result, the mathematical concepts we will be covering in this section are not "that" quantum mechanical and can be picked up in other areas of engineering (e.g., these concepts are prevalent in engineering electromagnetics). The interesting implications from a physics perspective are linked to this standard wave-like behavior applying to an electron where before electrons were thought to only act as point-like particles.

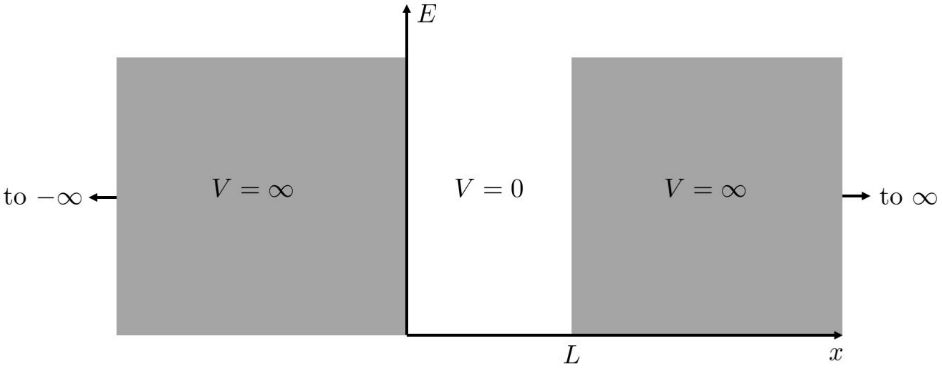


Figure 3.4: Illustration of the infinite potential well problem.

Considering that these wave physics concepts can be covered elsewhere, we will not spend a significant amount of time discussing them in this course. Instead, we will focus more on the broader concepts of quantum physics and interactions between different quantum systems that are more integral to the operation of next-generation quantum technologies than the wave behavior of electrons are. For those interested in learning more about the wave solutions to Schrödinger's equation, the reader is referred to many standard textbooks that cover this in greater depth; e.g., the introductory texts of [8, 9].

3.3.1 Infinitely-Deep Potential Well

We have already seen that if there is no potential energy term in Schrodinger's equation then the solution reduces to a simple "plane wave"-type solution. Of course, practical situations where there are no potential energy terms are rather limited, so we will now need to consider the more general case where a potential energy term does exist. We will begin with the simplest case; that of an infinitely-deep potential well.

The schematic of this situation is shown in Fig. 3.4. It consists of a region where the potential energy is taken to be 0 for $0 < x < L$, and the potential energy is assumed to be so large for $x < 0$ and $x > L$ that it can be considered infinite. Although the infinite potential energy is of course unrealistic, solutions to this problem will be a good approximation for cases where the potential energy is much larger than the energy of a particular solution (we will expand on what this means after solving the problem). One prevalent example of a practical scenario that exhibits physics similar to the infinitely-deep potential well is in semiconductor devices that use *quantum wells*. These systems consist of a semiconductor material (e.g., gallium arsenide) sandwiched between two layers of materials with wider bandgaps (e.g., aluminum gallium arsenide). This structure forms a potential well in one dimension, which can be used to manipulate the quantum effects in a device. This has been found to be useful in a large range of devices, but most notably in quantum well lasers that are a key component in fiber optic communication that serves as the backbone of the Internet.

Now, the basic process for solving a differential equation like Schrödinger's equation for this situation can typically be broken up into three basic steps. In *Step 1*, we first identify the different regions of our problem and the boundaries between them. In *Step 2*, we propose a suitable general solution to the differential equation in each region of the problem. These general solutions will involve some number of unknown constants in them, which we will need to determine the values of to solve the overall problem. This is done in *Step 3*, where we apply the *boundary conditions* at the interfaces between the different regions of our problem to find a set of (generally, but not always) algebraic equations to solve for the unknown constants.

In general, the process of specifying a suitable general solution can become rather complicated depending on the type of potential energy we are considering. However, it turns out that there are so few problems that can be exactly solved "by hand" that there are only so many general solutions we need to be able to actually remember. In this course, we will only focus on the *especially simple* or *especially important* general solutions. Details on other general solutions can be found as needed in standard textbooks or other mathematical resources.

Let's now apply our three-step process to solving Schrödinger's equation to this particular situation. We have already more or less done *Step 1* by drawing Fig. 3.4. More explicitly, we can say that

1. Region 1: $x \leq 0$, $V = \infty$,
2. Region 2: $0 < x < L$, $V = 0$,
3. Region 3: $x \geq L$, $V = \infty$.

Now, in *Step 2*, we need to propose a general solution in each region of the problem. This is relatively straightforward in Regions 1 and 3 for this case. In particular, because the potential energy is infinite the only way that Schrödinger's equation can be satisfied in that region is if $\psi(x) = 0$. For Region 2, we need to actually propose a non-trivial general solution. Since $V(x) = 0$ in this region, Schrödinger's equation reduces to

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi(x) = E\psi(x). \tag{3.23}$$

For convenience, it will be nice to "lump" all of the different constants in the problem on one side to come up with our general solution. This gives us

$$-\frac{d^2}{dx^2}\psi(x) = \frac{2mE}{\hbar^2}\psi(x) = k^2\psi(x), \tag{3.24}$$

where we have rewritten all of the constants as

$$k = \sqrt{\frac{2mE}{\hbar^2}} \tag{3.25}$$

for simplicity. As for choosing what our general solution should be, we can inspect this equation and readily see that whatever functions we choose for $\psi(x)$ will need to "return" back to what they started as after taking two derivatives (times some constants). Options that should jump to mind would be cosines, sines, or complex exponentials. For this particular case, the algebra will be easier if we use cosine or sine as our general solution (this is generally the case when the region is "closed" on both sides). Hence, we can write our general solution as

$$\psi(x) = A\sin(kx) + B\cos(kx). \tag{3.26}$$

We can now move on to *Step 3*, where we apply boundary conditions to determine our unknown constants in our general solution. For this problem, we can determine our boundary condition from the fact that the wavefunction must be *continuous* (this is related to the fact that we need to be able to differentiate it twice as part of Schrödinger's equation). Since we have already said that $\psi(x)$ is exactly 0 in Regions 1 and 3, we can find that the *boundary condition* for $\psi(x)$ will be

$$\psi(0) = 0 \quad \text{and} \quad \psi(L) = 0. \tag{3.27}$$

Beginning with the $\psi(0) = 0$ constraint, we quickly find that $\boxed{B = 0.}$ The next boundary condition is a little more nuanced to work with. In particular, we find that

$$\psi(L) = A\sin(kL) = 0. \tag{3.28}$$

We cannot choose $A = 0$ as our solution because this would lead to a trivial solution of $\psi = 0$ everywhere. Hence, we must instead have that

$$k = \frac{n\pi}{L}, \quad n \in \mathbb{Z}, \tag{3.29}$$

where $n \in \mathbb{Z}$ is a mathematical shorthand for specifying that n must be an integer (more explicitly, only positive integers will be of interest). The "trickiness" comes about because we have already specified that *k had to equal $\sqrt{2mE/\hbar^2}$ earlier in our solution.* The way "out" of this issue is that we have actually found that this problem will not be solvable for *any* value of E. Instead, we will only be able to have a valid solution when

$$E_n = \frac{\hbar^2}{2m}\left(\frac{n\pi}{L}\right)^2, \quad n \in \mathbb{Z}, \tag{3.30}$$

with the corresponding solution being of the form

$$\psi_n(x) = A_n \sin\left(\frac{n\pi}{L}x\right), \quad n \in \mathbb{Z}, \tag{3.31}$$

where A_n is a constant.

In most general problems (i.e., not in quantum mechanics), we will not be able to determine what A_n should be explicitly. However, *in quantum mechanics*, we have the additional requirement that our solutions be *normalized* so that we can successfully use our probabilistic interpretation of the wavefunction. In this particular case, we find that our normalization can be determined by evaluating

$$\int_0^L |\psi_n(x)|^2 dx = \int_0^L |A_n|^2 \sin^2\left(\frac{n\pi}{L}x\right) dx = |A_n|^2 \frac{L}{2}. \tag{3.32}$$

For this result to equal 1 (and thus be normalized), we need to choose

$$A_n = \sqrt{\frac{2}{L}}. \tag{3.33}$$

Hence, our normalized wavefunctions will then be

$$\psi_n(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi}{L}x\right), \quad n \in \mathbb{Z}, \tag{3.34}$$

which completes our mathematical solution to this problem. As a general note on terminology, we will usually refer to *each* of these possible wavefunctions as a *state* of the system.

From a physical perspective, we can conclude a few interesting things from this solution that are generally applicable when we confine a particle in some region of space (sometimes referred to as *quantum confinement*). First, we have found that the total energy of the particle can *only* take on a discrete set of possible values associated with a particular spatially-varying wavefunction. This marks a kind of *"quantization"*, i.e., the energy in our system can only change by discrete "steps". Each of these different values of energy are often referred to as *energy levels* of the system. Another interesting point is that there is a minimum possible energy for the particle confined in this potential well. In this case, it would correspond to a value of $E_1 = (\hbar^2/2m)(\pi/L)^2$. This lowest possible value is generally called the *zero point energy*. The fact that the zero point energy of a system is not generally "0" has some very profound consequences in how different quantum systems interact, e.g., an atom with an electromagnetic field (this leads to *spontaneous emission* of radiation, which we will learn about later in the course). Finally, we also see that the probability of finding the electron at a particular location is different for all of the different energy levels, as plotted for a few different cases in Fig. 3.5.

3.3.2 General Properties of Eigenvalue Problems

Typically, we refer to a differential equation like the time-independent Schrödinger equation as an *eigenvalue problem*. You are likely already somewhat familiar with eigenvalue problems from your studies of linear algebra. There, we describe the matrix equation

$$\overline{\mathbf{A}}\mathbf{v}_n = \lambda_n \mathbf{v}_n \tag{3.35}$$

as being an "eigenvalue problem" with *eigenvalues* λ_n and *eigenvectors* $\mathbf{v}_n$. In similarity, if we write the time-independent Schrödinger equation in its abstract operator form as

$$\hat{H}\psi_n = E_n \psi_n, \tag{3.36}$$

we see that it has a similar mathematical structure with eigenvalues of E_n and eigenvectors of ψ_n.

Generally, the terminology surrounding eigenvalue problems will take various different forms depending on the context. When we are working with a differential equation, it is common to refer to the eigenvectors as *eigenfunctions* or *eigenmodes*, although they still have the same underlying meaning. Sometimes, we will specifically refer to the ψ_n's of the

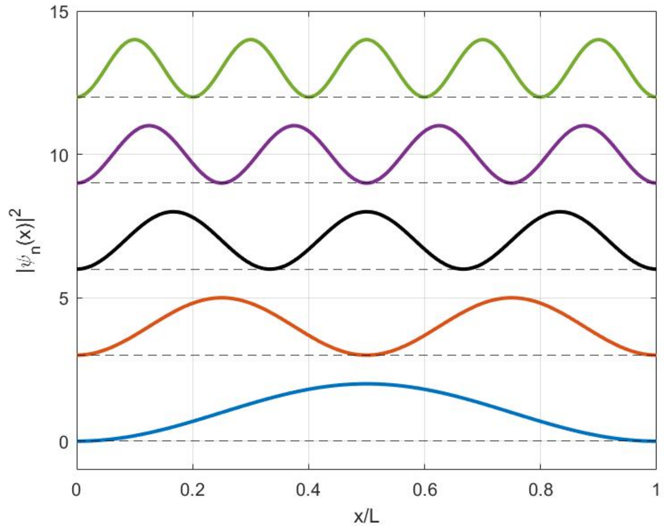


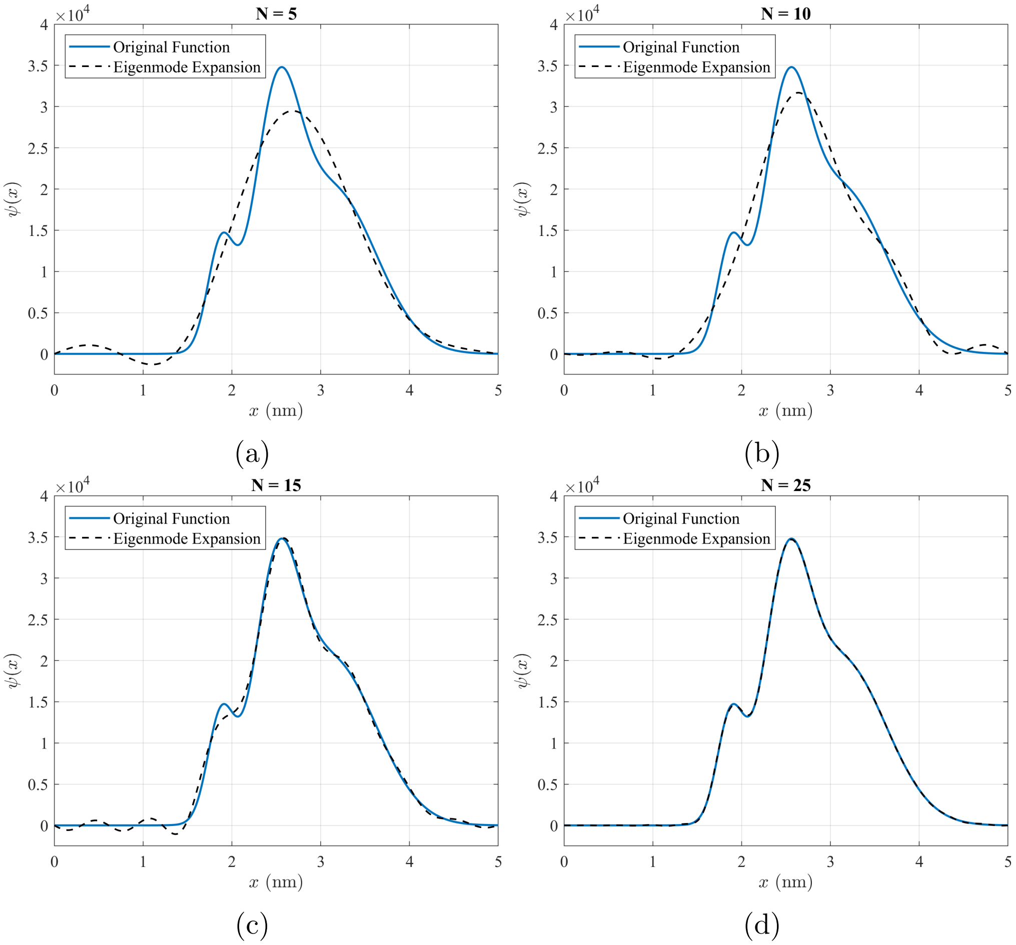
Figure 3.5: Plots of $|\psi_n(x)|^2$ for the first few energy levels of the infinite potential well. Note that the vertical offset between energy levels is only for clarity of plotting purposes, with each dashed line corresponding to the 0 probability line for a particular curve.

Schrödinger equation as *eigenstates* due to our terminology of referring to a general ψ as a *state* of the system. Likewise, we will also sometimes refer to the E_n's as *eigenenergies* rather than eigenvalues. Basically, it is common to add the prefix "eigen" on about anything we can think of in relation to an eigenvalue problem in quantum mechanics, much as it is common to add the word "quantum" in front of something to describe it (e.g., quantum superposition, quantum teleportation, quantum tunneling, etc.).

You will hopefully recall from your linear algebra studies that the eigenvalues and eigenvectors of a matrix (or more generally, a linear map) play a very important role in understanding the properties of that matrix. For instance, they can play an important role in diagonalizing a matrix and also are important in forming a *basis* for the vector spaces that the linear map operates on. In similarity to this, we will find that determining the eigenvalues and eigenvectors of the time-independent Schrödinger equation play a vital role in understanding a particular quantum system. More generally, we will find that determining the eigenvalues and eigenvectors of most *operators* that we can apply to one of our quantum states will be an important step in conceptualizing many results of quantum mechanics. As a result, we will come back to these concepts repeatedly throughout this course.

For now, we will mention a few important properties of the eigenvalues and eigenfunctions associated with the time-independent Schrödinger equation. These properties all come from the fact that the Hamiltonian operator $\hat{H}$ is an example of a special kind of operator, known as a *Hermitian operator* (we will define what this is more rigorously when we discuss the mathematical framework of quantum mechanics).

First, the eigenfunctions are *complete*. The exact notion of *completeness* can become somewhat technical, but at a high level what it means is that we can expand "any" function in terms of our eigenfunctions (there are some caveats to this statement, but these are not important at the level of mathematical rigor that we will require in this course). As a



(a) (b) (c) (d)

Figure 3.6: Illustration of how the eigenmodes of the infinitely-deep potential well can be used to recreate an arbitrary function. The recreations are performed when using the first (a) 5 eigenmodes, (b) 10 eigenmodes, (c) 15 eigenmodes, and (d) 25 eigenmodes. As more eigenmodes are included, the fidelity of the recreation continually increases.

result, we will be able to use our eigenfunctions as a *basis* to describe *any* possible state of our system. As a simple example of this, we show how an arbitrary wavefunction can be expanded in terms of the eigenmodes of the infinitely-deep potential well in Fig. 3.6. It is clear that as the number of eigenmodes used in the expansion is increased, the fidelity of the expansion can be made arbitrarily high. If you have already taken a linear systems course, this should come as no surprise to you as using the eigenmodes of the infinitely-deep potential well in this way is equivalent to a *Fourier series expansion*. In quantum mechanics, we often perform this type of expansion using eigenmodes from different quantum systems that do not correspond directly to a Fourier series. In this case, we will sometimes refer to the eigenmode expansion as being a *generalized Fourier series*. However, the important point is that the eigenmodes we will work with are *complete*, and so can be used to expand any wavefunction of interest with an arbitrarily high precision.

Second, the eigenfunctions are *orthogonal*. Mathematically, we will typically denote this as

$$\int \psi_m^*(x)\psi_n(x)dx = \delta_{mn}, \tag{3.37}$$

where δ_{mn} is a *Kronecker delta* that is equal to 1 when $m = n$ and is equal to 0 otherwise. Again, this property is fairly straightforward for the infinite potential well case since we are already familiar with sinusoids of different frequencies being orthogonal to each other. This orthogonality is also of significant importance for making the eigenfunctions a useful basis to expand a general state of our system in terms of.

To see how this works, we can assume that we have a general function $f(x)$. Since our set of eigenfunctions are *complete*, it is possible to write $f(x)$ in terms of them as

$$f(x) = \sum_n c_n \psi_n(x), \tag{3.38}$$

where c_n are called the *expansion coefficients*. We can determine what the mth expansion coefficient should be by multiplying by ψ_m^* and integrating. This gives

$$\begin{aligned}\int \psi_m^*(x)f(x)dx &= \int \psi_m^*(x)\left[\sum_n c_n\psi_n(x)\right]dx \\ &= \sum_n c_n \int \psi_m^*(x)\psi_n(x)dx \\ &= \sum_n c_n \delta_{mn} \\ &= c_m.\end{aligned} \tag{3.39}$$

If we were to try and expand $f(x)$ using a set of functions that were *not* orthogonal, then we would end up with a complicated “matrix” of relationships between how all of our basis functions “interact” with each other in the expansion. When we use an orthogonal set, this simplifies down to a “diagonal matrix” that is simple to work with to the point that we do not need to even bother with thinking of it as a “matrix” at all.

3.3.3 Quantum Harmonic Oscillator

We will now briefly discuss the time-independent Schrödinger equation for a quantum harmonic oscillator. We already discussed harmonic oscillators when learning about Lagrangian and Hamiltonian mechanics. There, we saw that the potential energy for the harmonic oscillator could be expressed as

$$V = \frac{1}{2}m\omega^2x^2, \tag{3.40}$$

where $\omega = \sqrt{k/m}$ was the natural oscillation frequency of the oscillator. Considering this, the time-independent Schrödinger equation that will describe a quantum harmonic oscillator can be written as

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi(x) + \frac{1}{2}m\omega^2x^2\psi(x) = E\psi(x), \tag{3.41}$$

which we see corresponds to having a "quadratic potential well".

Solving this problem is somewhat challenging because it requires the use of what are referred to as *special functions* that are relatively unfamiliar to engineers. Basically, a special function is a mathematical function that has an established name and notation because it has been found to be of particular importance for solving a specific type of problem, and, as a result, has been studied in depth by mathematicians. Special functions occur in solving differential equations quite frequently, and are especially prevalent in quantum mechanics.

In this particular case, we first simplify the notation by defining a dimensionless unit of distance of

$$\xi = \sqrt{\frac{m\omega}{\hbar}}x. \tag{3.42}$$

We can change to this variable and divide by $-\hbar\omega$ to write (3.41) as

$$\frac{d^2}{d\xi^2}\psi - \xi^2\psi = -\frac{2E}{\hbar\omega}\psi. \tag{3.43}$$

Through prior experience (i.e., not your prior experience, but the experiences of clever mathematicians), it is possible to find solutions to this equation of the form

$$\psi_n(\xi) = A_n e^{-\xi^2/2} H_n(\xi), \tag{3.44}$$

where H_n is a function that is still to be determined. This solution can be substituted into (3.43) to get (after some work) a differential equation that the H_n's must satisfy. In particular, we find that

$$\frac{d^2}{d\xi^2}H_n(\xi) - 2\xi\frac{d}{d\xi}H_n(\xi) + \left(\frac{2E}{\hbar\omega} - 1\right)H_n(\xi) = 0. \tag{3.45}$$

Admittedly, this looks like a mess, but it turns out that if

$$\frac{2E}{\hbar\omega} - 1 = 2n, \quad n \in \mathbb{Z} \tag{3.46}$$

then this differential equation is one that has been studied extensively in the past. Its solutions are known as *Hermite polynomials*, which are a particular example of a set of special functions (more specifically, a set of *orthogonal polynomials*).

Overall, we find that the valid eigenenergies for the quantum harmonic oscillator are

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega \tag{3.47}$$

and the normalized eigenfunctions are (after converting back to the original distance units)

$$\psi_n(x) = \sqrt{\frac{1}{2^n n!}\sqrt{\frac{m\omega}{\pi\hbar}}}\exp\left(-\frac{m\omega}{2\hbar}x^2\right)H_n\left(\sqrt{\frac{m\omega}{\hbar}}x\right). \tag{3.48}$$

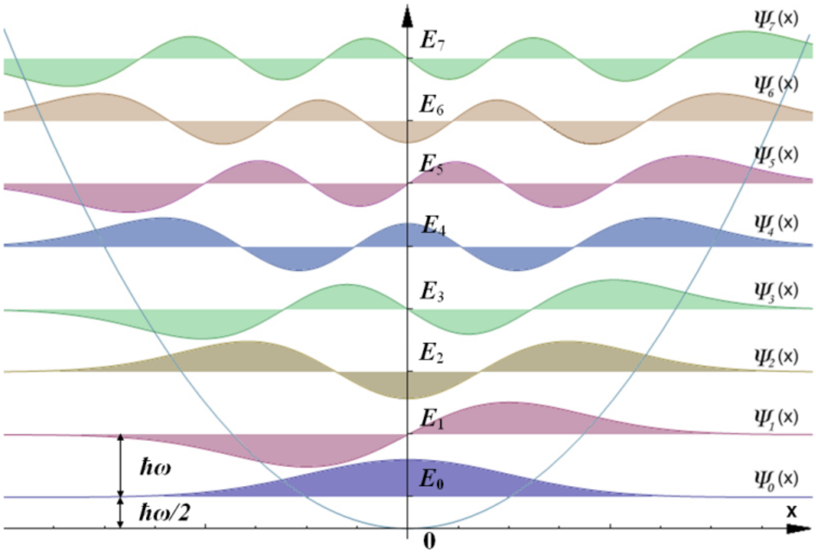


Figure 3.7: The first few eigenfunctions of the quantum harmonic oscillator. Note the definite *parity* of the functions (i.e., they are either even or odd functions) and how higher energies are able to have more probability density at larger values of x (image from Wikipedia [10]).

Remembering this exact expression for the eigenfunctions is not particularly important, but it is good to have an intuitive idea of what they "look" like when plotted. To help with this, the first few eigenfunctions are shown in Fig. 3.7.

The important takeaways from this example are the following. First, the eigenenergies are *equally spaced* for a quantum harmonic oscillator. This is a very rare phenomena, and has important consequences in next-generation quantum technologies, as we will discuss later in this course. Second, as with the infinite potential well case, we again see that we have a *non-zero zero point energy.* Finally, although we only considered this example as a special case for a potential well for an electron, it turns out that quantum harmonic oscillators are an excellent approximate description of a huge number of practical quantum systems. We will be more explicit about this later, but we will find that LC circuits behave like quantum harmonic oscillators, as do certain aspects of electromagnetic fields and acoustic vibrations along atomic lattices. As a result, the quantum harmonic oscillator is one of the most important "simple" quantum systems that we can analyze, whose properties have far-reaching effects in the design of modern quantum technologies.

3.4 Practice Problems

1. Explicitly evaluate the following expressions if $\psi_1(x) = \sqrt{2/L}\sin(\pi x/L)$ and $\psi_2(x) = \sqrt{2/L}\sin(2\pi x/L)$.

   (a) $\int_0^L \psi_1 \psi_2 dx$

   (b) $\int_0^L \psi_1 \frac{d}{dx} \psi_2 dx$

(c) $\int_0^L \psi_1 \frac{d}{dx}\psi_1 dx$

2. Suppose an electron has wavefunction $\psi(x) = \exp[ikx]$, where k is the wavenumber from de Broglie's hypothesis.

   (a) Show that this wavefunction is an eigenvector of the differential operator $\hat{p} = -i\hbar\frac{\partial}{\partial x}$.

   (b) From your answer to (a), what then is the eigenvalue of $\hat{p}$ equal to physically?

3. Normalize the following wavefunctions over the interval $-1 \leq x \leq 1$.

   (a) $\psi(x) = \frac{1}{2}(3x^2 - 1)$

   (b) $\psi(x) = \frac{1}{8}(35x^4 - 30x^2 + 3)$

   Note: these two polynomials are examples of *Legendre polynomials*. These find many uses in mathematical analysis and the solution of various differential equations.

4. Show that the two polynomials defined in 3(a) and 3(b) are orthogonal over the interval $-1 \leq x \leq 1$.

5. In this problem, we will consider details about generalized Fourier series expansions using eigenmodes of the time-independent Schrödinger equation for an infinitely-deep potential well that is 5 nm wide. In this potential well, we have been given the *un-normalized* wavefunction

$$\psi(x) = 0.5\exp\left[-\frac{(x-L/2)^2}{2(0.25\times10^{-9})^2}\right] + 0.4\exp\left[-\frac{(x-5L/8)^2}{2(0.5\times10^{-9})^2}\right] + 0.25\exp\left[-\frac{(x-3L/8)^2}{2(0.15\times10^{-9})^2}\right], \quad (3.49)$$

   where $L = 5\,\text{nm}$. Write a computer program to complete the following tasks, however, no use of symbolic computation (e.g., SymPy or Matlab symbolic math toolbox) is allowed.

   Note: so long as you use a reasonable number of points to discretize the interval $0 \leq x \leq L$ (e.g., 200) all integrations can be evaluated accurately using simple numerical integration functions like `trapz` in Matlab.

   (a) Normalize the wavefunction given in (3.49) and plot the result (Note: plot the wavefunction, not the probability density). It should look like Fig. 3.8.

   (b) Compute the generalized Fourier series expansion of the normalized wavefunction using the first 5 eigenmodes of the potential well. Plot the generalized Fourier series expansion and the normalized wavefunction on a single plot. The result should look like Fig. 3.6(a).

   (c) Repeat (b), but use the first 10 eigenmodes. The result should look like Fig. 3.6(b).

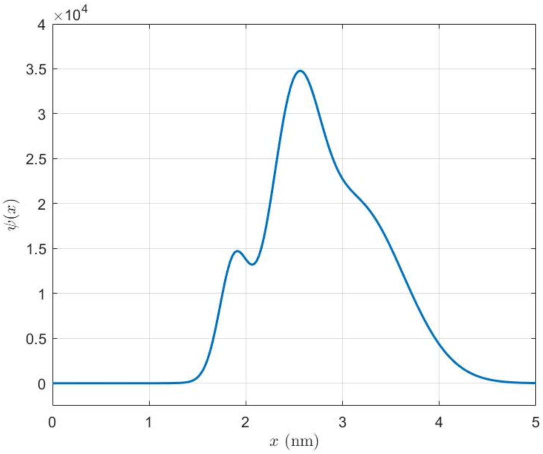


Figure 3.8: Normalized version of the wavefunction given in (3.49).

(d) Repeat (b), but use the first 25 eigenmodes. The result should look like Fig. 3.6(d).

6. Write a *finite difference method* program to discretize and solve the time-independent Schrödinger equation for an infinite potential well problem. In a finite difference method, the x-axis is discretized into a set of equally-spaced grid points, where Δx is the grid spacing. A derivative at the jth grid point is then approximated as

$$f'(j\Delta x) \approx \frac{f\big((j+1)\Delta x\big) - f\big((j-1)\Delta x\big)}{2\Delta x}, \tag{3.50}$$

and a second derivative as

$$f''(j\Delta x) \approx \frac{f\big((j+1)\Delta x\big) - 2f\big(j\Delta x\big) + f\big((j-1)\Delta x\big)}{(\Delta x)^2}. \tag{3.51}$$

(These approximations are known as *central differences*, and can be derived by truncating various Taylor series expansions.)

Using these approximations, the time-independent Schrödinger equation can be converted into a matrix equation, where you get one equation for each grid point. For instance, the matrix representation of the second derivative away from boundary points looks like

$$\frac{d^2}{dx^2} \rightarrow \frac{1}{(\Delta x)^2}\begin{bmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \iddots \\ \cdots & 1 & -2 & 1 & 0 & 0 & 0 & \cdots \\ \cdots & 0 & 1 & -2 & 1 & 0 & 0 & \cdots \\ \cdots & 0 & 0 & 1 & -2 & 1 & 0 & \cdots \\ \cdots & 0 & 0 & 0 & 1 & -2 & 1 & \cdots \\ \iddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \tag{3.52}$$

Some logic inspired by the knowledge of what happens in a region with an infinite potential can help you to determine how to find the complete finite difference discretization of d^2/dx^2 including next to the boundary points of the potential well.

Once the matrix form of the Schrödinger equation has been fully assembled, you can compute the eigenvalues and eigenvectors using built-in functions like Matlab's `eig`. Note that this entire process can be done in about 10 lines of code if you're efficient, but shouldn't take substantially more than that if you're not.

For consistency, do this problem for an electron confined in the interval $0 < x < 5\,\text{nm}$ and use a total of 400 grid points (excluding the boundary points of $x = 0$ and $x = 5\,\text{nm}$ which are accounted for implicitly since we know that $\psi(x) = 0$ at those points and so do not need to solve for its value there). Verify that the first 3 eigenvalues and eigenvectors closely match the analytical results derived in class. Note that the eigenvectors computed by the numerical method will not have the same normalization as the analytical solution. For comparison purposes, you will want to multiply the eigenvectors computed by the numerical method by $1/\sqrt{\Delta x}$.

7. For the following questions, assume that an electron has been confined inside an infinitely-deep potential well in the region $-L/2 < x < L/2$. For this potential well, the wavefunctions can be found to be

$$\psi_n(x) = \begin{cases} A\cos\left(\dfrac{n\pi}{L}x\right), & n \text{ is odd} \\ B\sin\left(\dfrac{n\pi}{L}x\right), & n \text{ is even} \end{cases}, \tag{3.53}$$

where A and B are constants to be determined and the energy associated with each state is given by

$$E_n = \frac{\hbar^2}{2m}\left(\frac{n\pi}{L}\right)^2, \tag{3.54}$$

where m is the mass of the electron.

   (a) Determine what A and B should equal so that the wavefunctions given in (3.53) are normalized.

   (b) Assume now that the initial state of the electron is given by

$$\psi(x,0) = \frac{1}{\sqrt{2}}\Big(\psi_1(x) + e^{i\varphi}\psi_2(x)\Big) \tag{3.55}$$

   and that the potential well is 10 nm wide. Evaluate the expectation value of the electron position $\langle x(t)\rangle$ at $t = 0.1\,\text{ps}$ and $t = 0.2\,\text{ps}$ given that the position operator $\hat{x} = x$ for this problem (i.e., it is just the x-coordinate) and that $\varphi = \pi/4$. You may compute the requested expectation value using hand calculations or by writing a simple numerical program (no symbolic computation tools allowed).

Chapter 4

Time-Dependent Schrödinger Equation

4.1 Introduction to the Time-Dependent Schrödinger Equation

Up to now, we have only considered the *time-independent* Schrödinger equation to begin understanding quantum mechanical systems. Obviously, a completely time-independent theory is not going to be able to fully describe quantum mechanical effects. Hence, we now turn our attention to discussing the *time-dependent* Schrödinger equation. As we will see, the results of the *time-independent* Schrödinger equation still typically play a vital role in building our understanding of the behavior of the *time-dependent* Schrödinger equation, and are also still of valuable use in practical calculations.

Thus far, we have worked with the time-independent Schrödinger equation in the form of

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}+V(x)\right)\psi(x)=E\psi(x), \tag{4.1}$$

which we saw defined an eigenvalue problem that could be solved for a particular potential energy to find the valid energies and wavefunctions for this system. The time-dependent Schrödinger equation looks very similar, and is given for this same kind of system (e.g., an electron subject to some potential energy) by

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}+V(x,t)\right)\psi(x,t)=i\hbar\frac{\partial}{\partial t}\psi(x,t). \tag{4.2}$$

In many situations, the potential energy may be time-independent (e.g., a fixed potential well) so that we can simplify (4.2) to

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}+V(x)\right)\psi(x,t)=i\hbar\frac{\partial}{\partial t}\psi(x,t). \tag{4.3}$$

In this case, we see that the only substantive change between (4.1) and (4.3) comes by "replacing" the energy by $i\hbar\partial_t$. It is because of this "small" change that we will find solutions of (4.1) are still very useful in understanding (4.3). Even when the total potential energy is not time-independent, it can still be a useful strategy to separate it into time-varying and time-stationary parts so that the time-independent Schrödinger equation is still useful. We will discuss how to deal with this kind of common occurrence later in this course when we discuss time-dependent perturbation theory in quantum mechanics.

Before moving on, we will comment that we can again write (4.3) in the same kind of more abstract operator notation that we discussed in relation to the time-independent Schrödinger equation. In this case, we would be able to write (4.3) as

$$\hat{H}\psi = i\hbar\frac{\partial}{\partial t}\psi. \tag{4.4}$$

This form of Schrödinger's equation is of significant interest as it is applicable to an extremely broad set of systems. As we discussed previously, we "simply" must find the suitable form of $\hat{H}$ and ψ and then we can use this equation to calculate the time-dependence of most practical quantum systems of interest for developing quantum technologies. For instance, even something as complex as the operation of a quantum computer can be thought of in terms of (4.4).

4.1.1 Stationary States

Let's now look at how a solution of the time-independent Schrödinger equation behaves if we consider it as the spatial part of a wavefunction in the time-dependent Schrödinger equation. In particular, we will assume that we have some eigenstate ψ_n with corresponding energy E_n. We will then assume as a "trial solution" that we can write our time-dependent wavefunction as

$$\psi(x,t) = c_n(t)\psi_n(x), \tag{4.5}$$

where $c_n(t)$ describes the time-dependence of the wavefunction that we would like to determine by solving (4.3). (Note: this kind of trial solution is known as a *separable solution.* We will comment more on this concept at the end of this section.)

If we plug (4.5) into (4.3), we will get

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\right)\big(c_n(t)\psi_n(x)\big) = i\hbar\frac{\partial}{\partial t}\big(c_n(t)\psi_n(x)\big). \tag{4.6}$$

By factoring c_n through the terms in the large parentheses on the left-hand side and ψ_n through ∂_t on the right-hand side, we get

$$c_n(t)\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\right)\psi_n(x) = i\hbar\psi_n(x)\frac{\partial}{\partial t}c_n(t). \tag{4.7}$$

We can recognize from the time-independent Schrödinger equation that

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\right)\psi_n(x) = E_n\psi_n(x), \tag{4.8}$$

so we can rewrite (4.7) as

$$c_n(t)E_n\psi_n(x) = i\hbar\psi_n(x)\frac{\partial}{\partial t}c_n(t). \tag{4.9}$$

We can now cancel the common terms of $\psi_n(x)$ on both sides of this equation to arrive at the very simple *ordinary* differential equation

$$E_n c_n(t) = i\hbar\frac{\partial}{\partial t}c_n(t). \tag{4.10}$$

You should recall from your differential equations course that the solution to this equation is

$$c_n(t) = c_n(0)e^{-i(E_n/\hbar)t}, \tag{4.11}$$

where $c_n(0)$ is a constant that comes from the initial conditions to the problem. If we assume that the initial condition is simply $c_n(0) = c_{n,0}$, then we can write our solution as

$$c_n(t) = c_{n,0}e^{-i(E_n/\hbar)t}. \tag{4.12}$$

We can then find that the time-dependence of (4.5) calculated by the time-dependent Schrödinger equation is

$$\psi(x,t) = c_n(t)\psi_n(x) = c_{n,0}\psi_n(x)e^{-i(E_n/\hbar)t}. \tag{4.13}$$

The state ψ_n is referred to as a *stationary state* because all properties of measurable values (typically referred to as *observables* in quantum mechanics) of this state are constant in time. As a simple illustration of this, we can calculate the probability density of finding our particle at a position x by evaluating

$$\begin{aligned} P(x,t) &= |\psi(x,t)|^2 \\ &= \left(c_{n,0}\psi_n(x)e^{-i(E_n/\hbar)t}\right)^* \times \left(c_{n,0}\psi_n(x)e^{-i(E_n/\hbar)t}\right) \\ &= |c_{n,0}|^2|\psi_n(x)|^2, \end{aligned} \tag{4.14}$$

which we see does not depend on time. (Note that in this very simple example $|c_{n,0}| = 1$ so that $\psi(x,t)$ is appropriately normalized.) Although we only showed this for one simple "observable", this property does hold in general for *all* observables of an eigenstate of the time-independent Schrödinger equation, and is the origin of the terminology *stationary state.*

Before moving on, it is worth briefly commenting on some of the general mathematical techniques that were utilized here. To begin, we usually refer to a function like

$$\psi(x,t) = c_n(t)\psi_n(x), \tag{4.15}$$

as being *separable* because we have been able to write its behavior in terms of independent functions that only depend on a single variable. Although this kind of solution is not possible for all partial differential equations, it is a very common strategy that works for many problems that can be solved "by hand". The general technique is known as the *separation of variables*, and is valuable because it allows us to simplify a complicated partial differential equation into a set of simpler differential equations that can be solved by themselves. In this case, our separation of variables in the form of (4.15) reduced the time-dependent Schrödinger equation into two simpler differential equations: namely, the time-independent Schrödinger equation of (4.1) and the temporal ordinary differential equation of (4.10). You will come across this general technique in many different areas of physics that must solve partial differential equations, including in the area of electromagnetics in electrical engineering.

4.1.2 Superpositions of Stationary States

We have just seen how solutions to the time-independent Schrödinger equation are stationary states whose observable properties do not change as a function of time. This is interesting, but obviously to describe a general quantum system evolving in time we are going to need to find solutions whose observable properties *do* evolve in time. To see one simple way for this to occur, let's start by looking at an example of a superposition of two stationary states.

In this case, we will assume that the initial state of our quantum system is given by

$$\psi(x,0) = c_{m,0}\psi_m(x) + c_{n,0}\psi_n(x), \tag{4.16}$$

where $c_{m,0}$ and $c_{n,0}$ are complex constants that "weight" the contributions of the two eigenstates ψ_m and ψ_n, which have corresponding energies E_m and E_n. For this to be a well-defined state, we will need it to be normalized. We can make sure this is the case by evaluating the integral of $|\psi(x,0)|^2$ and ensuring that its value is 1. To see what this works out to be, we note that

$$\int |\psi(x,0)|^2 dx = \int \Big[|c_{m,0}|^2\psi_m^*\psi_m + c_{m,0}^*c_{n,0}\psi_m^*\psi_n + c_{n,0}^*c_{m,0}\psi_n^*\psi_m + |c_{n,0}|^2\psi_n^*\psi_n\Big]dx \tag{4.17}$$

can be simplified greatly using the *orthonormality* of the normalized eigenstates ψ_m and ψ_n. This causes the cross terms like $\psi_m^*\psi_n$ to integrate to 0 and the common terms like $\psi_m^*\psi_m$ to evaluate to 1. The final normalization condition then reduces to

$$|c_{m,0}|^2 + |c_{n,0}|^2 = 1. \tag{4.18}$$

Assuming that this relationship holds for $c_{m,0}$ and $c_{n,0}$, we can now proceed with determining the time-dependence of our system given the initial condition of (4.16).

To do this, we propose

$$\psi(x,t) = c_m(t)\psi_m(x) + c_n(t)\psi_n(x) \tag{4.19}$$

as a "trial solution" to the time-dependent Schrödinger equation. We can substitute this in and separate terms out to get

$$c_m(t)\hat{H}\psi_m(x) + c_n(t)\hat{H}\psi_n(x) = i\hbar\frac{\partial}{\partial t}\big(c_m(t)\psi_m(x) + c_n(t)\psi_n(x)\big), \tag{4.20}$$

where we are using $\hat{H}$ as the shorthand operator notation to denote the Hamiltonian operator from the typical Schrödinger equation we have been working with. We can now use the properties of the different eigenstates to then arrive at

$$c_m(t)E_m\psi_m(x) + c_n(t)E_n\psi_n(x) = i\hbar\frac{\partial}{\partial t}\big(c_m(t)\psi_m(x) + c_n(t)\psi_n(x)\big). \tag{4.21}$$

Solving this equation looks slightly tricky at first glance, but we can isolate the time dependence of the different parts of the equation by multiplying by the conjugate of one of our eigenstates and integrating over the spatial region of interest. (As an aside, performing this kind of step is a very common in the formulation of numerical methods used to solve a wide variety of partial differential equations.) For instance, we can multiply by ψ_m^* and integrate to get

$$\int \psi_m^*\Big[c_m(t)E_m\psi_m(x) + c_n(t)E_n\psi_n(x)\Big]dx = \int \psi_m^*\Big[i\hbar\frac{\partial}{\partial t}\big(c_m(t)\psi_m(x) + c_n(t)\psi_n(x)\big)\Big]dx. \tag{4.22}$$

We can use the orthonormality of the eigenstates to evaluate the spatial integrals to then see that this reduces to

$$c_m(t)E_m = i\hbar\frac{\partial}{\partial t}c_m(t). \tag{4.23}$$

This will have as its solution

$$c_m(t) = c_m(0)e^{-i(E_m/\hbar)t}. \tag{4.24}$$

To determine what the initial condition is for $c_m(t)$, we can return to (4.16) and multiply by ψ_m^* and integrate. This gives us

$$\int \psi_m^*(x)\psi(x,0)dx = \int \Big[c_{m,0}\psi_m^*(x)\psi_m(x) + c_{n,0}\psi_m^*(x)\psi_n(x)\Big]dx, \tag{4.25}$$

which we can quickly recognize will evaluate to $c_{m,0}$ due to the orthonormality of the eigenstates. Hence, we find that the complete solution for $c_m(t)$ will be

$$c_m(t) = c_{m,0}e^{-i(E_m/\hbar)t}. \tag{4.26}$$

A similar set of steps allows us to find that

$$c_n(t) = c_{n,0}e^{-i(E_n/\hbar)t}. \tag{4.27}$$

Putting these results together, we finally have that our solution to the time-dependent Schrödinger equation is

$$\psi(x,t) = c_{m,0}\psi_m(x)e^{-i(E_m/\hbar)t} + c_{n,0}\psi_n(x)e^{-i(E_n/\hbar)t}. \tag{4.28}$$

Let's now look at the probability density of finding the particle at a particular position for this superposition of stationary states. We find that

$$\begin{aligned} P(x,t) = |\psi(x,t)|^2 = |c_{m,0}|^2|\psi_m(x)|^2 + |c_{n,0}|^2|\psi_n(x)|^2 \\ + c^*_{m,0}c_{n,0}\psi^*_m(x)\psi_n(x)e^{-i([E_n - E_m]/\hbar)t} + c^*_{n,0}c_{m,0}\psi^*_n(x)\psi_m(x)e^{i([E_n - E_m]/\hbar)t}, \end{aligned} \tag{4.29}$$

which now clearly exhibits time dependence due to the interference between the oscillations of the two stationary states. It is interesting to note that this interference oscillates at a frequency of

$$\omega_{mn} = (E_n - E_m)/\hbar, \tag{4.30}$$

where we have assumed that E_n is from a higher energy level so that the frequency is positive. We will find in many areas of quantum mechanics that the observable dynamical effects depend on frequencies associated with the *difference in energies* of the eigenstates involved. For instance, when an atom transitions between two energy levels the electromagnetic radiation that is absorbed or emitted as part of this process will be found to have an angular frequency defined by (4.30). Due to this dependence of the dynamics on the relative spacing of energy levels, we are often unconcerned with our "reference value" to measure energies from (i.e., what our 0 value is) since any constant shift to all our energies will not impact the dynamics of the system.

General Procedure

Although we have only considered the specific case of the superposition of two stationary states, it should hopefully be clear that this basic process immediately generalizes to as many stationary states as we need to work with for a particular problem due to their orthonormality. The *key result* is that due to the completeness of the eigenstates of the time-independent Schrödinger equation we can expand *any* valid initial spatial condition of the time-dependent problem in terms of these eigenstates. For example, if we have a specified initial condition of $\psi(x,0)$ we can readily compute the expansion coefficients of each eigenstate by evaluating

$$c_{n,0} = \int \psi^*_n(x)\psi(x,0)dx. \tag{4.31}$$

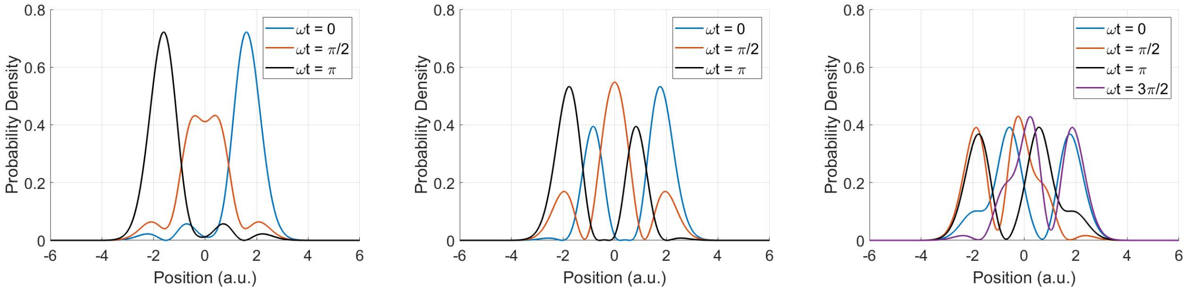
We can then write $\psi(x,0)$ in terms of the eigenstates as

$$\psi(x,0) = \sum_n c_{n,0}\psi_n(x), \tag{4.32}$$

subject to the normalization constraint that

$$\sum_n |c_{n,0}|^2 = 1. \tag{4.33}$$

(Note: checking your calculated coefficients to ensure they are normalized can be good test to make sure you evaluated your integrals correctly.) Now that we have expressed our initial



(a) (b) (c)

Figure 4.1: Time evolution of a quantum harmonic oscillator for (a) equal superposition of the first four eigenstates (b) equal superposition of states $\{\psi_0, \psi_2, \psi_3\}$, and (c) equal superposition but random phase of states $\{\psi_0, \psi_2, \psi_3\}$.

state in terms of the eigenstates of the system, the time evolution of this system can be easily found to be

$$\psi(x,t) = \sum_n c_{n,0}\psi_n(x)e^{-i(E_n/\hbar)t}. \tag{4.34}$$

Visualizing the interference between all of these terms can be relatively challenging, but the important point is that the time evolution can be found quite easily assuming we are able to solve the time-independent Schrödinger equation for our particular system of interest. To help visualize this, we will now look at some examples.

Example 1: Time Evolution in a Harmonic Oscillator

We now briefly look at the time evolution of states that correspond to a superposition of the first four eigenstates of the harmonic oscillator. We have previously discussed the spatial profile of these eigenstates and will not review that here. We will simply show the plots for a few different snapshots of the time evolution in Fig. 4.1. The main takeaway is that the time variation is periodic, as would be implied by (4.34), but that the spots where we are most likely to find the particle can significantly deviate from what we would expect for a classical particle trapped in this kind of potential well. We also see that the time variation can appear qualitatively quite different by simply changing the starting phase of the different expansion coefficients. This serves as a reminder to us that it is the complex-valued nature of the wavefunction that is so important in allowing for the interference between different eigenstates in the probability density.

Example 2: Time Evolution of a Coherent State in a Harmonic Oscillator

We have just seen that superpositions of the first few eigenstates of the harmonic oscillator can lead to trajectories that significantly differ from the expected classical motion. A natural question to then ask is whether there is some way for the quantum mechanical trajectory to closely follow the expected classical trajectory. It is sensible for us to expect that this must somehow be the case, since we expect that quantum mechanics as a more “fundamental”

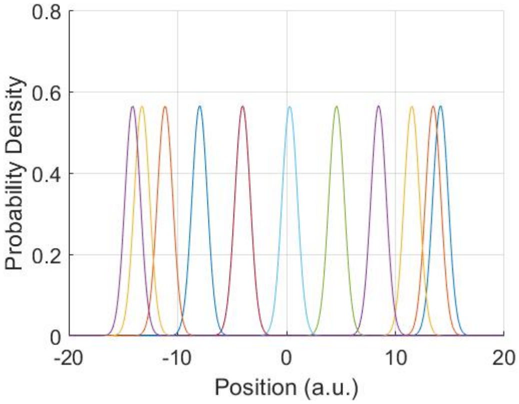


Figure 4.2: Snapshots of a coherent state oscillating in a quadratic potential well.

theory should still be able to (at some point) replicate the physics we experience in our day-to-day lives that classical mechanics describes. This is the basic idea behind the *correspondence principle* (due to Niels Bohr), which provides more structure to the notion of how quantum mechanics should reproduce classical physics. Typically, we will sometimes use the phrase the *classical limit* to describe the conditions for which the quantum and classical results agree.

In the context of the quantum harmonic oscillator, a very special superposition of eigenstates can be generated to most closely replicate the classical motion expected for this problem. This particular combination of eigenstates is known as a *coherent state*. Coherent states play a very important role in quantum mechanics. This is particularly the case in *quantum optics*, where they are found to provide a good description of widely-used electromagnetic fields, such as those produced by a laser. We will postpone a more detailed discussion of coherent states until later in this course, but will refer the interested reader to an introductory description in [8, Ch. 3] or a more advanced description that is useful in quantum optics [11, 12] for now.

An image showing the snapshots of the time evolution of a coherent state is included in Fig. 4.2. Although it is not easy to tell from the snapshots, the oscillations of the mean value of the coherent state occurs at the frequency of the classical motion. It is also seen that the spatial profile of the superposition of eigenstates maintains a consistent shape as it oscillates, which is in sharp contrast to the examples shown in Fig. 4.1.

4.2 Quantum Mechanical Measurements

Now that we have some understanding of time-dependent wavefunctions we can begin to discuss the very peculiar, but essential, role that *measurements* play in quantum mechanics. Although measurements form a very important piece of quantum mechanics, precisely defining what constitutes a "measurement" in a given context can be difficult. As a result, the "rules" surrounding quantum mechanical measurements can seem quite bizarre and can often lead to conceptual difficulties. This problem has puzzled physicists for many decades, and still continues to be an important question that is asked. However, recent work (well, in the 1980's) gave sufficient experimental evidence in support of one particular interpretation of quantum mechanics and measurements that has made that interpretation become the standard "correct" answer up to this point. This groundbreaking experimental work was so important that the main contributors were awarded the Nobel Prize in Physics in 2022 for their efforts. However, it should be mentioned that further work on these fronts continues to this day, and so our understanding of quantum mechanical measurements may yet develop in the future. Until then, we are left with what is typically known as the *Copenhagen interpretation*, which is associated with Niels Bohr and his followers.

We have already been following the Copenhagen interpretation up to this point based on our probabilistic interpretation of the wavefunction. We will now add to that interpretation an "explanation" of what happens when a quantum system is measured. In particular, when a measurement is performed, the system is said to *collapse* into an eigenstate of the quantity being measured (e.g., position, momentum, energy, etc.). If we have a wavefunction described by eigenstates ψ_n through

$$\psi = \sum_n c_n \psi_n, \tag{4.35}$$

then the probability that the wavefunction will collapse into state ψ_n upon measurement is determined by

$$P_n = |c_n|^2, \tag{4.36}$$

with the result of the measurement being the eigenvalue associated with the state ψ_n. We have already discussed states like (4.35) and have seen that their normalization constraint is that $\sum_n |c_n|^2 = 1$, so we see that the definition of (4.36) can indeed be interpreted as a probability. If the result of the measurement is the eigenvalue associated with state ψ_n, then immediately after the measurement has occurred the wavefunction of the system becomes

$$\psi = \psi_n; \tag{4.37}$$

i.e., it has *collapsed* into only the state ψ_n. Now, after the measurement the system begins to evolve in time again according to the time-dependent Schrödinger equation for the system. A consequence of this is that if we make a repeated measurement of the same quantity within a short enough time we are guaranteed to get the same result. For instance, we could measure the position of an electron, wait a *very* short time, and then measure the position again. If the evolution time was short enough, the wavefunction of the electron will not have enough time to "spread out" from the point of the first measurement, and so we will get the same result.

We can now consider preparing our quantum system in an identical manner many times and repeating our measurement process each time. Every time we perform the measurement, we will in principle get some "new" result based on the probability distribution described by the $|c_n|^2$'s. We can then determine the *expectation value* of our measurement based on all of these results. We can calculate what this expectation value would be using standard statistics, which give us

$$\langle A \rangle = \sum_n A_n P_n = \sum_n A_n |c_n|^2, \tag{4.38}$$

where A is the quantity we are measuring (e.g., energy, position, momentum, etc.) and A_n is the eigenvalue of this measurable quantity associated with the nth eigenstate. This is an important result, which we will return to shortly when we consider a slightly broader discussion of operators in quantum mechanics than we have considered up to this point.

Now, an *extremely important* (and weird!) aspect of the Copenhagen interpretation is that prior to us performing the measurement the system is in an "indeterminate state" described by the wavefunction. More explicitly, if we are measuring the position of an electron we cannot interpret the result of the measurement as being a consequence of the electron just already "happening" to be in the position where we found it. Rather, the electron *wasn't really anywhere* prior to the measurement, and it is the *act of measuring* the system that *forced* the electron into a particular position that we found it at. This is one of the central aspects of "quantum weirdness" at the heart of the famous Schrödinger's cat thought experiment, which was coincidentally postulated to try and express how *absurd* this view of the world could appear to be. However, this seems to be our best way to successfully perform quantum mechanical calculations to date, so we simply must live with this particularly weird aspect of quantum mechanics! It should also be emphasized that this apparent "weirdness" is central to the operation of many next-generation quantum technologies. For instance, if quantum systems didn't behave this way, many applications of quantum information (e.g., quantum communication and quantum computing) simply wouldn't be able to work!

4.2.1 Stern-Gerlach Experiment

To reinforce some of the important aspects of quantum mechanical measurements and wavefunction collapse, we will now discuss the famous *Stern-Gerlach experiment* and some key extensions to it. This experiment was originally performed in the early 1920's and played an important role in decisively demonstrating the quantization of particular aspects of atomic systems (in particular, the angular momentum). The basic setup of this experiment is shown in Fig. 4.3.

For simplicity, we will imagine that we are performing this experiment with a set of electrons rather than the more complex atoms that the experiment was originally performed with. In addition to this, we will need to have a basic notion of what *spin* is to understand these experiments. At a high-level, spin is a kind of *intrinsic angular momentum* of a particle that can take on only discrete values. The spin is an *intrinsic property* of the particle being considered, much like the charge of a particle is an intrinsic property, and does not depend on any actual movement of the particle (i.e., it is simply always there). For the purposes of this experiment, it suffices to understand that an electron can only have a spin of $\pm 1/2$ and that one consequence of spin is that it causes the electron to have what we can consider to be a small magnetic dipole moment.

In this experiment, an electron is "fired" through a device that produces an inhomogeneous magnetic field between the two poles of the magnet. If we consider the electron with its spin to behave like a tiny "bar magnet" we would expect the inhomogeneous magnetic field to deflect the trajectory of the electron depending on how the magnetic dipole moment is aligned with the inhomogeneous magnetic field of our device. For the purposes of this experiment, we will consider that the inhomogeneous magnetic field is inhomogeneous purely along one of our coordinate axes. As a result, if we take no effort to align the magnetic

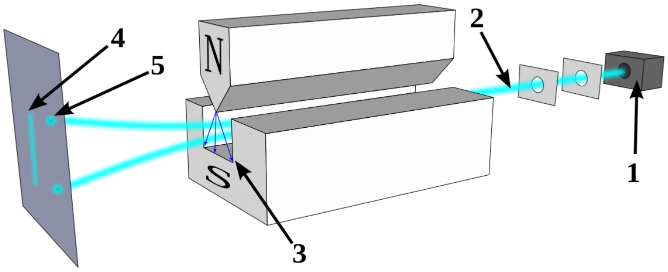
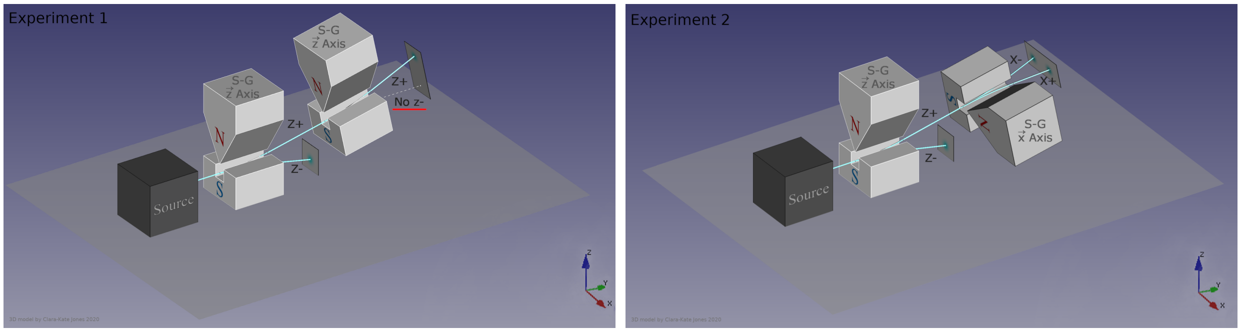


Figure 4.3: Basic configuration of the Stern-Gerlach experiment, where 4 denotes the expected classical result and 5 denotes the observed quantum mechanical result (image from Wikipedia [13]).

dipole moment of our electron prior to firing it through our system, we would expect that the distribution of points we find the electron at would be a continuous line along the direction of the inhomogeneous magnetic field. This expected distribution is marked by the number "4" in Fig. 4.3.

However, when we actually perform this measurement with a set of electrons we observe a result that significantly differs from our classical expectation. Instead of seeing a continuous set of deflections, we see that there are exactly two points that every electron fired through our setup deflects to. These points are marked by the number "5" in Fig. 4.3. The quantum mechanical explanation for this peculiar result is that our experimental setup has *measured* the spin of the electron as it passes through the magnetic field. More specifically, it has measured the spin along the axis of the inhomogeneous magnetic field. Since the spin property of the electron only has the two possible spin states of $\pm 1/2$, the measurement collapses our originally randomly prepared electron state into either a *spin up* or *spin down* orientation with respect to our measurement axis. After the wavefunction has collapsed into either the spin up or spin down orientation, the magnetic field then deflects the particle accordingly to one of the two positions where we see the electrons "pile up" in the actual measurement.

To emphasize a few other aspects of quantum measurements, it is instructive to consider a sequence of "Stern-Gerlach experiments" as illustrated in Fig. 4.4. The first experiment shown in Fig. 4.4(a) illustrates the *wavefunction collapse* property of repeated measurements; namely, that performing the same measurement in quick succession yields the same result. This is shown by taking one of the deflected electron paths from the first Stern-Gerlach experiment and passing this immediately into another Stern-Gerlach experiment oriented along the same axis. At the output of the second Stern-Gerlach experiment we see all the electrons "pile up" at a single point because the wavefunction has not had enough time to evolve between the two measurements to allow the second possibility to occur.



(a) (b)

Figure 4.4: Examples of sequential Stern-Gerlach experiments and the final detected results. Discussion of the two experiments is in the main text (images from Wikipedia [13]).

In the next experiment, shown in Fig. 4.4(b), we twist the second Stern-Gerlach experiment so that the magnetic field is oriented along a new orthogonal axis to our original measurement. The outcome of this measurement is again two deflected trajectories for all electrons. The reason for this is that our input state has its magnetic dipole moment oriented along an axis orthogonal to our measurement direction. The effect of the measurement is to project the magnetic dipole moment onto the axis of our measurement system, which in this case corresponds to an equal probability of the dipole moment being oriented along the + or − directions of our measurement axis. The end result is the two deflected trajectories that we observe.

4.3 Operators and Expectation Values

Earlier, one of the principles of quantum mechanical measurements that we discussed was that we can compute the *expectation value* of our measurement from

$$\langle A\rangle = \sum_n A_n P_n = \sum_n A_n |c_n|^2, \tag{4.39}$$

where A is the quantity we are measuring, A_n is the eigenvalue of this measurable quantity, and c_n is the expansion coefficient for the nth eigenstate of the measurable quantity. Although this formula is useful, it does have the drawback that it requires us to have expanded our wavefunction in terms of the eigenstates of the quantity we are measuring. To avoid needing to do this, it is useful to develop a more general method of computing expectation values.

The particular formula that achieves this is

$$\langle A(t)\rangle = \int \psi^*(x,t)\hat{A}\psi(x,t)dx, \tag{4.40}$$

where $\hat{A}$ is the operator associated with the quantity we are attempting to measure. To see an example of how this works, let's consider the case of calculating the expectation value of energy for a particular system. We have already seen that if we have expanded our wavefunction in terms of energy eigenstates as

$$\psi(x,t) = \sum_n c_n(t)\psi_n(x), \tag{4.41}$$

then the result for the expectation value of energy would be

$$\langle E(t)\rangle = \sum_n E_n |c_n(t)|^2. \tag{4.42}$$

We can also arrive at this result using our formula in (4.40). In particular, we initially have

$$\langle E(t)\rangle = \int \left(\sum_m c_m(t)\psi_m(x)\right)^* \hat{H}\left(\sum_n c_n(t)\psi_n(x)\right) dx. \tag{4.43}$$

Note that it is essential that when substituting summations like this into a general expression you ***must*** *use independent summation indices for each summation (the m and n in this case)!* Due to the linearity of our operators, we can apply the Hamiltonian operator to each of the eigenstates in the summation over n to get

$$\langle E(t)\rangle = \int \left(\sum_m c_m(t)\psi_m(x)\right)^* \left(\sum_n c_n(t)E_n\psi_n(x)\right) dx. \tag{4.44}$$

We can then flip the order of integration and summation to get

$$\langle E(t)\rangle = \sum_{m,n} \int E_n c_m^*(t) c_n(t) \psi_m^*(x)\psi_n(x) dx. \tag{4.45}$$

Using the orthogonality of the eigenstates, we can easily evaluate this integral to find

$$\langle E(t)\rangle = \sum_{m,n} E_n c_m^*(t) c_n(t)\delta_{mn}. \tag{4.46}$$

Due to the Kronecker delta function, we can eliminate one of these summations by forcing $m = n$ to finally arrive at

$$\langle E(t)\rangle = \sum_n E_n |c_n(t)|^2, \tag{4.47}$$

which matches our earlier result of (4.42).

Although we have only shown this result work for the expectation value of energy, the formula given in (4.40) holds *for any* measurable quantity in quantum mechanics (albeit, with some slight generalizations to the mathematical form that we will discuss when covering the mathematical framework of quantum mechanics). For now, it is important to emphasize why (4.40) is often a more convenient way to compute the expectation value than (4.39). The main disadvantage of (4.39) is that it *requires* us to have expressed our wavefunction in terms of the eigenstates of the particular operator we are computing the expectation value of. If we find ourselves wanting to evaluate the expectation value of multiple properties of our system (which is often a reasonable thing to do), we are then "stuck" with the laborious process of constantly rewriting our wavefunction in terms of different eigenstates. The advantage of (4.40) is that as long as we can evaluate the effect of $\hat{A}$ on $\psi(x,t)$ we can then compute the desired expectation value, completely avoiding any need to re-express our wavefunction in terms of different eigenstates.

As an example, another operator that is commonly used when working with the Schrödinger equation for a particle is the *momentum operator*. In our current setting, this would be given by

$$\hat{p} = -i\hbar\frac{d}{dx}, \tag{4.48}$$

which allows us to write our overall Hamiltonian as

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x) = \frac{\hat{p}^2}{2m} + V(x). \tag{4.49}$$

This form of expression makes this look even more like the Hamiltonian of a particle we are familiar with from classical mechanics. However, the important point is that the momentum operator defined in (4.48) can be easily applied to a general wavefunction $\psi(x,t)$ so long as we know how to evaluate the necessary derivatives. This can often be far simpler than first expressing our wavefunction in terms of momentum eigenstates and then using (4.39).

As a short aside, it is interesting to note the mathematical form of momentum eigenstates. By definition, we will need

$$\hat{p}\psi = -i\hbar\frac{d}{dx}\psi = p\psi. \tag{4.50}$$

It turns out that we have already encountered a form of ψ that would satisfy this property. In particular, it is the "plane wave" expression for an electron that we considered when originally rationalizing the Schrödinger equation. To see this, we recall that these wavefunctions were

$$\psi(x) = e^{ikx}, \tag{4.51}$$

where through the de Broglie relation $k = p/\hbar$. We can then explicitly evaluate

$$\hat{p}e^{ikx} = -i\hbar\frac{d}{dx}e^{ikx} = -i\hbar(ik)e^{ikx} = pe^{ikx} \tag{4.52}$$

to see that it is indeed the eigenstate of the momentum operator.

We can now think about expressing a particular eigenstate ψ_n of the time-independent Schrödinger equation in terms of momentum eigenstates rather than using a representation as a function of position (often called a *position basis* or *coordinate space representation*). Following our rule for evaluating the expansion coefficient for a particular state [e.g., see (4.31) as a motivating example], we find that we can determine the coefficients in the momentum representation for a particular value of k as

$$c_k = \int \left(e^{ikx}\right)^* \psi_n(x)dx. \tag{4.53}$$

Since k is a continuous variable, it makes more sense to index these coefficients like

$$C(k) = \int e^{-ikx}\psi_n(x)dx. \tag{4.54}$$

Inspecting this formula, we see that we can find the coefficients for all of the different momentum eigenstates by taking the Fourier transform of a state in the position basis! We can similarly think of taking the inverse Fourier transform of a state expressed in a momentum representation to return to a coordinate space representation. This correspondence ties into many deep concepts in quantum mechanics, like the *uncertainty principle*. We will return to these concepts in a more general setting after discussing the mathematical framework of quantum mechanics that generalizes some of the concepts we have just introduced.

4.4 Practice Problems

1. You have been given an electron in a 5 nm wide infinitely-deep potential well that is in a superposition of two low-lying eigenstates (for our purposes here, let's say $1 \leq n \leq 4$). You have observed that the electron probability density oscillates at a frequency of 29.097 THz (this is a *linear* frequency, not an *angular* frequency). Considering this, which two eigenstates is the electron in a superposition of?

   Note: there isn't a procedure to directly calculate which two eigenstates are involved. Rather, it will take some basic "trial and error" to see which frequencies of oscillation are possible. If you keep trying different combinations of eigenstates you will eventually land on the frequency mentioned above.

2. An electron in a 5 nm wide infinitely-deep potential well is placed in an equal superposition of the first three eigenstates. List all of the frequencies that the electron probability density function will be observed to oscillate at.

   Hint: you should find 3 different frequencies.

3. Repeat Problem 2 for an electron in a quadratic potential well with potential energy given by $V = \frac{1}{2}m\omega^2x^2$. Give your answer in terms of ω.

   Note: the hint from Problem 2 does not apply to this problem.

4. For the wavefunctions we have been working with up to this point, the position operator $\hat{x}$ can be found to just equal the regular position variable x. Considering this, evaluate $\langle x\rangle$ for an electron in a 5 nm wide infinitely-deep potential well for the following cases.

   (a) Perform the calculation "by hand" for the electron in the first eigenstate.

   (b) Perform the calculation by writing a simple numerical program for the electron in the second eigenstate.

   Note: this program should be something you write. It cannot be something evaluated via a tool like Wolfram Alpha or symbolic computation in Python/Matlab.

   (c) Provide a simple explanation for why the answers to (a) and (b) are the same. Your explanation should be more than just "that's what it evaluated to". It may help to look at Fig. 3.5 in coming up with your explanation.

5. Evaluate $\langle p(t)\rangle$ for an electron that has been placed in an equal superposition of the first two states of a 5 nm wide infinitely-deep potential well in the form

$$\psi(x,t) = \frac{1}{\sqrt{2}}\Big(\psi_1(x)e^{-i(E_1/\hbar)t} + \psi_2(x)e^{-i(E_2/\hbar)t}\Big). \tag{4.55}$$

   Make sure to write your answer in a manner that it is clear the result is real-valued and do not forget to label the units of your answer.

6. For the following questions, assume that an electron has been confined inside an infinitely-deep potential well in the region $-L/2 < x < L/2$.

   (a) Assume that the initial state of the electron is

$$\psi(x,0) = \begin{cases} \dfrac{2}{\sqrt{L}}\cos\left(\dfrac{2\pi}{L}x\right), & -\dfrac{L}{4} \le x \le \dfrac{L}{4}, \\ 0, & \text{elsewhere.} \end{cases} \tag{4.56}$$

   Determine what the initial value should be for the expansion coefficients of the first two stationary states of this potential well if we were to use these as part of a basis expansion to describe the time dynamics of this system.

   (b) Assume now that the initial state of the electron is given by

$$\psi(x,0) = \frac{1}{\sqrt{2}}\Big(\psi_1(x) + \psi_2(x)\Big), \tag{4.57}$$

   where ψ_n is the nth stationary state of the potential well. Evaluate $\langle x(t)\rangle$ and give your answer in a form that it is clear that the result is real-valued at all times.

   (c) Assume now that the initial state of the electron is given by

$$\psi(x,0) = \frac{1}{\sqrt{2}}\Big(\psi_1(x) + e^{i\varphi}\psi_2(x)\Big), \tag{4.58}$$

   where φ is a real-valued constant and ψ_n is the nth stationary state. Evaluate $\langle p(t)\rangle$ and give your answer in a form that it is clear that the result is real-valued at all times.

7. For this problem, consider a double slit experiment with electrons. For each statement below, determine whether the statement is true or false according to the Copenhagen interpretation of quantum mechanics. Clearly state your answer (i.e., true or false) and provide an explanation for why this should be the correct interpretation.

(a) It is possible to measure the location of the electron immediately before the slits, but having made this measurement we will no longer see an interference pattern on the detecting screen.

(b) It is possible to measure which slit each electron went through and see an interference pattern on the detecting screen.

(c) If we do not make a measurement before the detecting screen, each electron can be considered to have gone through one slit or the other and the averaging of the results of many experiments will produce an interference pattern at the detector screen.

8. At a given moment in time t_1, the state of an electron in an infinitely-deep potential well that is 10 nm wide is given by

$$\psi(x,t_1) = \frac{1}{\sqrt{8}}\psi_1(x) + \frac{i}{\sqrt{2}}\psi_2(x) + \frac{i}{2}\psi_3(x) + \frac{1}{\sqrt{8}}\psi_4(x), \tag{4.59}$$

where ψ_n is the nth stationary state of the potential well.

(a) What is $\langle E(t_1)\rangle$ for this system?

(b) Assume we allow the system to evolve freely in time until time $t_2 > t_1$. What is $\langle E(t_2)\rangle$?

(c) At time t_2, we now make a measurement of the energy of the system and get a result of E_2 (the energy of the second stationary state). Write down the state of the system just after this measurement has been made.

Chapter 5

Mathematical Framework of Quantum Mechanics

5.1 Introduction to Function Spaces

Up to now, we have focused on solving Schrödinger's equation for a charged particle in very simple settings. At times, we have alluded to the fact that the general form of Schrödinger's equation, written as

$$\hat{H}\psi = i\hbar\partial_t\psi, \tag{5.1}$$

will still be very useful when considering very complex systems like the operation of a quantum computer. To be able to consider these more sophisticated cases, we need to develop a suitable mathematical framework to better describe quantum mechanics of systems where a representation of $\hat{H}$ as a differential equation and ψ as a simple position-dependent function may no longer be relevant. This more general mathematical framework is described in terms of *function spaces* and their properties. Function spaces can be viewed as an extension of the concept of a vector space you will have learned about in a linear algebra course to the more complex mathematical objects of functions. Considering this, it will be useful to recall some basic properties of finite-dimensional vector spaces.

5.1.1 Finite-Dimensional Vector Spaces

You should recall from linear algebra that a vector space is a very broad concept. At a high level, we can say that a vector space can be defined by *anything* that satisfies the following general properties. In describing these properties, we must specify what *field* our vector space is defined over. A *field* is a general mathematical concept, but for our purposes we will only be concerned with cases where the field is the set of complex-valued numbers. We then refer to elements of the vector space V as *vectors* and elements of the field as *scalars.*

In terms of these, our vector space must have two well-defined operations that it is *closed* with respect to. The first operation is vector addition, and the second is scalar multiplication. For our vector space to be *closed* with respect to these operations, we require that the vector addition of any two vectors in our space returns a new vector in our vector space and that the scalar multiplication of any vector by any scalar also produces a new vector in our vector space. In addition to these two properties, there are *eight axioms* of a vector space. We will not review these in detail, but they include properties like the associativity or commutativity of vector addition, and that there are identity elements for vector addition (i.e., a zero vector exits) and scalar multiplication (e.g., the scalar 1).

The important result is that once we have determined that *something* is a vector space, we can use our powerful tools from linear algebra to think about and process aspects of a situation being considered. This includes tasks like determining whether a set of vectors are linearly independent, establishing a basis for our vector space, or developing a procedure to best approximate a vector using only a subspace of our overall vector space. All of these operations have many applications in a broad set of areas, especially in electrical engineering (e.g., signal processing and control theory make heavy use of many of these concepts).

For many practical purposes, engineers tend to try and work with *finite-dimensional vector spaces.* These are simply vector spaces for which a basis that spans the entire space can be built out of a finite number of vectors in the space. Unfortunately, many practical situations do not fit into this finite-dimensional picture particularly well. As a result, it has traditionally been a very important area of mathematics to determine how to work with *infinite-dimensional vector spaces.* In general, this can be a difficult subject to grasp. Fortunately, there are special examples of infinite-dimensional vector spaces for which much of our intuition built from finite-dimensional vector spaces is still typically applicable. The types of spaces that we will work with in quantum mechanics will generally fall into this important category of infinite-dimensional spaces that are easier to work with.

Before moving on to discuss these more complex spaces, it will be useful to discuss some typical notations used with vector spaces. You should recall that when we are going to be working with matrices, it is typical to write our finite-dimensional vectors as a *column vector* like

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{bmatrix}, \tag{5.2}$$

where N would be the dimension of the vector space and all the v_i's would be some element of the field the vector is defined with respect to. When working with complex-valued vector spaces, it is also typically very useful to take the *Hermitian transpose* (also called *conjugate transpose*) of a vector. This is generally denoted as $\mathbf{v}^\dagger$ and can be found by taking the transpose of the vector $\mathbf{v}$ and then taking the complex conjugate of every element of the vector (hence, the alternate name conjugate transpose). As an example, we would have

$$\mathbf{v}^\dagger = \begin{bmatrix} v_1^* & v_2^* & \dots & v_N^* \end{bmatrix}. \tag{5.3}$$

With this shorthand notation, we can write down important operations like the inner product of two vectors as

$$\mathbf{v}^\dagger \mathbf{w} = \begin{bmatrix} v_1^* & v_2^* & \dots & v_N^* \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix} = \sum_{i=1}^{N} v_i^* w_i. \tag{5.4}$$

In quantum mechanics, it is typical to use the *Dirac bra-ket notation* as a shorthand for working with all the different kinds of vectors that we will be concerned with. This notation takes a little time to become acquainted with in its most abstract form, but it is relatively simple to understand if we consider vectors like in (5.2). Then, we would denote our vector using a *ket* as

$$|v\rangle \to \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{bmatrix}. \tag{5.5}$$

It should be emphasized that the *ket* notation will be used in much more abstract settings where we may not be able to write our vector down as a simple column vector. Regardless, the ket is still meant to denote our abstract vector, where we will typically replace the symbol inside the $|\ \ \rangle$ to be some suggestive symbol to remind us what the ket represents. For instance, it will be very common to work with kets that are denoted as $|\psi\rangle$ to represent the state of our quantum mechanical system. In addition to the ket, there is the *bra* which is defined as the "Hermitian transpose" of the ket (interpreting exactly what the Hermitian transpose means for a particular vector must be determined by the context of what is being worked with). For our simple case here, the *bra* would be denoted as

$$\langle v| \to \begin{bmatrix} v_1^* & v_2^* & \dots & v_N^* \end{bmatrix}. \tag{5.6}$$

The inner product between two vectors would then be denoted by the shorthand

$$\langle v|w\rangle \to \mathbf{v}^\dagger \mathbf{w}. \tag{5.7}$$

You will become more comfortable with manipulating Dirac's bra-ket notation as you work with it more throughout this course. With these basics now understood, we can move to generalizing our notions of a vector space.

5.1.2 Function Spaces

A *function space* is a generic name for a broad class of vector spaces where the *vectors* are actually *functions*. If you think of the properties required of a vector space, it should not be too difficult to come to terms with the idea that functions could act as suitable vectors. For instance, we can easily take a vector $\mathbf{v}_1 = x^2$ and add to it $\mathbf{v}_2 = \exp(x)$ and see that the result is obviously just another function.

We have also already touched on other important aspects of function spaces when we discussed solutions to the time-independent and time-dependent Schrödinger equations. There, we discussed how the eigenfunctions of the time-independent Schrödinger equation were *complete* so that we could use them to express any other function of interest as a linear combination of eigenfunctions. As a result, we can consider these eigenfunctions to form a *basis* for the function space defined by all solutions to Schrödinger's equation for a particular system. Considering that there are an infinite number of eigenfunctions to the Schrödinger equation, we see that this then is an example of an infinite-dimensional function space (which we have already been successful in working with).

As an example of how the Dirac bra-ket notation will typically be used for this more general setting, consider the wavefunction

$$\psi(x) = \sum_n c_n \psi_n(x), \tag{5.8}$$

where ψ_n is an eigenfunction of the relevant Schrödinger equation. In quantum mechanics, we can use the ket $|\psi\rangle$ as a shorthand for (5.8), or to denote a different way to express the same function, e.g.,

$$|\psi\rangle \iff \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \end{bmatrix}, \tag{5.9}$$

where the length of this "column vector" of expansion coefficients would be infinitely long. As mentioned previously, we can use a variety of different symbols inside the ket notation to denote a different kind of vector. For instance, we might write

$$|1\rangle \iff \psi_1(x) \iff \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{5.10}$$

as a shorthand for the first state of the system we are considering.

Likewise, we can use the bra $\langle\psi|$ to denote equally well the expression

$$\langle\psi| \iff \sum_n c_n^* \psi_n^*(x) \tag{5.11}$$

or

$$\langle\psi| \iff \begin{bmatrix} c_1^* & c_2^* & c_3^* & \ldots \end{bmatrix}. \tag{5.12}$$

Whether we choose to use the expression (5.11) or (5.12) just depends on the context and on what calculation we are attempting to perform at any given moment. As a result, we often find it much more convenient to write expressions using the bra-ket shorthand rather than constantly carrying around the "baggage" of a particular representation of each of the vectors involved. Once we simplify things as far as we can using general mathematical properties of our vector spaces we can then choose a particular representation of our vectors to finish computing a result.

As a simple example, we can consider evaluating the inner product $\langle\psi|\psi\rangle$. If we use a representation like (5.8), we would evaluate this as

$$\langle\psi|\psi\rangle = \int \Big(\sum_m c_m \psi_m(x)\Big)^* \Big(\sum_n c_n \psi_n(x)\Big) dx = \sum_{m,n} \delta_{mn} c_m^* c_n = \sum_n |c_n|^2. \tag{5.13}$$

We can see that we will get the same result if we use an expression like (5.9) by evaluating

$$\langle\psi|\psi\rangle = \begin{bmatrix} c_1^* & c_2^* & c_3^* & \dots \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \end{bmatrix} = \sum_n |c_n|^2. \tag{5.14}$$

The key thing is that *regardless of how we calculate* $\langle\psi|\psi\rangle$ *we get the same result* because we have defined our representations of the bras and kets in an appropriate manner. It is then our job to determine the most suitable representation for a given problem to make the mathematics simpler.

As a further short example, we can also see how to express the orthonormality and calculation of expansion coefficients in terms of the bra-ket notation. In the first case, we can easily see that the orthonormality would be expressed by various expressions like

$$\langle\psi_m|\psi_n\rangle = \langle m|n\rangle = \delta_{mn}. \tag{5.15}$$

Similarly, we can determine an expansion coefficient by taking the inner product of our basis vector with the function of interest. This would be denoted as

$$c_m = \langle\psi_m|f\rangle \tag{5.16}$$

if we are computing the expansion coefficient for vector ψ_m of the function f.

5.1.3 Types of Function Spaces

As mentioned previously, working with infinite-dimensional function spaces can be rather challenging. In the course of considering these complicated spaces, mathematicians have found it useful to categorize function spaces into different groups based on additional properties they have.

One of the simplest examples of an "extra property" that is particularly useful for many engineering applications is that the function space have a *norm*. A function space that is complete and has a norm is referred to as a *Banach space*. A norm is a generalization of the geometric notion of the *length* or *magnitude* of a vector. Although it can be useful to compare the "lengths" of two functions, it is typically of most interest to use a norm as a *metric*; i.e., as a way to compute the "distance" between two functions. This provides us with a way of comparing functions within an abstract setting, e.g., by determining whether two functions are more similar to each other than another set of functions. Some of the most widely used norms come from the function spaces known as *Lebesque spaces.* These are denoted as L^p spaces, and have a norm defined by

$$||g||_p = \left(\int |g|^p dx \right)^{1/p}. \tag{5.17}$$

We would compute the similarity of two functions g and h using a norm by evaluating $||g - h||_p$. Typically, the most useful Lebesque spaces are those with $p = 1, 2$. These L^p spaces are particularly useful in signal processing and optimization applications, where being able to define the similarity of two functions is vital in the formulation and solution of many practical engineering problems.

Although having a norm is very useful, it turns out that being able to define an inner product for a function space is *extremely valuable.* This allows us to not only determine the relative "distance" between two functions, but to also generalize our geometric intuition about the angle between two vectors. Amazingly, if we do have an inner product in our function space, in many cases we can safely extrapolate our geometric intuition from working with something as simple as a three-dimensional Cartesian space to the abstract setting of working with an infinite-dimensional function space! As a result, inner product spaces are of intense interest. These are typically referred to as *Hilbert spaces*, and are the particular type of function space that quantum mechanics is mathematically described in.

5.2 Operators and Function Spaces

Now that we have a basic understanding of what function spaces are, we can begin to discuss the role of operators in function spaces. As we have already seen, operators can take various forms depending on the situation we are considering. At its most basic, we can think of an operator as a mathematical "object" that defines some set of operations that we should perform on a function in our function space that yields another function as its output. This definition is extremely broad, and obviously encompasses many different things. Just as the basic definition of a function space was broader than we will typically need, we will find that there are special kinds of operators that we will be most interested in for quantum mechanics.

5.2.1 Preliminaries

As a reminder, we have already seen that a simple example of an operator would be the derivative. For example, we can define the operator

$$\hat{A} = \frac{d}{dx}, \tag{5.18}$$

and write its operation on a function $f(x)$ in a generic manner as

$$g(x) = \hat{A} f(x). \tag{5.19}$$

In this case, we see that the output function $g(x)$ still "lives" in the same space as $f(x)$ in the sense that it depends on the same variable x. However, we can also have cases where our operator *maps* a function into a different kind of space. A common example of this would be a Fourier transform, which we could define as

$$g(y) = \hat{F}f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)e^{-iyx}dx. \tag{5.20}$$

(Note that physicists often define the Fourier transform with this $1/\sqrt{2\pi}$ normalization so that the only difference between the Fourier transform and its inverse transform is a change of sign in the exponential function). These two examples also highlight that operators can involve only "local" values of the function, like for the differential operator in (5.18), or they can involve all values of the function, like for the Fourier transform in (5.20).

Earlier, we used the phrase *map* to describe how the Fourier transform "took" our input function and returned one in a new space. This term should "ring a bell" for you with respect to the concept of a *linear map* that you would have learned about during a linear algebra course. You should recall that one of the key aspects of a linear map was that we could always express its operation on a vector space through the use of a *matrix.* In a similar manner we can think of *linear operators* in our function spaces as being like a (potentially infinitely-sized) matrix.

For an operator $\hat{A}$ to be linear, we require that for any complex number c and any two functions f and h that

$$\hat{A}\Big[f(x) + h(x)\Big] = \hat{A}f(x) + \hat{A}h(x), \tag{5.21}$$

$$\hat{A}\Big[cf(x)\Big] = c\hat{A}f(x). \tag{5.22}$$

In quantum mechanics, we will be exclusively interested in linear operators, which is another consequence of the linearity of quantum mechanics. One important consequence of this is that we will be able to form a *matrix representation* of an operator in such a way that we can describe the effect of the operator on a function in the form of a matrix-vector product following the rules of linear algebra you are already familiar with.

5.2.2 Matrix Representation of an Operator

To see how to form the matrix representation of an operator, we need to establish the bases we will use to form it. As an example, we will assume the basis for the *domain space* (the space of functions/vectors we apply the operator to) to be composed of a set of $|\psi_n\rangle$ and the *range space* (the space of output functions/vectors) to be composed of a set of $|\varphi_m\rangle$. To form the matrix representation, we can begin by applying the operator $\hat{A}$ on one of our basis vectors $|\psi_n\rangle$ to get

$$|g\rangle = \hat{A}|\psi_n\rangle, \tag{5.23}$$

where $|g\rangle$ is some function in the range space. Now, because $|g\rangle$ is in the range space, we can always express it in terms of our basis $|\varphi_m\rangle$ as

$$|g\rangle = \sum_m A_{mn}|\varphi_m\rangle, \tag{5.24}$$

where we keep the index n on our coefficient A_{mn} to remind ourselves that this coefficient is due to the input vector $|\psi_n\rangle$. We can then isolate how the operator maps our basis vector $|\psi_n\rangle$ into the range space by taking the inner product of $|g\rangle$ with a range space basis vector $|\varphi_m\rangle$. In bra-ket notation, we would find that

$$\langle\varphi_m|g\rangle = \langle\varphi_m|\left(\sum_{m'} A_{m'n}|\varphi_{m'}\rangle\right) = \sum_{m'} A_{m'n}\langle\varphi_m|\varphi_{m'}\rangle = \sum_{m'} A_{m'n}\delta_{mm'} = A_{mn}. \tag{5.25}$$

We can then determine that the value A_{mn} is the coefficient for the matrix representation of $\hat{A}$ in the mth row and nth column. We can also see that by comparing the first expression in (5.25) with (5.23) that we can alternatively write

$$A_{mn} = \langle\varphi_m|\hat{A}|\psi_n\rangle. \tag{5.26}$$

We can in principle do this for all of our basis vectors $|\psi_n\rangle$ and $|\varphi_m\rangle$ to eventually find that the full matrix representation is

$$\hat{A} \iff \begin{bmatrix} \langle\varphi_1|\hat{A}|\psi_1\rangle & \langle\varphi_1|\hat{A}|\psi_2\rangle & \langle\varphi_1|\hat{A}|\psi_3\rangle & \dots \\ \langle\varphi_2|\hat{A}|\psi_1\rangle & \langle\varphi_2|\hat{A}|\psi_2\rangle & \langle\varphi_2|\hat{A}|\psi_3\rangle & \dots \\ \langle\varphi_3|\hat{A}|\psi_1\rangle & \langle\varphi_3|\hat{A}|\psi_2\rangle & \langle\varphi_3|\hat{A}|\psi_3\rangle & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \tag{5.27}$$

Finding the matrix representation of an operator is typically a very useful operation involved in solving complex quantum mechanical calculations. Generally, we must only use a finite-sized basis to represent our operator numerically, but when done properly we can still achieve very accurate numerical results.

An important point to note about operators is with respect to their *commutativity.* You should recall from linear algebra that matrix multiplication is generally *not commutative.* That is, we generally have that

$$\overline{\mathbf{A}}\,\overline{\mathbf{B}} \neq \overline{\mathbf{B}}\,\overline{\mathbf{A}} \tag{5.28}$$

for matrices $\overline{\mathbf{A}}$ and $\overline{\mathbf{B}}$. Since a general matrix will have non-square dimensions, it should be easy to convince yourself of this fact as it could very easily be the case that while $\overline{\mathbf{A}}\,\overline{\mathbf{B}}$ is well-defined via matrix multiplication the product $\overline{\mathbf{B}}\,\overline{\mathbf{A}}$ may be completely meaningless. Considering that we can always form a matrix representation of our linear operators, it should hopefully not surprise you then that linear operators will also typically not commute. This is such an important concept in quantum mechanics that we will frequently work with the *commutator* of two operators, defined as

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}. \tag{5.29}$$

In fact, we will see later that there is a very deep connection between the commutator of two quantum operators and the Poisson bracket of the corresponding classical quantities. This correspondence is exploited frequently in determining how to “quantize” a particular classical theory (we will discuss this more later).

In addition to finding the matrix representation of an operator, another very important way to represent an operator is through a *bilinear expansion*. In essence, this can be viewed like a generalization of expanding a function in terms of basis vectors to expanding the operator in terms of the basis vectors. We will motivate the form of this expansion to illustrate some other useful operations with manipulating bra-ket notation. Considering this, we begin with

$$|g\rangle = \hat{A}|f\rangle \tag{5.30}$$

and the basis expansions of $|g\rangle$ and $|f\rangle$ given by

$$|g\rangle = \sum_m d_m|\varphi_m\rangle, \tag{5.31}$$

$$|f\rangle = \sum_n c_n|\psi_n\rangle. \tag{5.32}$$

From our earlier matrix representation of $\hat{A}$, we know that the d_m's can be computed from the c_n's as

$$d_m = \sum_n A_{mn}c_n. \tag{5.33}$$

We also know that the expansion coefficients c_n can be found by

$$c_n = \langle\psi_n|f\rangle. \tag{5.34}$$

Substituting this expression into (5.33), we find that

$$d_m = \sum_n A_{mn}\langle\psi_n|f\rangle. \tag{5.35}$$

We can substitute this new expression for d_m into (5.31) to then find that

$$|g\rangle = \sum_{m,n}\Big(A_{mn}\langle\psi_n|f\rangle\Big)|\varphi_m\rangle. \tag{5.36}$$

Recognizing that $\langle\psi_n|f\rangle$ is simply a scalar, we can rearrange this expression in a more suggestive form as

$$|g\rangle = \sum_{m,n}\Big(A_{mn}|\varphi_m\rangle\langle\psi_n|\Big)|f\rangle. \tag{5.37}$$

Comparing this to (5.30), we see that we can write the operator as

$$\hat{A} = \sum_{m,n} A_{mn}|\varphi_m\rangle\langle\psi_n|, \tag{5.38}$$

which is known as the *bilinear expansion* of $\hat{A}$.

In your linear algebra course you may have discussed a similar expression for expanding a matrix through the use of an *outer product*. In contrast to the inner product that "collapses" two vectors into a scalar, an outer product "expands" the two vectors into a matrix. The outer product of two vectors can be easily computed through the matrix multiplication of vectors $\mathbf{v}$ and $\mathbf{w}$ as

$$\text{Outer Product} = \mathbf{v}\mathbf{w}^{\dagger}. \tag{5.39}$$

Expanding matrices through the outer products of vectors is very common in more advanced studies due to its role in the *singular value decomposition* of a matrix (we won't discuss singular value decompositions in this course, but it is a very valuable tool to learn about if you intend to pursue advanced studies in your career).

5.3 Some Important Types of Operators

We will now discuss a few examples of "important" kinds of operators that are used frequently in quantum mechanics. Although not an exhaustive list, the kinds of operators we do discuss will continue to acquaint us with common techniques used in manipulating quantum mechanical calculations.

5.3.1 Identity Operator

The identity operator is abstractly defined as

$$|\psi\rangle = \hat{I}|\psi\rangle \tag{5.40}$$

for any $|\psi\rangle$ that exists in the function space the identity operator is defined for. It's matrix representation is simply the identity matrix for the basis being used to define the representation.

Although this appears quite simple at face value, utilizing identity operators is a very useful technique in various quantum mechanical calculations. The particular form of the identity operator that is generally most useful is it's bilinear expansion, which in this case is also sometimes referred to as the *resolution of the identity operator*. This can be written as

$$\hat{I} = \sum_n |\psi_n\rangle\langle\psi_n|, \tag{5.41}$$

where the $|\psi_n\rangle$ are *any* complete basis on the function space of interest. It is important to note that in contrast to the previous bilinear expansions we discussed, this one only involves a single index that is being summed over. The reason for this is that the identity operator is "diagonal" in the basis we are using. In view of this, we will generally refer to an operator as being *diagonalized* if the bilinear expansion of it only involves a single index to sum over. However, it is also important to recognize that even though we think of (5.41) as being a "diagonalized" form of the operator, it's matrix representation need not be.

To illustrate this subtle point, let's look at the matrix representation of (5.41) expressed in a different complete basis composed of $|\varphi_m\rangle$. From earlier, we know that an element of the matrix representation in this new basis will look like $\langle\varphi_i|\hat{I}|\varphi_j\rangle$. We can focus on any of these arbitrary elements of the matrix and see that

$$\langle\varphi_i|\hat{I}|\varphi_j\rangle = \langle\varphi_i|\bigg(\sum_n |\psi_n\rangle\langle\psi_n|\bigg)|\varphi_j\rangle = \sum_n \bigg(\langle\varphi_i|\psi_n\rangle\bigg)\bigg(\langle\psi_n|\varphi_j\rangle\bigg). \tag{5.42}$$

This final result is "just" a scalar value (as it should be for being an element of a matrix representation), but will clearly *not* be a diagonal matrix since there is not guaranteed to be any kind of orthogonality relationship between the $|\psi_n\rangle$ and $|\varphi_i\rangle$.

To see how using the resolution of the identity operator can be useful in more complex cases, it is instructive to prove that the *trace* of an operator is independent of the basis used to express it. The trace is defined as the sum of the diagonal elements of an operator, which we would typically write using the orthonormal basis $|\psi_n\rangle$ as

$$\mathrm{Tr}\big(\hat{A}\big) = \sum_n \langle\psi_n|\hat{A}|\psi_n\rangle. \tag{5.43}$$

From this general expression, it is not at all obvious that the trace would be an intrinsic property of the operator $\hat{A}$ that does not depend on the basis used in expressing it. However, we can demonstrate this is the case by inserting the identity operator before or after $\hat{A}$ in our expression. This is guaranteed to not change the result, since $\hat{I}\hat{A} = \hat{A}$ by definition. If we expand $\hat{I}$ using a new basis $|\varphi_m\rangle$ we find that

$$\mathrm{Tr}\big(\hat{A}\big) = \sum_n \langle\psi_n|\hat{A}|\psi_n\rangle = \sum_n \langle\psi_n|\hat{I}\hat{A}|\psi_n\rangle = \sum_n \langle\psi_n|\bigg(\sum_m |\varphi_m\rangle\langle\varphi_m|\bigg)\hat{A}|\psi_n\rangle. \tag{5.44}$$

We can rearrange this as

$$\mathrm{Tr}\big(\hat{A}\big) = \sum_n \sum_m \bigg(\langle\psi_n|\varphi_m\rangle\bigg)\langle\varphi_m|\hat{A}|\psi_n\rangle, \tag{5.45}$$

which can be further rearranged as

$$\mathrm{Tr}\big(\hat{A}\big) = \sum_n \sum_m \langle\varphi_m|\hat{A}|\psi_n\rangle\langle\psi_n|\varphi_m\rangle \tag{5.46}$$

since all of the quantities in the above expression are simply scalars so their order doesn't matter. However, we now realize that we can rewrite this expression in the more suggestive form as

$$\mathrm{Tr}\big(\hat{A}\big) = \sum_m \langle\varphi_m|\hat{A}\bigg(\sum_n |\psi_n\rangle\langle\psi_n|\bigg)|\varphi_m\rangle. \tag{5.47}$$

We recognize the terms in the parentheses as being a resolution of the identity operator using the $|\psi_n\rangle$ basis, so that we can finally arrive at the result that

$$\mathrm{Tr}\big(\hat{A}\big) = \sum_m \langle\varphi_m|\hat{A}\hat{I}|\varphi_m\rangle = \sum_m \langle\varphi_m|\hat{A}|\varphi_m\rangle. \tag{5.48}$$

Comparing this to (5.43), we see that the result of the trace will be the same regardless of whether we use the $|\psi_n\rangle$ or $|\varphi_m\rangle$ basis to compute it. This is an important result in its own right, as we will see later on that the trace of operators plays an important role in quantum mechanical calculations. More generally, we often find that many *intrinsic properties* of operators that *do not depend* on the basis used are of significant value.

5.3.2 Unitary Operators

A broad class of operators that are very important in quantum mechanics are *unitary operators.* A unitary operator is one whose *inverse* is given by the Hermitian transpose of itself (in the context of operators, this is sometimes referred to as an *adjoint*). Mathematically, this is written as

$$\hat{U}^{-1} = \hat{U}^{\dagger}, \tag{5.49}$$

so that

$$\hat{U}^{\dagger}\hat{U} = \hat{U}\hat{U}^{\dagger} = \hat{I}. \tag{5.50}$$

One of the key properties of a unitary operator is that it does not change the "length" of a vector, or more generally, it does not change the inner product between two vectors. To see this, let's define two new vectors through the operation of $\hat{U}$ on vectors $|f\rangle$ and $|g\rangle$ as

$$|f'\rangle = \hat{U}|f\rangle, \quad |g'\rangle = \hat{U}|g\rangle. \tag{5.51}$$

We then wish to see that $\langle g'|f'\rangle = \langle g|f\rangle$. To do this, we need to recognize that taking the Hermitian transpose of $|g'\rangle$ follows the rules similar to a matrix transpose in that

$$\langle g'| = \langle g|\hat{U}^{\dagger}. \tag{5.52}$$

That is, we take the Hermitian transpose of each "matrix" and "flip" the order. Considering this, we then have that

$$\langle g'|f'\rangle = \langle g|\hat{U}^{\dagger}\hat{U}|f\rangle = \langle g|\hat{I}|f\rangle = \langle g|f\rangle, \tag{5.53}$$

as desired. If we do this with $\langle f'|f'\rangle$, we would see that $\langle f'|f'\rangle = \langle f|f\rangle$, which establishes the preservation of the length of the vector as well.

Unitary operators appear in quantum mechanics in a number of different contexts. One prevalent use is for changing the basis representation of a vector. If we have a vector represented using the expansion coefficients from a particular basis set $|\psi_n\rangle$ and wish to convert this into a vector of expansion coefficients for the basis set $|\varphi_m\rangle$, we can use a unitary operator to handle this. In particular, the unitary operator that will achieve this can be found using two resolutions of the identity as

$$\hat{U}_{\varphi\psi} = \sum_{m,n} |\varphi_m\rangle\langle\varphi_m|\ |\psi_n\rangle\langle\psi_n| = \sum_{m,n} \big(\langle\varphi_m|\psi_n\rangle\big)|\varphi_m\rangle\langle\psi_n|. \tag{5.54}$$

Intuitively, we can think of this operator as finding how "much" of an arbitrary function $|f\rangle$ lies along the vector $|\psi_n\rangle$. The factors $\langle\varphi_m|\psi_n\rangle$ then help us understand how much the basis vector $|\psi_n\rangle$ overlaps with one of our new basis vectors $|\varphi_m\rangle$ to then reconstruct $|f\rangle$ in our new basis. That this operator is unitary can be seen by evaluating

$$\begin{aligned}\hat{U}^\dagger_{\varphi\psi}\hat{U}_{\varphi\psi} &= \sum_{m',n'} \big(\langle\varphi_{m'}|\psi_{n'}\rangle\big)^* |\psi_{n'}\rangle\langle\varphi_{m'}| \sum_{m,n} \big(\langle\varphi_m|\psi_n\rangle\big)|\varphi_m\rangle\langle\psi_n| \\ &= \sum_{m,m',n,n'} \big(\langle\varphi_{m'}|\psi_{n'}\rangle\big)^*\big(\langle\varphi_m|\psi_n\rangle\big)|\psi_{n'}\rangle\langle\varphi_{m'}|\varphi_m\rangle\langle\psi_n|.\end{aligned} \tag{5.55}$$

We can proceed by recognizing that $\langle\varphi_{m'}|\varphi_m\rangle = \delta_{mm'}$ due to the orthonormality of the basis vectors. Also, because $\big(\langle\varphi_{m'}|\psi_{n'}\rangle\big)^*$ is just a scalar resulting from an inner product, we can recognize that this also equals $\big(\langle\psi_{n'}|\varphi_{m'}\rangle\big)$. Using these two properties, we get that

$$\begin{aligned}\hat{U}^\dagger_{\varphi\psi}\hat{U}_{\varphi\psi} &= \sum_{m,m',n,n'} \big(\langle\psi_{n'}|\varphi_{m'}\rangle\big)\big(\langle\varphi_m|\psi_n\rangle\big)|\psi_{n'}\rangle\langle\psi_n|\delta_{mm'} \\ &= \sum_{m,n,n'} \langle\psi_{n'}|\varphi_m\rangle\langle\varphi_m|\psi_n\rangle|\psi_{n'}\rangle\langle\psi_n|.\end{aligned} \tag{5.56}$$

We can now rearrange the order of summations to write this as

$$\hat{U}^\dagger_{\varphi\psi}\hat{U}_{\varphi\psi} = \sum_{n,n'}\langle\psi_{n'}|\left(\sum_m |\varphi_m\rangle\langle\varphi_m|\right)|\psi_n\rangle|\psi_{n'}\rangle\langle\psi_n|. \tag{5.57}$$

Recognizing the resolution of the identity operator using the $|\varphi_m\rangle$ basis, we can simplify this to become

$$\hat{U}^\dagger_{\varphi\psi}\hat{U}_{\varphi\psi} = \sum_{n,n'}\langle\psi_{n'}|\psi_n\rangle|\psi_{n'}\rangle\langle\psi_n|. \tag{5.58}$$

Using the orthonormality of the $|\psi_n\rangle$ basis, we can further simplify this to find

$$\hat{U}^\dagger_{\varphi\psi}\hat{U}_{\varphi\psi} = \sum_{n,n'}\delta_{nn'}|\psi_{n'}\rangle\langle\psi_n| = \sum_n |\psi_n\rangle\langle\psi_n| = \hat{I}, \tag{5.59}$$

which establishes the unitarity of $\hat{U}_{\varphi\psi}$, as desired.

Similarly, we can use a unitary operator defined like (5.54) to change the basis representation of an operator. In particular, we have that

$$\hat{A}_{\text{new}} = \hat{U}\hat{A}_{\text{old}}\hat{U}^\dagger. \tag{5.60}$$

We can interpret this formula by recognizing that $\hat{U}^\dagger$ takes a vector in our new basis and converts it back into the old basis. We can then apply $\hat{A}_{\text{old}}$ to this vector expressed in the old basis representation before converting the result back into our new basis using $\hat{U}$. In quantum mechanics, it can often be convenient to use unitary operators to transform the Hamiltonian operator into a new representation. Although this can be done rigorously, this is most frequently done when attempting to make some approximation to simplify the Hamiltonian of a complex system.

5.3.3 Time Evolution Operator

A special example of a unitary operator often used in quantum mechanics is the *time evolution operator*. This operator can be beneficial when we want to consider a somewhat more abstract description of the time evolution of a state of a quantum system. In certain cases, this is simply a *formal result* that illustrates the particular principle of something without necessarily being a viable computational approach. However, in other cases, this more abstract view can be very beneficial in providing us a way to manipulate a result into a form that is more conducive to improving our understanding of a system or developing a suitable numerical computation technique.

To develop the desired result, we return to the time-dependent Schrödinger equation written in its operator form as

$$\frac{\partial}{\partial t}|\psi(t)\rangle = -\frac{i}{\hbar}\hat{H}|\psi(t)\rangle, \tag{5.61}$$

where we have rearranged a few of the constants for later convenience. Looking at this equation, we see that it has a very similar format to the simple ordinary differential equations like

$$\frac{\partial}{\partial t}c_n(t) = -\frac{i}{\hbar}E_n c_n(t) \tag{5.62}$$

that we solved previously. Considering this, we may be tempted to solve (5.61) directly and write the answer as

$$|\psi(t)\rangle = e^{-i\hat{H}t/\hbar}|\psi(0)\rangle. \tag{5.63}$$

This looks nice and compact, but raises the fundamental question of whether writing something like

$$e^{-i\hat{H}t/\hbar} \tag{5.64}$$

is even meaningful (or allowed for that matter). It turns out that it is meaningful to write such an expression, but we first need to understand how to work with a *function of an operator*.

The basic idea of how to work with a function of an operator is that to evaluate something like

$$e^{-i\hat{H}t/\hbar}|\psi(0)\rangle \tag{5.65}$$

we must rewrite the function of the operator in terms of its Taylor series and then apply this to the function $|\psi(0)\rangle$. In the case of an exponential, we know that the Taylor series is

$$e^x = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \ldots, \tag{5.66}$$

so we interpret the exponential of an operator to mean

$$e^{\hat{X}} = 1 + \hat{X} + \frac{1}{2!}\hat{X}^2 + \frac{1}{3!}\hat{X}^3 + \ldots\ . \tag{5.67}$$

In using an expression like (5.67), we interpret the square (or higher power) of an operator as simply meaning that we apply the operator repeatedly two (or more) times. For instance, we have that

$$\hat{X}^2|\psi\rangle = \hat{X}\big(\hat{X}|\psi\rangle\big), \tag{5.68}$$

with the generalization to higher powers like $\hat{X}^m$ being straightforward.

Now, considering that an operator like $\hat{H}$ can have a complicated form like

$$\hat{H} = \bigg(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\bigg) \tag{5.69}$$

it should not be surprising that applying something like $\hat{H}^m$ to a generic function is a rather daunting task that is not particularly likely to end well. This is true, which is why we generally think of expressions involving functions of operators to mostly be meaningful when we apply them to an eigenstate of the operator under consideration. For instance, if $|\psi_n\rangle$ is an energy eigenstate of $\hat{H}$, then we can easily see that

$$\hat{H}^3|\psi_n\rangle = \hat{H}\hat{H}(\hat{H}|\psi_n\rangle) = \hat{H}\hat{H}(E_n|\psi_n\rangle) = \hat{H}(E_n^2|\psi_n\rangle) = E_n^3|\psi_n\rangle, \tag{5.70}$$

where we have used the linearity of our operator to note that the successive multiplications by E_n do not affect how $\hat{H}$ acts on $|\psi_n\rangle$. We can then return to our expression (5.65) and see that it evaluates on an energy eigenstate $|\psi_n\rangle$ as

$$\begin{aligned} e^{-i\hat{H}t/\hbar}|\psi_n\rangle &= \bigg[1 + \big(-i\hat{H}t/\hbar\big) + \frac{1}{2!}\big(-i\hat{H}t/\hbar\big)^2 + \frac{1}{3!}\big(-i\hat{H}t/\hbar\big)^3\bigg]|\psi_n\rangle \\ &= \bigg[1 + \big(-iE_nt/\hbar\big) + \frac{1}{2!}\big(-iE_nt/\hbar\big)^2 + \frac{1}{3!}\big(-iE_nt/\hbar\big)^3\bigg]|\psi_n\rangle \\ &= e^{-iE_nt/\hbar}|\psi_n\rangle, \end{aligned} \tag{5.71}$$

which we can see matches our expressions for the time evolution of an energy eigenstate derived previously. Due to the linearity of all the aspects involved here, we see that this expression also generalizes immediately to a superposition of eigenstates, and hence, is applicable to any function that can be expanded in terms of the eigenstates (which, again, is all states we would be interested in).

More generally, it is common in quantum mechanics to talk of a *time evolution operator* that when applied to a quantum state advances it forward (or backward) in time a given amount. We would typically write this as

$$\hat{U}(t_f, t_i) = e^{-i\hat{H}(t_f - t_i)/\hbar}, \tag{5.72}$$

which when applied to a state $|\psi(t_i)\rangle$ gives the result

$$\hat{U}(t_f, t_i)|\psi(t_i)\rangle = |\psi(t_f)\rangle. \tag{5.73}$$

Again, writing an expression like this is not always meaningful from a numerical calculation perspective, but it can tend to be very useful in establishing general results to guide our understanding. This will become particularly prevalent if you pursue more advanced studies in quantum mechanics, but is something that we will find ourselves not needing to use substantially in this course.

The importance of $\hat{U}(t_f, t_i)$ being a *unitary operator* stems from the probabilistic interpretation of quantum mechanics. So long as our time evolution is defined through unitary operators, we guarantee that the "length" of our state vector $|\psi(t)\rangle$ remains a constant value normalized to one. Hence, we can continue to use our probabilistic interpretation without fear of our state vector becoming un-normalized as it advances in time.

5.4 Hermitian Operators in Quantum Mechanics

One of the most fundamentally important kind of operator in quantum mechanics is a *Hermitian operator*. A Hermitian operator is one that is its own adjoint, which is often referred to as being *self-adjoint*. Mathematically, we have that

$$\hat{A}^\dagger = \hat{A} \tag{5.74}$$

for a Hermitian operator.

5.4.1 Preliminaries

As you may expect, a Hermitian operator is a generalization of a Hermitian matrix, which is an important type of matrix in linear algebra. In matrix form, we see that having $\overline{\mathbf{A}}^\dagger = \overline{\mathbf{A}}$ can only occur if $A_{mn} = A^*_{nm}$. Along the diagonal of our matrix, we also see that $A_{mm} = A^*_{mm}$, which can only be true if the diagonal of the matrix is purely real.

Although these definitions are important to have, we need to have a more general way of determining if an operator is Hermitian without insisting on having some matrix representation of it. Toward this end, the Hermitian property of an operator is most generally defined by viewing its role in the course of evaluating an inner product. In particular, the adjoint of a linear operator $\hat{A}$ is defined as the operator $\hat{A}^\dagger$ for which

$$\langle \hat{A}x, y\rangle = \langle x, \hat{A}^\dagger y\rangle \tag{5.75}$$

for any vectors x and y in the vector space. We have used the mathematical notation for the inner product here with $\langle x, y\rangle = \langle y|x\rangle$ in bra-ket notation. The reason for temporarily using the mathematical notation here is to emphasize that we think of the adjoint $\hat{A}^\dagger$ as being applied to y in the inner product on the right-hand side of (5.75). This point is more subtle to recognize in bra-ket notation (somewhat by design). Returning to the idea of a Hermitian operator, we then see that what we actually need is for

$$\langle \hat{A}x, y\rangle = \langle x, \hat{A}y\rangle \tag{5.76}$$

to hold for any vectors x and y in the vector space.

To see an important result surrounding Hermitian operators in bra-ket notation, it is instructive to consider the quantity $\big(\langle y|\hat{A}|x\rangle\big)^*$. Since this is just a scalar, we can consider this to be the same as $\big(\langle y|\hat{A}|x\rangle\big)^\dagger$. We can now use some manipulations of "matrices" (which also hold for our state vectors and operators) due to the Hermitian transpose to arrive at our desired result. In particular, we find that

$$\big(\langle y|\hat{A}|x\rangle\big)^* = \big(\langle y|\hat{A}|x\rangle\big)^\dagger = \big(\langle y|\big[\hat{A}|x\rangle\big]\big)^\dagger = \big(\hat{A}|x\rangle\big)^\dagger\big(\langle y|\big)^\dagger = \big(|x\rangle\big)^\dagger\big(\hat{A}\big)^\dagger\big(\langle y|\big)^\dagger = \langle x|\hat{A}^\dagger|y\rangle. \tag{5.77}$$

Using the Hermitian property of the operator $\hat{A}$ we then arrive at the final result that

$$\langle x|\hat{A}|y\rangle = \big(\langle y|\hat{A}|x\rangle\big)^* \tag{5.78}$$

for a Hermitian operator $\hat{A}$. This is a particularly useful result for proving certain general properties that Hermitian operators have, which we will turn to shortly.

Before continuing, it is worth mentioning that in some situations it can be useful to think of applying our operators to the "left" in bra-ket notation rather than just acting to the "right". By this, we mean that we can group different parts of an inner product together. For instance, we have that

$$\langle x|\hat{A}|y\rangle = \big(\langle x|\hat{A}\big)|y\rangle = \langle x|\big(\hat{A}|y\rangle\big) \tag{5.79}$$

due to the associativity of the operations involved (namely, matrix multiplication). This can be useful in simplifying certain intermediate calculations before using a specific representation of all the quantities involved to evaluate a final result. As an example, the applicability of this kind of intermediate simplification is not as obvious to determine when working with operators written explicitly in their differential form where we need to be very careful with their ordering, chain rules, etc. to get the correct result.

5.4.2 Important Properties of Hermitian Operators

There are three crucial properties of Hermitian operators that make them extremely useful. In addition to this, it turns out that they play a central role in quantum mechanics, which we will comment on after discussing the important properties of Hermitian operators.

Reality of Eigenvalues

It is quite easy to prove that all eigenvalues of Hermitian operators are *real-valued*. To see this, we assume that we have a normalized eigenvector of the Hermitian operator $\hat{A}$. If the eigenvalue of eigenvector $|\psi_n\rangle$ is λ_n, then we have by definition that

$$\hat{A}|\psi_n\rangle = \lambda_n|\psi_n\rangle. \tag{5.80}$$

We can then take the inner product of this equation with $|\psi_n\rangle$ to get

$$\langle\psi_n|\hat{A}|\psi_n\rangle = \lambda_n\langle\psi_n|\psi_n\rangle = \lambda_n. \tag{5.81}$$

Now, because $\hat{A}$ is Hermitian, we can use the property established in (5.78) to see that

$$\lambda_n = \langle\psi_n|\hat{A}|\psi_n\rangle = \big(\langle\psi_n|\hat{A}|\psi_n\rangle\big)^* = \lambda_n^*. \tag{5.82}$$

The only way for $\lambda_n = \lambda_n^*$ is for λ_n to be purely real, which shows the desired property.

Orthogonality of Eigenvectors

Another easy property to show is that the eigenvectors that correspond to different eigenvalues must be orthogonal. We can start with the trivial expression that

$$\langle\psi_m|\hat{A}|\psi_n\rangle - \langle\psi_m|\hat{A}|\psi_n\rangle = 0. \tag{5.83}$$

The key to the proof is to now use $\hat{A}$ to operate to the "left" in one of these expressions and operate to the "right" in the other. To see this, we explicitly group our operations as

$$\big(\langle\psi_m|\hat{A}\big)|\psi_n\rangle - \langle\psi_m|\big(\hat{A}|\psi_n\rangle\big) = 0. \tag{5.84}$$

To evaluate the first term, we recognize that

$$\big(\langle\psi_m|\hat{A}\big) = \big(\hat{A}^\dagger|\psi_m\rangle\big)^\dagger = \big(\hat{A}|\psi_m\rangle\big)^\dagger = \big(\lambda_m|\psi_m\rangle\big)^\dagger = \lambda_m\langle\psi_m|, \tag{5.85}$$

where we have used the Hermitian property of $\hat{A}$ and that λ_m must be real to simplify the result. Using this result in (5.84), we see that

$$\lambda_m\langle\psi_m|\psi_n\rangle - \langle\psi_m|\big(\hat{A}|\psi_n\rangle\big) = 0. \tag{5.86}$$

We can now simplify the second term easily to arrive at

$$\lambda_m\langle\psi_m|\psi_n\rangle - \lambda_n\langle\psi_m|\psi_n\rangle = (\lambda_m - \lambda_n)\langle\psi_m|\psi_n\rangle = 0. \tag{5.87}$$

Now, because by assumption we have that $\lambda_m \neq \lambda_n$, the only way for this equation to hold is for $\langle\psi_m|\psi_n\rangle = 0$, which is simply a statement of the orthogonality of the two eigenvectors.

It is possible for a Hermitian operator to have multiple eigenvectors with the *same* eigenvalue. This is referred to as *degeneracy*, and invalidates our proof for the orthogonality of the eigenvectors. However, it is still generally possible to form linear combinations of the degenerate eigenvectors to find an orthogonal set in the degenerate space. This can complicate working with them somewhat, but this is not something we will need to consider in this course.

Completeness of Eigenvectors

We will not provide a proof for this property, but it turns out that for the Hermitian operators we will typically be interested in working with in quantum mechanics it can be shown that the eigenvectors of the operator are *complete*. As with our earlier discussions, the primary importance of this result is that we can then use the eigenvectors as a basis set to represent *any* function/state that we would find ourselves needing to work with in practice. As a result, all Hermitian operators we will consider "come" with a basis set that we can use in analyzing particular problems. This can often be a very useful result in interpreting certain quantum mechanical expressions and in computing specific results.

Hermitian Operators in Quantum Mechanics

Clearly, Hermitian operators have many useful properties that would make us want to work with them. Fortunately, it turns out that all physically measurable quantities in quantum mechanics (referred to as *observables*) can be represented by a Hermitian operator. As a result, Hermitian operators play a central role in quantum mechanics.

Up to now, the main Hermitian operator we have worked with is the Hamiltonian operator. However, due to the general properties of Hermitian operators, many of the mathematical techniques we have used to manipulate expressions involving the eigenvectors and eigenvalues of the Hamiltonian operator will be directly applicable to many other operators we will work with. As a result, we have built up a fairly useful set of mathematical tools to begin considering more complicated quantum mechanical calculations.

5.5 Commutation of Operators and the Uncertainty Principle

As alluded to previously, one of the most important concepts in quantum mechanics is related to the *commutation* of two operators. We defined the *commutator* as

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}, \tag{5.88}$$

which helps determine whether two operators commute or not. We say that two operators *commute* if $[\hat{A}, \hat{B}] = 0$. The result of the commutator between two quantum mechanical operators (e.g., if it is 0 or equals something else) plays a very important role in understanding quantum systems, and is also a central component of the *uncertainty principle* that we will discuss shortly.

5.5.1 Commutation of Hermitian Operators

Before considering the uncertainty principle, it is useful to establish that two Hermitian operators commute if and only if they share the same set of eigenfunctions. To prove this, we need to show both "directions" of this statement. To begin, we will show that if Hermitian operators $\hat{A}$ and $\hat{B}$ commute then they have the same eigenfunctions. We will assume that the functions $|\psi_n\rangle$ are eigenfunctions of $\hat{A}$ with eigenvalue A_n. We then have that

$$\hat{A}\hat{B}|\psi_n\rangle = \hat{B}\hat{A}|\psi_n\rangle = \hat{B}A_n|\psi_n\rangle = A_n\hat{B}|\psi_n\rangle. \tag{5.89}$$

We can then equate the first and last expressions and write them in the suggestive form

$$\hat{A}\big(\hat{B}|\psi_n\rangle\big) = A_n\big(\hat{B}|\psi_n\rangle\big). \tag{5.90}$$

For this expression to hold, we must have that the function $\hat{B}|\psi_n\rangle$ is an eigenfunction of $\hat{A}$. If we ignore the case of degenerate eigenvalues (which can be handled, but we won't worry about this), we actually conclude that the function $\hat{B}|\psi_n\rangle$ *must* be proportional to $|\psi_n\rangle$ for

(5.90) to have eigenvalue A_n. Another way to write that $\hat{B}|\psi_n\rangle$ is proportional to $|\psi_n\rangle$ is that

$$\hat{B}|\psi_n\rangle = B_n|\psi_n\rangle, \tag{5.91}$$

where B_n is the proportionality constant. However, we can recognize that (5.91) is also the statement that $|\psi_n\rangle$ is an eigenfunction of $\hat{B}$ with eigenvalue B_n. As a result, we see that it is the case that if $\hat{A}$ and $\hat{B}$ commute then they have the same eigenfunctions.

Let's now show that if two Hermitian operators have the same eigenfunctions they commute. We will assume we have an arbitrary function expanded in terms of the eigenfunctions $|\psi_n\rangle$ as

$$|x\rangle = \sum_n c_n|\psi_n\rangle. \tag{5.92}$$

We now look to show that

$$[\hat{A}, \hat{B}]|x\rangle = 0. \tag{5.93}$$

To begin, we expand the commutator out as

$$\hat{A}\hat{B}|x\rangle - \hat{B}\hat{A}|x\rangle = \sum_n c_n\hat{A}\hat{B}|\psi_n\rangle - \sum_n c_n\hat{B}\hat{A}|\psi_n\rangle. \tag{5.94}$$

If we denote the eigenvalues of $\hat{A}$ and $\hat{B}$ for eigenfunction $|\psi_n\rangle$ as A_n and B_n, respectively, then we can easily see that

$$\begin{aligned}\sum_n c_n\hat{A}\hat{B}|\psi_n\rangle - \sum_n c_n\hat{B}\hat{A}|\psi_n\rangle &= \sum_n c_n\hat{A}B_n|\psi_n\rangle - \sum_n c_n\hat{B}A_n|\psi_n\rangle \\ &= \sum_n c_nA_nB_n|\psi_n\rangle - \sum_n c_nB_nA_n|\psi_n\rangle = 0.\end{aligned} \tag{5.95}$$

As desired, this shows that if two Hermitian operators share the same eigenfunctions they commute.

We are almost in a position to consider a general form of the uncertainty principle. What we will see is that there exists an uncertainty principle relationship between *any* observables associated with operators that *do not commute*. We will comment more on the physical meaning of this after establishing the particular result.

5.5.2 Variance and Standard Deviation

Before considering the uncertainty principle, we need to determine how to compute some additional statistical properties from our operators and wavefunctions. In particular, we will find that the uncertainty principle involves the *variance* or *standard deviation* of the operators involved, so we will need to be able to compute these. Toward this end, we will need recall that we can compute the mean value of an observable A through the expectation value of an operator $\hat{A}$ with the state of our quantum mechanical system. For our current purposes, it will be useful to simplify our notation, so we will write this mean value as

$$\overline{A} = \langle A \rangle = \langle x|\hat{A}|x\rangle, \tag{5.96}$$

where $|x\rangle$ is the state of the system. Part of computing the variance requires "re-centering" the mean value, so we will define a "new" operator $\widehat{\delta A}$ as

$$\widehat{\delta A} = \hat{A} - \overline{A}. \tag{5.97}$$

To continue, the variance will involve the expectation value of the operator $\widehat{\delta A}^2$, which we now turn to computing. To do this, we assume that $|x\rangle$ is expanded on our set of eigenfunctions $|\psi_n\rangle$ for $\hat{A}$ as

$$|x\rangle = \sum_n c_n|\psi_n\rangle. \tag{5.98}$$

We then compute

$$\overline{(\delta A)^2} = \langle x|\big(\hat{A} - \overline{A}\big)^2|x\rangle = \Bigg(\sum_m c_m^*\langle\psi_m|\Bigg)\big(\hat{A} - \overline{A}\big)\Bigg(\sum_n c_n\big(\hat{A} - \overline{A}\big)|\psi_n\rangle\Bigg). \tag{5.99}$$

Using that $|\psi_n\rangle$ has eigenvalue A_n, we can simplify this as

$$\begin{aligned}\overline{(\delta A)^2} &= \Bigg(\sum_m c_m^*\langle\psi_m|\Bigg)\big(\hat{A} - \overline{A}\big)\Bigg(\sum_n c_n\big(A_n - \overline{A}\big)|\psi_n\rangle\Bigg)\\ &= \Bigg(\sum_m c_m^*\langle\psi_m|\Bigg)\Bigg(\sum_n c_n\big(A_n - \overline{A}\big)^2|\psi_n\rangle\Bigg)\end{aligned} \tag{5.100}$$

Since $\big(A_n - \overline{A}\big)^2$ is now just a number, we can further simplify this using the orthonormality of the eigenfunctions to get

$$\overline{(\delta A)^2} = \sum_{m,n} c_m^* c_n\big(A_n - \overline{A}\big)^2\langle\psi_m|\psi_n\rangle = \sum_{m,n} c_m^* c_n\big(A_n - \overline{A}\big)^2\delta_{mn} = \sum_n |c_n|^2\big(A_n - \overline{A}\big)^2. \tag{5.101}$$

Recalling the statistical interpretation of $|c_n|^2$ as the probability the system will return a value of A_n upon measurement, we see that the final result in (5.101) can be expressed in words as being the *expectation value of the squared deviation of the random variable A from its mean $\overline{A}$*. This is, by definition, the variance. In addition to variances, it is typical to work with *standard deviations*. These are simply the square root of the variance; i.e.,

$$\sigma_A = \sqrt{\overline{(\delta A)^2}}. \tag{5.102}$$

Often, it is more convenient to compute the variance directly in terms of our original operators rather than defining a "new" operator like $\widehat{\delta A}$. To determine the more convenient formula, we first note that

$$\overline{(\delta A)^2} = \langle x|\big(\hat{A} - \overline{A}\big)^2|x\rangle = \langle x|\Big[\hat{A}^2 - 2\overline{A}\hat{A} + \overline{A}^2\Big]|x\rangle. \tag{5.103}$$

We can simplify the term

$$\langle x|2\overline{A}\hat{A}|x\rangle = 2\overline{A}\langle x|\hat{A}|x\rangle = 2\overline{A}^2 = \langle x|2\overline{A}^2|x\rangle, \tag{5.104}$$

so that we find

$$\overline{(\delta A)^2} = \langle x|\left[\hat{A}^2 - \overline{A}^2\right]|x\rangle = \langle A^2\rangle - \overline{A}^2. \tag{5.105}$$

It is this final expression that is often used to compute the variance of a quantum mechanical operator.

5.5.3 Uncertainty Principles

In quantum mechanics, it is typical to write the commutator between two Hermitian operators in the general form

$$[\hat{A}, \hat{B}] = i\hat{C}, \tag{5.106}$$

where $\hat{C}$ is sometimes called the *commutation rest.* To establish the uncertainty principle, we will consider two non-commuting Hermitian operators according to (5.106). The proof used is mainly mathematical, but once completed the physical intuition of it is easy to interpret.

To derive the uncertainty principle, we consider for an arbitrary real number α the quantity

$$G(\alpha) = \langle(\alpha\widehat{\delta A} - i\widehat{\delta B})x|(\alpha\widehat{\delta A} - i\widehat{\delta B})x\rangle \geq 0. \tag{5.107}$$

We recognize that this quantity is always greater than 0 because $G(\alpha)$ is simply the inner product of the somewhat "odd" looking vector defined by $(\alpha\widehat{\delta A} - i\widehat{\delta B})|x\rangle$ with itself. Now, to proceed we rearrange our expression for $G(\alpha)$ as

$$G(\alpha) = \langle x|\big(\alpha\widehat{\delta A} - i\widehat{\delta B}\big)^\dagger\big(\alpha\widehat{\delta A} - i\widehat{\delta B}\big)|x\rangle = \langle x|\big(\alpha\widehat{\delta A}^\dagger + i\widehat{\delta B}^\dagger\big)\big(\alpha\widehat{\delta A} - i\widehat{\delta B}\big)|x\rangle. \tag{5.108}$$

Since all our operators here are Hermitian, we can further rewrite this as

$$G(\alpha) = \langle x|\big(\alpha\widehat{\delta A} + i\widehat{\delta B}\big)\big(\alpha\widehat{\delta A} - i\widehat{\delta B}\big)|x\rangle. \tag{5.109}$$

We proceed as

$$\begin{aligned} G(\alpha) &= \langle x|\left[\alpha^2\widehat{\delta A}^2 + \widehat{\delta B}^2 - i\alpha\big(\widehat{\delta A}\,\widehat{\delta B} - \widehat{\delta B}\,\widehat{\delta A}\big)\right]|x\rangle \\ &= \langle x|\left[\alpha^2\widehat{\delta A}^2 + \widehat{\delta B}^2 - i\alpha[\widehat{\delta A}, \widehat{\delta B}]\right]|x\rangle. \end{aligned} \tag{5.110}$$

Recognizing that

$$[\widehat{\delta A}, \widehat{\delta B}] = [\hat{A} - \overline{A}, \hat{B} - \overline{B}] = [\hat{A}, \hat{B}] = i\hat{C} \tag{5.111}$$

because the scalar numbers $\overline{A}$ and $\overline{B}$ commute, we find that

$$G(\alpha) = \langle x| \left[\alpha^2 \widehat{\delta A}^2 + \widehat{\delta B}^2 + \alpha \hat{C} \right] |x\rangle = \alpha^2 \sigma_A^2 + \sigma_B^2 + \alpha \overline{C}. \tag{5.112}$$

We can rearrange this in a non-obvious way to get

$$G(\alpha) = \sigma_A^2 \left(\alpha + \frac{\overline{C}}{2\sigma_A^2} \right)^2 + \sigma_B^2 - \frac{\overline{C}^2}{4\sigma_A^2} \geq 0. \tag{5.113}$$

Now, because this must hold for any α we can choose α in a way to simplify this expression. The sensible choice is to make α eliminate the first term in our earlier expression, which requires

$$\alpha = -\frac{\overline{C}}{2\sigma_A^2}. \tag{5.114}$$

Choosing α in this way, then rearranging and taking the square root gives us

$$\sigma_A \sigma_B \geq \frac{1}{2} |\overline{C}|. \tag{5.115}$$

This is the *generalized uncertainty principle*, which holds for *any* set of non-commuting quantum mechanical operators associated with two measurable quantities.

This is a rather profound result that has many consequences in quantum mechanics. It expresses the *relative minimum size* of the *distributions* of two measurable quantities if we were to perform *many measurements on identically prepared systems*. One important consequence of this result is that it is *impossible* for us to simultaneously know with arbitrary precision both quantities. To get a better feel for this, it is helpful to consider a few of the most “famous” uncertainty principles.

Heisenberg (Position-Momentum) Uncertainty Principle

The *Heisenberg (or position-momentum) uncertainty principle* expresses the relationship between the standard deviations of position and momentum. To determine this uncertainty principle, we need to evaluate $\overline{C}$. We are already acquainted with the momentum operator as

$$\hat{p} = -i\hbar \frac{d}{dx}, \tag{5.116}$$

but we will also need the position operator. It turns out that within this *coordinate space representation*, the position operator $\hat{x}$ is simply the *function* x. Hence, we need to evaluate $[\hat{p}, \hat{x}]$. Since this involves differential operators, it is easiest to think of this being applied to an arbitrary function $|f\rangle$. Then, we have that

$$[\hat{p}, \hat{x}]|f\rangle = -i\hbar \left(\frac{d}{dx} x - x \frac{d}{dx} \right) |f\rangle = -i\hbar \left(\frac{d}{dx} \big(x|f\rangle \big) - x \frac{d}{dx} |f\rangle \right). \tag{5.117}$$

It is essential to recognize that the first term in this last expression needs to be evaluated using the *product rule* from calculus. This then gives us

$$[\hat{p}, \hat{x}]|f\rangle = -i\hbar\bigg(|f\rangle + x\frac{d}{dx}|f\rangle - x\frac{d}{dx}|f\rangle\bigg) = -i\hbar|f\rangle. \tag{5.118}$$

Since $|f\rangle$ is arbitrary, we see that we can write in general that

$$[\hat{p}, \hat{x}] = -i\hbar. \tag{5.119}$$

The commutation rest operator in this case is then just $\hat{C} = -\hbar\hat{I}$, so that the expectation value is $\overline{C} = -\hbar$. Finally, we have the Heisenberg uncertainty principle that

$$\sigma_p \sigma_x \geq \frac{\hbar}{2}. \tag{5.120}$$

This tells us that it is impossible to known both the position and momentum of a particle simultaneously. For example, if we were to measure the position of our particle, then the uncertainty in the momentum would necessarily become very large. This result can actually be readily understood in terms of Fourier theory. If we have a very localized function in one space (e.g., a delta function), then we have a very broad function in our Fourier space. Similarly, we find that if we measure something and collapse it into an eigenfunction of one operator we will need to use a large number of eigenfunctions (generally, all of them) from our other non-commuting operator to express the corresponding state of the system.

Energy-Time Uncertainty Principle

There is also an almost identical uncertainty principle between energy and time. We already know that the operator corresponding to energy is the Hamiltonian operator. Similar to the function x being the position operator $\hat{x}$, we also have the time operator $\hat{t}$ just being the function t. Then, for an arbitrary quantum state, one can find that

$$[\hat{H}, \hat{t}] = i\hbar. \tag{5.121}$$

As a result, there is an energy-time uncertainty principle of

$$\sigma_E \sigma_t \geq \frac{\hbar}{2}. \tag{5.122}$$

This can also be recast in the form of a frequency-time uncertainty principle by noting $E = \hbar\omega$. We then have

$$\sigma_\omega \sigma_t \geq \frac{1}{2}, \tag{5.123}$$

which is a relation that is common in traditional Fourier analysis.

5.6 Practice Problems

1. Problem 4.1.1 from D. A. B. Miller.

2. Problem 4.2.1 from D. A. B. Miller.

3. Problem 4.10.3 from D. A. B. Miller.

4. Problem 4.11.4 from D. A. B. Miller.

5. There are two quantum-mechanically measurable quantities A and B associated with quantum operators $\hat{A}$ and $\hat{B}$, respectively. The possible eigenvalues of $\hat{A}$ are A_1 and A_2 associated with eigenvectors $|\psi_1\rangle$ and $|\psi_2\rangle$, while the possible eigenvalues of $\hat{B}$ are B_1 and B_2 associated with eigenvectors $|\phi_1\rangle$ and $|\phi_2\rangle$. Further, we know that the relation between these eigenvectors is

$$|\phi_1\rangle = \frac{1}{5}\big(3|\psi_1\rangle + i4|\psi_2\rangle\big), \tag{5.124}$$

$$|\phi_2\rangle = -\frac{1}{5}\big(i4|\psi_1\rangle + 3|\psi_2\rangle\big). \tag{5.125}$$

   We will now consider a sequence of measurements that are performed in quick succession such that you do not need to consider any free time evolution of the system according to the time-dependent Schrödinger equation occurs in between the measurements.

   (a) A measurement of quantity B is performed and the result is B_2. What is the state of the system immediately after this measurement?

   (b) We now immediately measure the quantity B again. What will the result be?

   (c) If we were now to make a measurement of quantity A, what is the probability the result will be A_1 and what is the probability the result will be A_2?

   (d) We now perform a measurement of quantity A and find that the result is A_1. If we were to now make a measurement of quantity B, what is the probability the result will be B_1 and what is the probability the result will be B_2?

6. You have been given a Hermitian operator $\hat{A}$ whose bilinear expansion is given by

$$\hat{A} = \sum_n A_n |\alpha_n\rangle\langle\alpha_n|, \tag{5.126}$$

   where A_n is the eigenvalue associated with the nth-eigenvector $|\alpha_n\rangle$ of a complete orthonormal basis. Now, find the inverse operator to $\hat{A}$ such that $\hat{A}^{-1}\hat{A} = \hat{I}$. By correctly manipulating the bra-ket notation of $\hat{A}^{-1}$ and $\hat{A}$, also explicitly show that $\hat{A}^{-1}\hat{A} = \hat{I}$.

7. In this problem, we will consider the manipulation of terms that occur in the Hamiltonian of many superconducting qubits. These systems involve a "charge operator" $\hat{n}$ (that we will learn about more later in the course) that has a complete set of orthonormal eigenstates typically denoted by $|N\rangle$ in Dirac bra-ket notation where N is a positive or negative integer. The action of the charge operator on these states follows the simple rule that

$$\hat{n}|N\rangle = N|N\rangle. \tag{5.127}$$

Now, using two "resolution of the identity operator" written in terms of charge states, show how to rewrite the left-hand side of (5.128) to become the right-hand side:

$$4E_C\big(\hat{n} - n_g\big)^2 \rightarrow 4E_C \sum_N \big(N - n_g\big)^2 |N\rangle\langle N|, \tag{5.128}$$

where E_C and n_g are real-valued constants.

Hint: To show this, you will need to put one identity operator before $\big(\hat{n} - n_g\big)^2$ and one identity operator after this. Don't forget that when you introduce these resolution of the identity operators that the dummy indices for the summations ***must*** be different!

8. Quantum logic gates that operate on a single quantum bit (qubit) are defined in a two-dimensional Hilbert space. Three canonical single qubit gates are the Pauli-X, Y, and Z gates (these are also called *Pauli spin matrices* for reasons we will discuss later). These are denoted as

$$\hat{X} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \hat{Y} = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \hat{Z} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{5.129}$$

   (a) Find the commutation relations between each pair of operators through explicit matrix multiplication. You should be able to simplify your answer so that the result is some scalar times one of the Pauli matrices (i.e., $\hat{X}$, $\hat{Y}$, or $\hat{Z}$).

   (b) Using the results of (a), construct an uncertainty principle between $\hat{Y}$ and $\hat{Z}$.

Chapter 6

Basics of Quantum Computing

6.1 Introduction to Quantum Computing

We will now begin to discuss the operation of a quantum computer at a high level. Quantum computing is one of the most intriguing technological applications of quantum information processing, and has driven much of the interest in pushing the boundaries of implementing next-generation quantum technologies. At a high level, a quantum computer is a machine designed to modify the quantum state of a collection of qubits to achieve some computational task. The interest in using a quantum computer arises from the fact that there are certain computational tasks of great importance that are known to become essentially impossible to solve once the number of variables involved gets large enough. Simple examples include many kinds of optimization problems, as well as simulating quantum systems.

We will begin our discussion by considering how to mathematically express the state of a quantum computer before then reviewing one of the simpler quantum algorithms that shows a speedup compared to classical algorithms. Later, we will go through a more detailed introduction to the different parts of a quantum computer from the physical layer through to more abstracted interfaces. We will then continue through much of the remaining part of the course by learning in more detail about the various physical components that make up a quantum computer to have a more fundamental understanding of what goes into building and controlling these fascinating new technologies.

To begin, we will first note that the most common form of quantum computer involves a number of interconnected *quantum bits (qubits)*. As its name suggests, a qubit is abstractly defined as any quantum system that either exactly or approximately can be considered to consist of two states. It is conventional to then consider these two states to be represented as $|0\rangle$ or $|1\rangle$. One of the defining features of a qubit is that the complete state of it can then be considered to be given by a *superposition* of these two states, such as

$$|\psi\rangle = c_0|0\rangle + c_1|1\rangle, \tag{6.1}$$

where c_0 and c_1 are complex-valued coefficients subject to the usual normalization constraint so that they can be interpreted in a probabilistic manner.

Of course, a quantum computer with only one qubit is hardly of interest, so we also need to be able to describe the state of a more complex system such as a quantum computer composed of many qubits. The mathematical "tool" that is used to achieve this is known as a *tensor product*. Tensors are a generalization of matrices to be able to extend to higher dimensional sets of data. We can think of a tensor as having a number of indices that can run over the entire range of the Hilbert space that makes up each dimension of the tensor. If we initially consider the case of a system with only two qubits, we would have the Hilbert spaces of qubits 1 and 2, which we denote as $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively. The total Hilbert space that our system is analyzed in then becomes the tensor product of these two spaces, denoted as

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2. \tag{6.2}$$

A state within the overall Hilbert space then becomes the tensor product of states from the individual Hilbert spaces. As an example, if we have qubit 1 in its 0 state and qubit 2 in its 1 state, we could write the state as $|0\rangle_1 \otimes |1\rangle_2$. This notation becomes rather cumbersome, especially if we have many coupled qubits, and so simpler notations are often adopted. For instance, we could write this same state equivalently as

$$|0\rangle_1 \otimes |1\rangle_2 = |0\rangle_1 |1\rangle_2 = |0\rangle|1\rangle = |0, 1\rangle = |01\rangle. \tag{6.3}$$

All of these notations are entirely equivalent, it is just a matter of preference for picking between one or another. Another point to note is that because each set of qubit states represents an orthonormal basis for that particular qubit, our tensor product of these different states can be used as an orthonormal basis in the larger Hilbert space $\mathcal{H}$. One important consequence of this is that all of the states are mutually orthonormal, i.e.,

$$\langle m', n' | m, n\rangle = \delta_{mm'}\delta_{nn'}. \tag{6.4}$$

It should also be emphasized, that each state can exist in a quantum superposition with the same rules as before applied in regards to normalization of coefficients and statistical interpretation. For example, we could have the state

$$|\psi\rangle = \frac{1}{\sqrt{2}}|01\rangle - \frac{1}{\sqrt{2}}|00\rangle \tag{6.5}$$

as a valid quantum state. More generally, we can have

$$|\psi\rangle = \sum_{m,n} c_{mn} |mn\rangle, \quad \sum_{m,n} |c_{mn}|^2 = 1. \tag{6.6}$$

More broadly, we can quickly find that the size of the Hilbert space spanned by N qubits coupled to each other will be given by 2^N. Amazingly, if we had a quantum computer with only 300 qubits, the size of this state space would already be so large that there are not enough atoms in the known universe to be able to fully represent all of the possible states! One of the key aspects of a quantum computer is the idea of *quantum parallelism*, i.e., we can utilize a superposition of these enormous quantum states to in some sense "query" a problem a huge number of times simultaneously in a manner that is not possible with classical computing. We can also utilize *quantum entanglement* to process information in a manner distinct from any classical system, although we will have to defer our discussion on quantum entanglement until later in the course. The goal of quantum computers is to utilize these effects in novel ways to try and approach the solution of important computational tasks that are known to be nearly impossible for classical computers.

As such, the design of these quantum algorithms can be very challenging because it requires us to think of ways to utilize non-intuitive quantum effects in a manner to solve a computational problem. The difficulty is further enhanced because we are only interested in these quantum algorithms if they solve a particularly important problem in a manner that is better than *any* known classical algorithm, otherwise the effort in using a quantum computer would not be worth it. Hence, the bar is quite high to achieve success with a quantum algorithm!

Despite this, the interest in quantum computers has been steadily growing due to mounting evidence that they can be implemented experimentally and theoretical results demonstrating their performance improvement over classical computers for certain problems. One problem that significantly ignited interest was the proof that a quantum computer could provide an exponential speedup for finding the prime factors of an integer (this is *Shor's algorithm*). The assumed difficulty of this problem for very large numbers forms the backbone of many public-key cryptography schemes that underlie much of modern secure data transmission. As a result, the development of a quantum computer could significantly undermine large parts of how society currently functions, making many governments and companies very motivated to stay on top of the development of these revolutionary technologies.

6.2 Grover's Search Algorithm

We will now look at one of the simpler quantum algorithms to get a taste at how the effects of quantum mechanics can be leveraged to solve a particular computational task. In particular, we will discuss *Grover's search algorithm*, which is often described as being useful to locate an item in an unstructured database. However, Grover's search algorithm has application to a rather broad set of problems that require "searching" for something of interest. In many cases, a classical algorithm that must perform this "searching" will usually require $O(N)$ operations if there are N possible items to search through. Grover's search algorithm provides a *quadratic speedup* by being able to accomplish this task in $O(\sqrt{N})$ operations. Although this speedup is not as exciting as the exponential speedup of other quantum algorithms (e.g., Shor's algorithm), it still can be substantial when N is large. As a simple example to emphasize this point, an algorithm for solving a linear system of equations using a single processor that requires $O(N^2)$ operations will take over 100 days to perform this task for $N = 10^6$. Meanwhile, an algorithm that perform this in $O(N)$ operations will reduce this computational time to less than 1 hour. Hence, when the number N becomes large even a quadratic speedup can be exceptionally important in making a particular computational task practical!

6.2.1 Basic Process

The basic structure of Grover's search algorithm is shown in Fig. 6.1. The algorithm works with a set of n qubits that we use to index into our "search space" and another set of qubits that we use for auxiliary computations in an *oracle function.* The oracle function is where we embed information about the actual search problem we are attempting to solve, and we will discuss it in more detail later. The number of qubits we need depends on the size of the space we are searching and the specifics of the oracle function. In general, we will assume that there are N elements in the search space so that we can choose the number of qubits we need such that $N < 2^n$. For simplicity, we will assume that $N = 2^n$.

To begin the algorithm, we first use a set of Hadamard gates to place all of our n qubits into an equal superposition state. More specifically, a single-qubit Hadamard gate has as its matrix representation

$$H = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \tag{6.7}$$

If we apply this to a qubit beginning in its ground state we will have

$$H|0\rangle = \frac{1}{\sqrt{2}}\big(|0\rangle + |1\rangle\big). \tag{6.8}$$

If we were to instead perform a Hadamard gate to *all* of our n qubits starting in the ground state, then we would have

$$H^{\otimes n}\big(|0\rangle_1 \otimes |0\rangle_2 \ldots \otimes |0\rangle_n\big) = \frac{1}{\sqrt{2^n}}\big(|0\rangle_1 + |1\rangle_1\big) \otimes \big(|0\rangle_2 + |1\rangle_2\big) \ldots \otimes \big(|0\rangle_n + |1\rangle_n\big). \tag{6.9}$$

By explicitly expanding all the terms in this tensor product, one can find that it results in an equal superposition state of all are n qubits, as desired. As a shorthand, we will denote this superposition by

$$|\psi\rangle = \frac{1}{\sqrt{N}}\sum_{x=0}^{N-1}|x\rangle, \tag{6.10}$$

where $|x\rangle$ is a number corresponding to a particular bit string that indexes into our search space (also, recall that $N = 2^n$). The purpose of this superposition state is to exploit the quantum parallelism that is possible with quantum states. Next, we input our quantum state into a sequence of *Grover iterations.* Each Grover iteration is meant to modify the quantum state of our system so that it is closer to the correct answer. We repeat this Grover iteration a sufficient number of times so that we can maximize the statistics of our output state matching the solution to our search problem. Clearly, a lot is going on inside this Grover iteration, which we now turn to considering in more detail.

6.2.2 Grover Iteration

The layout of each Grover iteration is shown in Fig. 6.2. Conceptually, it performs two operations. The first involves evaluating the oracle function, with the second step involving

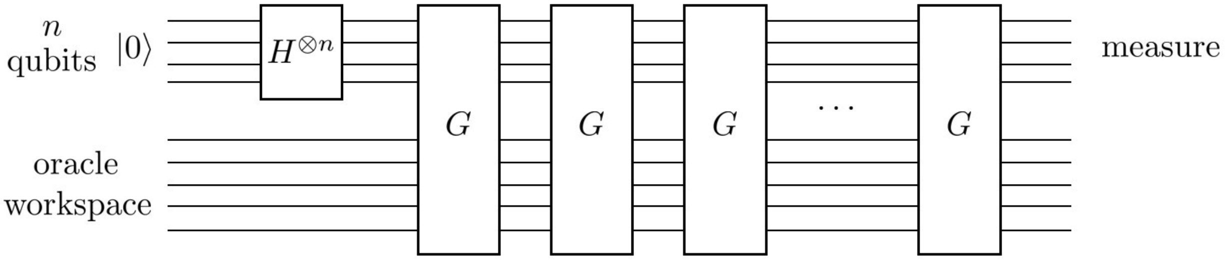


Figure 6.1: Circuit schematic of Grover's search algorithm. Each G block corresponds to a Grover iteration.

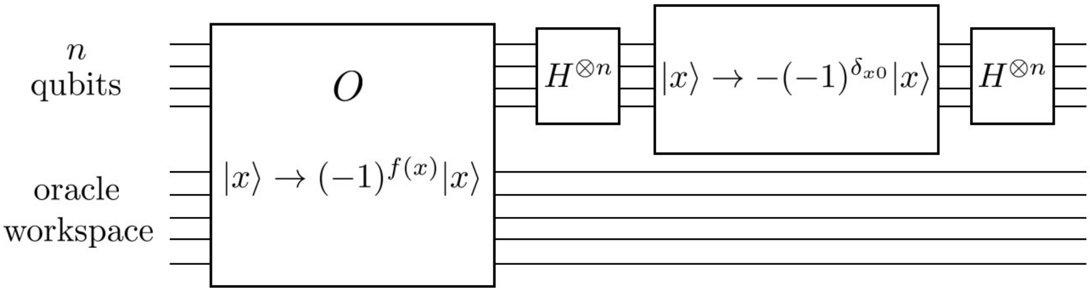


Figure 6.2: Circuit schematic of a single Grover iteration.

the three remaining gates. Hence, we now need to consider a more specific definition of what the oracle function actually does.

At a high level, the oracle function is meant to be an abstract concept to allow us to analyze quantum circuits in a generic sense without getting too "hung up" on the details. Unfortunately, this often leads to the operation of the oracle seeming particularly mysterious, which is certainly not helped by its name or the abstract language often used to describe it. In essence, the purpose of the oracle is to be a black box that performs a specific target operation on the input state. The design of the oracle function depends on the specific details of the search problem being considered, and as a result is not able to be defined explicitly for a general analysis. Overall, the oracle is designed to *recognize* whether something is a solution to the search problem and to then *mark* this state in some way.

It should be emphasized that *recognizing* whether something is a solution to the search problem does not require actually *computing* the solution, or even *knowing* the solution *a priori*. In many cases, we can utilize classical logic statements to test whether a particular item matches the "pattern" we are looking for. The oracle function encodes this logic into our quantum circuit.

It is typical to consider this operation in terms of some function $f(x)$ that takes as input an integer x that indexes into our search space and returns a value of $f(x) = 1$ if x is a solution and $f(x) = 0$ if x is not a solution to the search problem. To "mark" the result of the oracle function, we utilize an *oracle qubit* where the evaluation of $f(x)$ will be stored.

This is typically done in the form

$$O|x\rangle|q\rangle = |x\rangle|q \oplus f(x)\rangle, \tag{6.11}$$

where O is a unitary operator that implements the oracle function, $|q\rangle$ is the oracle qubit, and $q \oplus f(x)$ denotes addition modulo 2. We see that if $|x\rangle$ is a solution to the search problem, then the oracle function will flip the value of the oracle qubit.

For Grover's search algorithm, we actually prepare the oracle qubit in the superposition state

$$|q\rangle = \frac{1}{\sqrt{2}}\big(|0\rangle - |1\rangle\big). \tag{6.12}$$

The result of this is that if the trial state $|x\rangle$ is not a solution to the search problem then the oracle qubit will remain unchanged, but if the trial state $|x\rangle$ is a solution to the search problem then we get that

$$|q\rangle = \frac{1}{\sqrt{2}}\big(|0\rangle - |1\rangle\big) \to \frac{1}{\sqrt{2}}\big(|1\rangle - |0\rangle\big) = -|q\rangle. \tag{6.13}$$

Hence, we can write the overall action of the oracle function on our quantum state as

$$O|x\rangle\left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right) = (-1)^{f(x)}|x\rangle\left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right). \tag{6.14}$$

Since the oracle qubit state does not change (we just get the sign flip on the rest of the state), we can "drop" it when writing our expressions to simplify things down a bit. However, it is important to emphasize that it is still there and is essential in achieving the end goal of the Grover search algorithm by *marking* the states that are solutions to the search problem by flipping their sign.

Now, the remaining three gates in each Grover iteration are lumped together to achieve the other important step of the algorithm. We have already encountered the Hadamard gates, but the middle operation of a *conditional phase shift* is new to us. The idea is that we should flip the sign of every state in the computational basis *except* for the $|0\rangle$ state. We can write this operation mathematically as

$$|x\rangle \to -(-1)^{\delta_{x0}}|x\rangle, \tag{6.15}$$

where δ_{x0} is a Kronecker delta. The unitary operator that achieves this conditional phase shift is

$$2|0\rangle\langle 0| - I. \tag{6.16}$$

Combined with the two Hadamard gates we get that the remaining step of the Grover iteration implements

$$H^{\otimes n}\big(2|0\rangle\langle 0| - I\big)H^{\otimes n} = 2|\psi\rangle\langle\psi| - I. \tag{6.17}$$

From this, we see that an entire Grover iteration is equivalent to $G = \big(2|\psi\rangle\langle\psi| - I)O$. To see *why* this accomplishes our goal of finding the solution to the search problem, it is useful to look at a geometric visualization of what is happening.

6.2.3 Geometric Visualization

At this point, the purpose of the oracle function and Grover iteration are likely still rather perplexing. Fortunately, it turns out there is a very simple and intuitive geometric explanation for what these two operations are performing. To see this, it helps to think of our state as being decomposed into an equal superposition of states that are solutions to the search problem and an equal superposition of states that are *not* solutions to the search problem. If we have M valid solutions, then we can denote these two states as

$$|\alpha\rangle = \frac{1}{\sqrt{N-M}} \sum_{x\in\mathcal{N}} |x\rangle, \tag{6.18}$$

$$|\beta\rangle = \frac{1}{\sqrt{M}} \sum_{x\in\mathcal{S}} |x\rangle, \tag{6.19}$$

where $\mathcal{S}$ ($\mathcal{N}$) denotes the set of states that are (not) solutions to the search problem. The states $|\alpha\rangle$ and $|\beta\rangle$ are easily seen to be orthogonal to each other, and so we can think of using them as a basis to describe our initial state $|\psi\rangle$ as

$$|\psi\rangle = \sqrt{\frac{N-M}{N}}\,|\alpha\rangle + \sqrt{\frac{M}{N}}\,|\beta\rangle. \tag{6.20}$$

We can now introduce the geometric idea of what the different steps of the Grover iteration accomplish. To start, we recognize that the operation of the oracle function on a state described in terms of $|\alpha\rangle$ and $|\beta\rangle$ is

$$O\big(a|\alpha\rangle + b|\beta\rangle\big) = a|\alpha\rangle - b|\beta\rangle. \tag{6.21}$$

If we consider $|\alpha\rangle$ and $|\beta\rangle$ to define two orthogonal axes in a plane we can recognize that the operation of O is to *reflect* the state vector about the $|\alpha\rangle$ axis. The remaining part of the Grover iteration can also be interpreted as a reflection, but this time about the axis defined by the vector $|\psi\rangle$. The product of these two reflections is a *rotation* of the state vector in the direction of $|\beta\rangle$ – i.e., the equal superposition of all solutions to the search problem. By performing more and more Grover iterations, we keep rotating our state vector closer and closer to the correct answer $|\beta\rangle$. This process is illustrated for a single Grover iteration in Fig. 6.3.

6.2.4 Summary

Now that we've discussed the Grover search algorithm fully, let's reflect briefly on how it accomplishes its processing goal in a manner that is not possible with classical computers. The first and obvious point is that it relies on us beginning with a quantum state that is an equal superposition of *all* possible indexes in our search space. We then process this superposition state in a coherent manner to exploit the quantum parallelism that is possible with this kind of state. The second important point was that we were able to modify the phase of specific portions of our state space while operating on the coherent superposition.

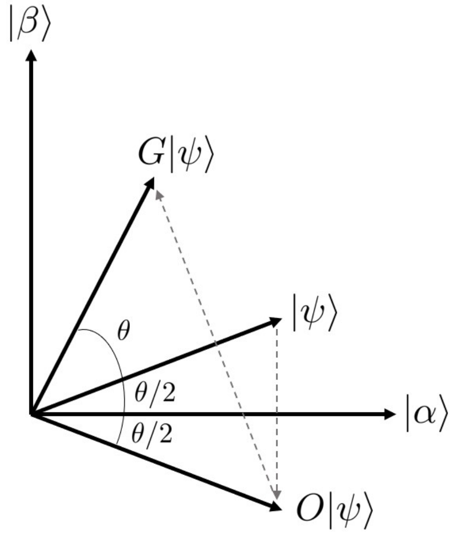


Figure 6.3: Geometric visualization of a single Grover iteration. The two reflections involved lead to a rotation of the state vector in the direction of the equal superposition of solutions to the search problem.

Once these portions had been “marked”, we could design a set of operations to *amplify* the probability amplitude of these states. Since our coefficients must remain normalized, this *amplification* is accompanied by reducing the portion of our probability amplitude on the states that were not marked by our algorithm. By repeating this process enough times, we will arrive at a quantum state that approximately only contains appreciable probability amplitudes for states that solve our search problem. We can then measure the output state of our computer and be relatively certain that it is a solution to our problem. We can always perform a quick check on our output state to verify that it is indeed one of the desired solutions. We can then repeat this process as many times as we wish to find additional solutions to the search problem (assuming they exist). Of course, we may also remeasure a solution we have already found, but given that all solutions have an equal probability of being measured it should not take us too long to find all the solutions to our search problem.

6.3 Building Blocks of a Quantum Computer

We will discuss the building blocks of a quantum computer from a bottom-up approach. We will not go into depth on each topic, but will give a very broad overview of some of the important concepts.

6.3.1 Physical Layer

As with any computer, the foundation of a quantum computer is its physical layer. As we have mentioned previously, a quantum computer is composed of a number of qubits that we need to be able to coherently control. The core capabilities needed of our physical system have been encapsulated in *DiVincenzo's criteria* [14]. These are an agreed upon set of characteristics needed of a physical system to have it be able to perform as a quantum computer. They are:

1. A scalable physical system with well-characterized qubits.
2. The ability to initialize the state of the qubits to a simple fiducial state.
3. Long relevant decoherence times.
4. A "universal" set of quantum gates.
5. A qubit-specific measurement capability.

The discussion we will have in a few weeks on the physics of superconducting circuit artificial atoms will be a good example of the kind of details that need to be understood for a physical system to be a "well-characterized qubit". The notion of scalability encapsulates the need for us to combine many of these qubits without fundamentally changing their properties. For instance, if each time we added a new qubit it reduced our ability to measure the state of all other qubits then we may have a system that is *not scalable.*

Next, it should not be surprising that for a particular quantum algorithm to run in a desirable manner we will generally need to have it begin from a known "starting point". If the initial state of all the qubits are not controllable, it becomes impossible to achieve this desired property. In many cases, the simple fiducial state that will be used to initialize the qubits will be the ground state. This is usually the most accessible option because the ground state is often the equilibrium state for the qubit that it "wants" to return to (e.g., through spontaneous emission of an excitation).

In order to successfully run a quantum algorithm, it is vital that uncontrollable errors do not ruin the coherence of our quantum state. We will discuss more details about this later in the course, but at a high level uncontrollable perturbations to our qubits in the form of various environmental fluctuations can over time cause the "phase coherence" of our quantum states (i.e., the relative phases involved in the superposition state) to be lost. This makes our superposition states "fall apart", leading to our possible quantum parallelism breaking down. Hence, we need to make sure that the *decoherence times* of our qubits are long enough that we can successfully complete our desired quantum algorithm before the phase coherence is lost.

Our physical system also must be able to be controlled in such a way that we can have a universal set of quantum gates. This is similar to the idea of a universal gate set for classical computing, where it is well known that we can synthesize any possible logical operation using only a subset of all possible logic gates. For instance, clever combinations of NAND gates can be used to realize any other logic gate (e.g., an OR gate or a NOT gate, etc.). In the quantum case, the number of quantum logic gates that can be performed is in one sense broader than the classical case, but also more restrictive in another sense. The possible quantum gates can be viewed as broader because we can define "partial" operations of gates that leave a quantum state in various superpositions. However, any quantum gate also must follow certain quantum mechanical principles, which restricts the types of gates that are possible. In particular, all quantum gates must be defined by *unitary operators*, which

Name	Circuit Symbol	Matrix	Name	Circuit Symbol	Matrix
Pauli-**X**	$-\boxed{\mathbf{X}}-$	$\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$	Hadamard	$-\boxed{\mathbf{H}}-$	$\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}$
Pauli-**Y**	$-\boxed{\mathbf{Y}}-$	$\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix}$	S-Gate	$-\boxed{\mathbf{S}}-$	$\begin{bmatrix}1 & 0\\ 0 & i\end{bmatrix}$
Pauli-**Z**	$-\boxed{\mathbf{Z}}-$	$\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}$	T-Gate	$-\boxed{\mathbf{T}}-$	$\begin{bmatrix}1 & 0\\ 0 & e^{i\pi/4}\end{bmatrix}$

Figure 6.4: Examples of the most commonly used single qubit gates. The circuit symbol is how these gates are denoted schematically when one draws out a quantum algorithm in terms of a sequence of gates. The matrix representation helps us understand the actual operation of the gate and prove certain correspondences between various gates (e.g., how to synthesize one gate through the use of other gates). Visualizing the action of these gates through a geometric depiction on the *Bloch sphere* is also very useful. Examples of this can be found in [15].

in turn correspond to unitary matrices in a matrix representation. This has a number of consequences, including that all quantum gates must be *reversible*. This is not the case with classical logic gates, such as an AND gate, where it is not possible to reconstruct what the inputs were given the output state. However, it is still possible to synthesize any "standard" logic gate we are familiar with using only reversible logic gates (a good example is the *Fredkin gate*). Due to this, it is possible to use only reversible gates and still implement any logical operation, although the reversible implementation may have some additional overhead associated with this.

Importantly, despite the many possible quantum gates it is still possible to reconstruct any possible quantum gate to sufficient fidelity using a smaller subset (i.e., a finite number) of quantum gates. This often involves needing to produce relatively arbitrary single qubit operations combined with some *controlled multi-qubit gate*. These controlled gates are a kind of conditional operation that only applies a particular function depending on the value of a control qubit. One particularly popular gate is a CNOT (controlled NOT) gate that only performs the NOT operation if the control qubit is in the $|1\rangle$ state. These controlled gates help facilitate entanglement between different qubits, and hence, play a vital role in quantum computation. Examples of many of the most commonly used single and multiple qubit gates are shown in Figs. 6.4 and 6.5, respectively.

Finally, we also must be able to measure the states of the qubits in our quantum computer. This certainly must be done at the end of an algorithm to get our result out of the computer, but can also be a handy tool at an intermediate state of a computation. This often occurs when we wish to take the result of the intermediate measurement, process it with a *classical* computer, and then use this result to condition further operations of the quantum algorithm. Although these classical computations can actually be done on a quantum computer (any classical computation *can* be done with a quantum computer), it may not be the most efficient

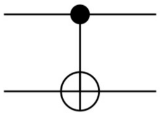
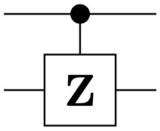
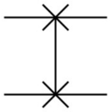
Name	**Circuit Symbol**	**Matrix**
Controlled-Not (CNOT)		$\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&0&1\\0&0&1&0\end{bmatrix}\begin{bmatrix}\vert 00\rangle\\\vert 01\rangle\\\vert 10\rangle\\\vert 11\rangle\end{bmatrix}$
Controlled-**Z** (CZ)		$\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&-1\end{bmatrix}\begin{bmatrix}\vert 00\rangle\\\vert 01\rangle\\\vert 10\rangle\\\vert 11\rangle\end{bmatrix}$
SWAP		$\begin{bmatrix}1&0&0&0\\0&0&1&0\\0&1&0&0\\0&0&0&1\end{bmatrix}\begin{bmatrix}\vert 00\rangle\\\vert 01\rangle\\\vert 10\rangle\\\vert 11\rangle\end{bmatrix}$

Figure 6.5: Examples of some of the most commonly used two qubit gates. These gates are important for generating entangled states.

way to do things. Hence, interfacing with a classical computer can still be a valuable resource. It should also be noted that these intermediate measurements can certainly be needed for the running of a particular quantum algorithm, but also form a key part of many quantum error correction protocols as well. Hence, the intermediate measurement of qubits in the process of an algorithm is an extremely important capability.

6.3.2 Compiler

The next layer up from the physical layer is the compiler. The purpose of the compiler is to take a quantum algorithm defined in terms of an arbitrary set of gates and convert this into an equivalent set of gates that can actually be implemented on the hardware platform being used. A good compiler will be aware of certain details about the performance of the underlying hardware so that it can attempt to choose equivalent gates that perform better on the specific hardware being used. It may also insert additional operations in the algorithm to try and stave off decoherence effects in idling qubits (i.e., one's that aren't actively being used at a given step of the algorithm). The compiler also must be aware of the interconnectivity of the different qubits in the quantum processor so that it can "move" the particular state of qubits around the processor to facilitate different multi-qubit gates. These "movements" themselves need to be accomplished via application of particular gates, which the compiler handles so that the algorithm designer does not need to think about these details. The compiler can also be viewed as converting all of these gates into a sequence of actual physical biases that should be applied to the qubits in the quantum computer.

Current quantum computers are limited in size, but as they become bigger the compilers will also need to help facilitate the use of quantum memories. In these situations, it is anticipated that quantum computing architectures will begin to resemble those of classical computers more and more. This includes having structures like RAM, optimized registers for facilitating different operations, etc.

6.3.3 Higher-Level Programming

The next layer above the compiler would be a quantum programming language. This would be something that is written on a classical computer and defines the steps of a quantum algorithm. The higher-level programming language can facilitate additional abstraction through user-defined "functions" that would be repeated in common algorithms. The higher-level language can also be used to interface with classical machines to support parts of a quantum algorithm. Due to how "far" away these higher-level languages are from the actual quantum hardware, they are able to resemble classical programming languages to a significant degree.

6.3.4 Additional Comments

This has, of course, been a very abbreviated discussion about the different parts of a quantum computer. As the field continues to evolve, it is anticipated that more and more sophisticated quantum computers will be built. This will lead to the kinds of abstraction layers that are familiar from classical computer engineering. However, current devices are still quite primitive in this grand scheme, and so some knowledge of all of the different layers of the machine are important. The current era of quantum computers are often referred to as *noisy intermediate-scale quantum (NISQ) devices*. This is meant to emphasize that current machines are still fraught with errors due to decoherence effects and only contain a relatively small number of qubits compared to what is needed for practical applications. For instance, the first claim of a "quantum advantage" was made with a machine that only contained 53 qubits [16], but current estimates for the number of qubits needed to solve practical problems is usually in the hundreds to thousands [17].

It should also be emphasized that to solve practical problems the qubits need to be error correctable. These are often referred to as *logical qubits* to make a distinction between them and a single *physical qubit*. The reason for this is because each logical qubit is actually composed of many physical qubits to add the needed redundancy in the system to be able to detect and then correct errors in the qubit state. Although each hardware platform is different, some of the current best estimates for superconducting circuit platforms suggest that as many as 1,000 physical qubits may be needed in each logical qubit of an error-correctable quantum computer [18]. Hence, there is still a long way to go before we have a full error-correctable quantum computer! This kind of device is often referred to as a *large-scale fault-tolerant quantum computer*, and is the kind of quantum computer that is often meant when one talks about the revolutionary potential of quantum computation.

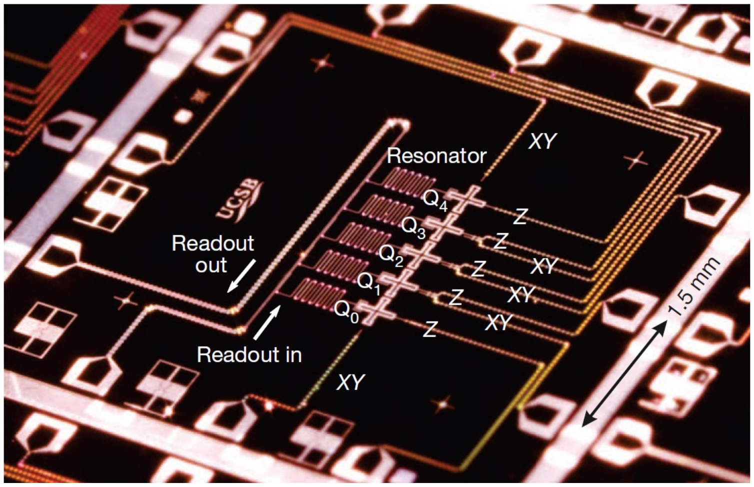


Figure 6.6: Example of a device with multiple transmon qubits that are directly coupled to each other through a capacitive interaction (image from [19]).

6.4 Physical Implementation of a Quantum Computer

There are many different hardware platforms being explored to realize a large-scale fault-tolerant quantum computer. Each requires having a sufficiently deep understanding of the physics involved in the hardware layer to gain a useful level of understanding of the advantages and disadvantages of the particular technological approach. Due to this, we will not attempt to review the various approaches exhaustively. Instead, we will focus on the approach using superconducting circuits that largely fits within the kind of analysis quantized circuit framework we will consider throughout this course.

Although many different superconducting circuit qubits have been developed, the most popular for current implementations is the transmon qubit, which we will learn about in more depth in the coming weeks. Being a planar technology, it is natural to array the transmon qubits in a 2D grid with each qubit designed to facilitate nearest-neighbor qubit-qubit interactions. This can be achieved through a direct capacitive coupling, or can be mediated through an electromagnetic resonator or tunable coupler. In a very simplistic description, a single mode of the electromagnetic resonator behaves like a parallel LC oscillator circuit that is coupled capacitively to the nearby transmon qubits. In the case of a tunable coupler, the coupler can take on various forms, but the most common approaches involve a tunable inductance (implemented with a SQUID) either directly connecting the qubits or implemented in a separate "transmon" that acts as the coupler between two actual transmon qubits. Example devices using these different coupling approaches are shown in Figs. 6.6 to 6.8.

Each of these different approaches have a number of tradeoffs that will dictate which kind of architecture is preferable for a given application. Some of the tradeoffs include how single and multi-qubit gates can be implemented, how many additional control lines are needed to operate a device, whether the design is robust to manufacturing imperfections, etc. All of these design decisions also impact other key figures of merit, such as how quickly different gates and measurements can occur and the ultimate susceptibility to noise of the

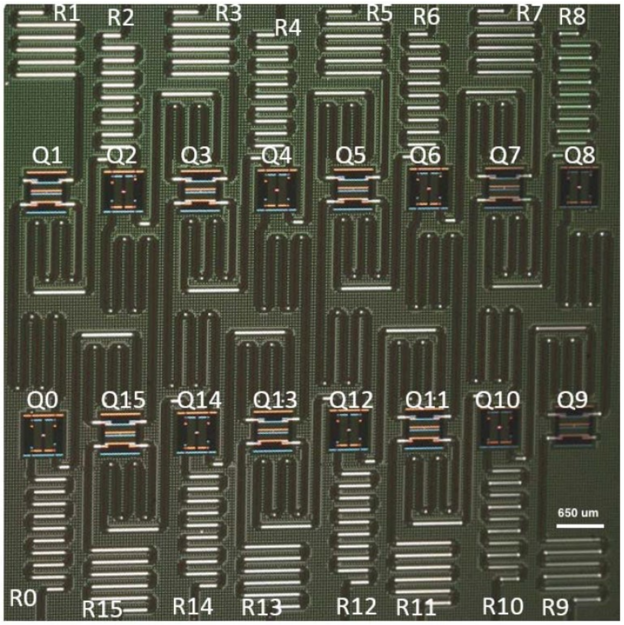


Figure 6.7: Example of a device with multiple transmon qubits that are coupled to each other through an electromagnetic resonator (image from [20]).

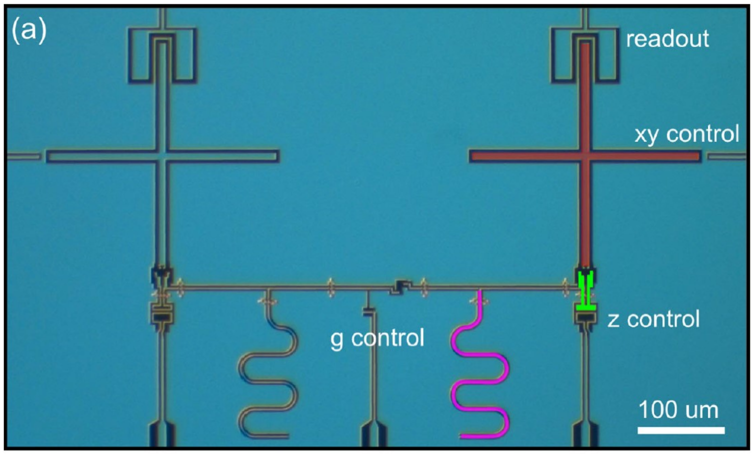


Figure 6.8: Example of a device with two transmon qubits that are coupled to each other through a tunable inductance (image from [21]).

different qubits. Overall, this represents a very challenging systems engineering problem, with significant progress still needing to be made to reach the desired performance levels to achieve a large-scale fault-tolerant quantum computer.

In addition to the coupling needed between different qubits, it is also necessary to measure the state of the qubit. This is accomplished by coupling the qubit to an electromagnetic resonator, known as the *readout resonator*. The effect of this interaction leads to a shift in the resonant frequency of the readout resonator. Importantly, the magnitude of this shift depends on the state of the qubit, which allows one to infer the state of the qubit by measuring the characteristics of the electromagnetic resonator. This is a “straightforward”

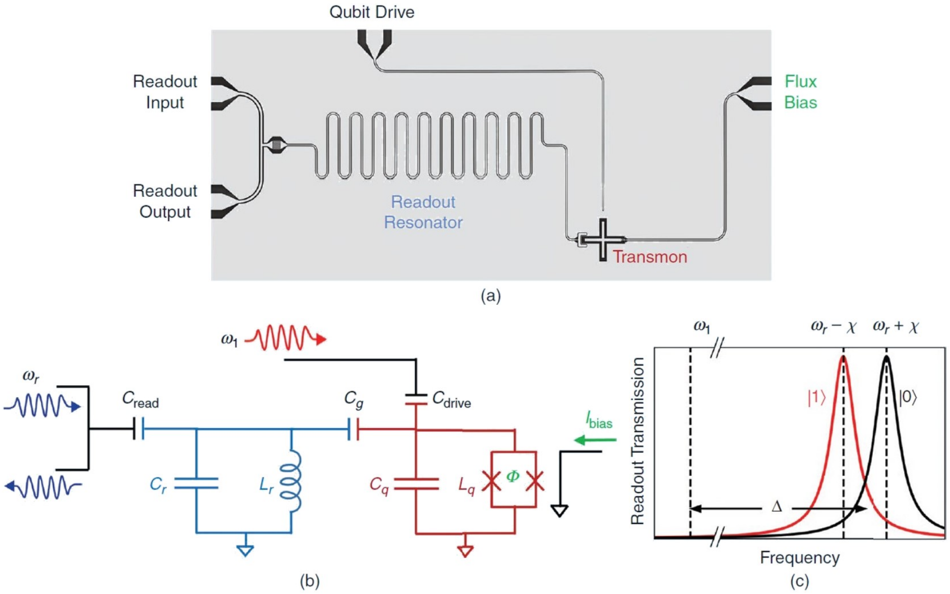


Figure 6.9: (a) Example of all of the features needed to control a single transmon qubit. (b) Simplified equivalent circuit schematic of the single qubit device. (c) Response of the electromagnetic resonator depending on the state of the transmon qubit. The shift in resonant frequency due to the transmon state can be used to infer the state of the qubit from a measurement of the electromagnetic resonator. Images are from [23].

measurement to perform in principle, although there are many important complexities that make this more challenging than a traditional measurement of an electromagnetic resonator. In particular, the amplitude of the probing signal must be kept very low to not accidentally "drive" the qubit and to allow for rapid measurements. This then requires a very sophisticated amplification chain to be implemented to successfully read out the state of the qubit [22].

Beyond reading the state of the qubit, it is also necessary to be able to implement drives on the different qubits to control their state. To achieve this, a drive line needs to be coupled to each of the qubits. Further, if the qubits are tunable via a superconducting circuit element known as a SQUID (which can be thought of as a magnetic-flux tunable inductor), an additional flux line needs to be brought close to the qubit as well to modify the magnetic flux intersecting the SQUID. Overall, this leads to a rather crowded number of control lines needed to interface with the various qubits, making this an area for future improvement as more and more qubits are integrated into larger quantum computers. An example of all the needed control lines to operate a single transmon qubit is shown in Fig. 6.9. In this figure, we also show the equivalent circuit schematic that can be used to (coarsely) analyze the behavior of this single qubit system. Simple generalizations to this circuit model can be used to conceptualize the behavior of more complicated multi-qubit circuits.

6.4.1 Summary

As we have just seen, the physical implementation of a quantum computer primarily consists of a few basic parts. We must have some physical implementation of qubits and a convenient scheme for interfacing them to each other, as well as interfacing them to measurement electronics. In current implementations, this coupling and interfacing to external electronics is most often achieved in the form of electromagnetic resonators, which at a very basic level behave similarly to LC oscillators.

To achieve a deeper understanding of the operation of a quantum computer, we will then need to study in more detail the quantum mechanical description of these constituent parts and (eventually) the quantum interactions between them. We will begin our study by revisiting the analysis of a quantum harmonic oscillator before linking this to the quantum description of an LC oscillator. We will study this system in depth, as well as consider what happens when we couple two quantum LC oscillators to each other. Later, we will consider the detailed analysis of a transmon qubit before then considering the interactions of such a qubit with our LC oscillators.

Chapter 7

Quantum Mechanics of Simple Circuits

7.1 Quantum Harmonic Oscillator – Revisited

As we discussed when first learning about the quantum harmonic oscillator, this problem represents one of the most commonly encountered models for many practical situations. Now that we have a better understanding of the mathematical framework of quantum mechanics, it is worth revisiting this problem to introduce a particularly useful approach to analyzing this system. We will use this description extensively as we go on to learn about how quantized circuits interact with complex quantum systems, such as artificial atoms. These interactions form a central part of hardware implementations of quantum information processing (such as in quantum computers), and as a result are of intense practical interest in the engineering of next-generation quantum technologies.

7.1.1 Factoring the Hamiltonian

To begin, recall that the Schrödinger equation for the quantum harmonic oscillator in dimensionless units could be written as

$$\frac{1}{2}\left[-\frac{d^2}{d\xi^2}+\xi^2\right]\psi(\xi)=\frac{E}{\hbar\omega}\psi(\xi), \tag{7.1}$$

where ω is the angular frequency of oscillations for the oscillator. We will find that the quantum harmonic oscillator provides a useful description of many physical systems, but that this occurs in a more abstract form where retaining an explicit "position" variable like ξ may not be particularly meaningful. As a result, it is useful to recast (7.1) into our more abstract operator notation as

$$\frac{1}{2}\left[\hat{\pi}^2+\hat{\xi}^2\right]|\psi\rangle=\frac{E}{\hbar\omega}|\psi\rangle, \tag{7.2}$$

where

$$\hat{\pi}=-i\frac{d}{d\xi} \quad \text{and} \quad \hat{\xi}=\xi \tag{7.3}$$

in the dimensionless coordinate space representation.

The trick now is to notice that $\hat{\pi}^2 + \hat{\xi}^2$ looks like a "sum of squares", so maybe it is useful to try and *factor* this operator expression somehow. Recalling basic algebra, we could factor the expression $a^2 + b^2$ as

$$a^2 + b^2 = (-ia + b)(ia + b). \tag{7.4}$$

Considering this, we wish to see whether the factorization

$$\frac{1}{\sqrt{2}}\big(-i\hat{\pi} + \hat{\xi}\big) \times \frac{1}{\sqrt{2}}\big(i\hat{\pi} + \hat{\xi}\big) \tag{7.5}$$

may be useful in expressing the Hamiltonian of (7.2). It turns out this doesn't quite work perfectly, but will still be very useful.

To see this, we expand out (7.5) to get

$$\begin{aligned}\frac{1}{\sqrt{2}}\big(-i\hat{\pi} + \hat{\xi}\big) \times \frac{1}{\sqrt{2}}\big(i\hat{\pi} + \hat{\xi}\big) &= \frac{1}{2}\big[\hat{\pi}^2 + \hat{\xi}^2\big] - \frac{i}{2}\big[\hat{\pi}\hat{\xi} - \hat{\xi}\hat{\pi}\big] \\ &= \frac{1}{2}\big[\hat{\pi}^2 + \hat{\xi}^2\big] - \frac{i}{2}[\hat{\pi}, \hat{\xi}],\end{aligned} \tag{7.6}$$

which we see differs from our desired Hamiltonian only by the commutator of $\hat{\pi}$ and $\hat{\xi}$. However, we can recall that the commutation relation for momentum and position has a very simple relationship of

$$[\hat{p}, \hat{x}] = -i\hbar. \tag{7.7}$$

Within our particular choice of dimensionless units, this reduces to

$$[\hat{\pi}, \hat{\xi}] = -i \tag{7.8}$$

so that (7.6) becomes

$$\frac{1}{\sqrt{2}}\big(-i\hat{\pi} + \hat{\xi}\big) \times \frac{1}{\sqrt{2}}\big(i\hat{\pi} + \hat{\xi}\big) = \frac{1}{2}\big[\hat{\pi}^2 + \hat{\xi}^2\big] - \frac{1}{2}. \tag{7.9}$$

If we define two "shorthand" operators (known as *ladder operators* for reasons we will get to shortly)

$$\hat{a}^\dagger = \frac{1}{\sqrt{2}}\big(-i\hat{\pi} + \hat{\xi}\big) \quad \text{and} \quad \hat{a} = \frac{1}{\sqrt{2}}\big(i\hat{\pi} + \hat{\xi}\big), \tag{7.10}$$

then we can express the Hamiltonian for the quantum harmonic oscillator in the simple form (after some slight rearranging)

$$\hat{H} = \frac{\hbar\omega}{2}\big[\hat{\pi}^2 + \hat{\xi}^2\big] = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right) \tag{7.11}$$

so that Schrödinger's equation would be written as

$$\hat{H}|\psi\rangle = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right)|\psi\rangle = E|\psi\rangle. \tag{7.12}$$

Before looking at some useful properties of the operators $\hat{a}$ and $\hat{a}^\dagger$, it is important to comment on our notation. First, we have denoted $\hat{a}^\dagger$ using our † symbol, which suggests that this is the adjoint to the operator $\hat{a}$. This is the case, although we won't go through the details of proving this. Second, we can quickly see from the explicit expressions for $\hat{a}$ and $\hat{a}^\dagger$ that these are distinct (i.e., they don't equal each other), and so we can conclude that these operators are *not Hermitian.* We have almost exclusively worked with Hermitian operators up to this point, so it is useful to emphasize that because $\hat{a}$ and $\hat{a}^\dagger$ are not Hermitian operators they do not correspond to measurable quantities and they are not guaranteed to have other useful properties such as completeness of their eigenfunctions. Despite this, we will see that these operators are still extremely useful.

7.1.2 Properties of the Ladder Operators

Now that we've seen how to write the Hamiltonian in terms of the *ladder operators* $\hat{a}$ and $\hat{a}^\dagger$, it is time to see how they can be used and determine some other useful properties of them. To begin, we can recall from our previous analysis of the quantum harmonic oscillator that the eigenvalues are all evenly spaced with specific values of

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega. \tag{7.13}$$

Comparing this to the eigenvalue equation of (7.12), we quickly see that

$$\hat{a}^\dagger\hat{a}|\psi_n\rangle = n|\psi_n\rangle. \tag{7.14}$$

(Note, this property can also be derived without already knowing the eigenvalues of the quantum harmonic oscillator.) We can interpret the value n here as denoting the number of quanta in the quantum harmonic oscillator, so it is typical to call the *Hermitian operator*

$$\hat{N} = \hat{a}^\dagger\hat{a} \tag{7.15}$$

the *number operator.* As a Hermitian operator, this *does* correspond to a measurable quantity even though our individual ladder operators were not themselves directly measurable.

As with many quantum operators, the commutator of the ladder operators is often of importance. The commutator for these operators is

$$[\hat{a}, \hat{a}^\dagger] = 1, \tag{7.16}$$

which is able to be established easily using the commutation relation for $\hat{\pi}$ and $\hat{\xi}$. Due to the prevalence of these kinds of operators, this commutation relation will also sometimes be referred to as the *standard bosonic commutation relation.* The phrase *boson* here is a particle physics term for subatomic particles that have integer values of spin. This is in contrast to *fermions*, which have odd half-integer values of spin. Bosons and fermions represent two fundamental classes of subatomic particles, and posses very different properties. We will not discuss these differences in significant depth in this course, but we will mention that an electron is an example of a fermion and a photon is an example of a boson, and that these different kinds of particles do exhibit significantly different behavior.

Now, with the commutation relation (7.16) we can examine some more properties of the ladder operators. In particular, we wish to see what happens to an eigenstate of the harmonic oscillator when we operate on it with $\hat{a}$. To determine what happens, we can operate on (7.14) with $\hat{a}$ to get

$$\hat{a}\big(\hat{a}^\dagger\hat{a}\big)|\psi_n\rangle = n\hat{a}|\psi_n\rangle. \tag{7.17}$$

We can make a trivial regrouping to write this as

$$\big(\hat{a}\hat{a}^\dagger\big)\big(\hat{a}|\psi_n\rangle\big) = n\big(\hat{a}|\psi_n\rangle\big). \tag{7.18}$$

We can now rearrange the commutation relation (7.16) as

$$\hat{a}\hat{a}^\dagger - \hat{a}^\dagger\hat{a} = 1 \rightarrow \hat{a}\hat{a}^\dagger = 1 + \hat{a}^\dagger\hat{a} \tag{7.19}$$

and substitute this result into the the left-hand side of (7.18) to get

$$\big(1 + \hat{a}^\dagger\hat{a}\big)\big(\hat{a}|\psi_n\rangle\big) = n\big(\hat{a}|\psi_n\rangle\big). \tag{7.20}$$

We can move one $\big(\hat{a}|\psi_n\rangle\big)$ to the right-hand side to find that

$$\hat{a}^\dagger\hat{a}\big(\hat{a}|\psi_n\rangle\big) = (n-1)\big(\hat{a}|\psi_n\rangle\big). \tag{7.21}$$

Considering that we know the operator $\hat{a}^\dagger\hat{a}$ returns as eigenvalue the number of the eigenstate, we see that the state $\big(\hat{a}|\psi_n\rangle\big)$ must in fact be proportional to $|\psi_{n-1}\rangle$. Specifically, this proportionality constant can be determined to be $\sqrt{n}$ so that

$$\hat{a}|\psi_n\rangle = \sqrt{n}|\psi_{n-1}\rangle. \tag{7.22}$$

We see that the operator $\hat{a}$ *lowers* our eigenstate by a single level. As a result, it is common to refer to $\hat{a}$ as the *lowering operator* or, more commonly, as the *annihilation operator* since it removes one quanta from the system.

A similar process can be performed by operating on (7.14) with $\hat{a}^\dagger$ to determine its properties. This shows that

$$\hat{a}^\dagger|\psi_n\rangle = \sqrt{n+1}|\psi_{n+1}\rangle. \tag{7.23}$$

We see that the operator $\hat{a}^\dagger$ *raises* our eigenstate by a single level. As a result, it is common to refer to $\hat{a}^\dagger$ as the *raising operator* or, more commonly, as the *creation operator* since it add one quanta to the system. The terminology of referring to $\hat{a}$ and $\hat{a}^\dagger$ as *ladder operators* can also now be understood. We can use these to go "up" and "down" the "ladder" of evenly-spaced eigenstates of the harmonic oscillator.

The ladder operators can also be used to determine the eigenstates of the harmonic oscillator in a simpler way than the "brute force" approach to solving a complicated differential equation we considered earlier. In particular, we can deduce that the harmonic oscillator has a lowest state that would correspond to $n = 0$. We would then determine that

$$\hat{a}|\psi_0\rangle = 0. \tag{7.24}$$

This can be written back in the coordinate space representation as

$$\frac{1}{\sqrt{2}}\left(\frac{d}{d\xi} + \xi\right)\psi_0(\xi) = 0. \tag{7.25}$$

This is a much simpler differential equation that can be solved with

$$\psi_0(\xi) = \pi^{-1/4} e^{-\xi^2/2}. \tag{7.26}$$

Higher eigenstates can then be found through repeated application of the creation operator to find

$$\left(\hat{a}^\dagger\right)^n |\psi_0\rangle = \sqrt{n!}|\psi_n\rangle \tag{7.27}$$

so that the normalized eigenstates become

$$|\psi_n\rangle = \frac{1}{\sqrt{n!}}\left(\hat{a}^\dagger\right)^n |\psi_0\rangle, \tag{7.28}$$

which will replicate the Hermite polynomial solution approach we discussed briefly earlier. Although this approach is not necessarily needed for solving the simple quantum harmonic oscillator problem, other more complex forms of the Schrödinger equation can be solved using a similar strategy if it is possible to factorize the differential operator [2]. As a result, this can be a useful strategy for solving other problems in quantum mechanics.

A final note is in order about a common terminology used when dealing with quantum harmonic oscillators. When we use ladder operators to work with harmonic oscillators in the abstract cases where a reference to "position" and "momentum" may no longer make sense, it is common to use a simplified representation for our basis set. In particular, we index them simply with the energy level number as $|\psi_n\rangle \rightarrow |n\rangle$. We typically refer to these states as *numbers states* or *Fock states*, and then sometimes refer to a *number-state* or *Fock-state representation* of a problem. This Fock-state representation is particularly useful when considering fundamental interactions between two physical systems that swap small numbers of quanta; for example, the absorption or emission of a finite number of photons by an atom is conveniently analyzed in the Fock-state representation. It should be noted that these Fock states do still correspond to different energy states in the quantum harmonic oscillator, and so they are still constitute an orthonormal basis. Hence, we have that $\langle m|n\rangle = \delta_{mn}$.

7.2 Canonical Quantization of an LC Oscillator

Now that we have built up the necessary tools, we can begin to analyze the quantum mechanics of broader classes of physical systems. In particular, we will consider the parallel

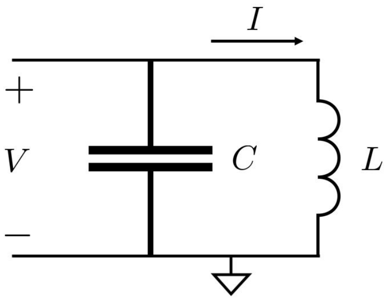


Figure 7.1: Circuit schematic of a parallel LC circuit (also referred to as an *LC tank circuit.*)

LC oscillator within the framework of *canonical quantization*. This example will illustrate some extremely fundamental techniques that can be applied with minimal changes to a surprisingly wide range of physical systems, including more complicated electrical circuits, electromagnetic fields, vibrations in crystal lattices, and many more!

Canonical quantization earns its name due to its link with Hamiltonian mechanics, where canonical position and momentum variables play a central role in the physical description of a system. Recall that for the LC oscillator shown in Fig. 7.1, we could easily write the total Hamiltonian in terms of voltage and current as

$$H = \frac{1}{2}CV^2 + \frac{1}{2}LI^2. \tag{7.29}$$

Unfortunately, our familiar quantities of voltage and current are not canonically conjugate variables, and so we had to determine a new set of variables to use in our Hamiltonian description of this system. Although there are different ways to do this, the particular option that will be most useful for us was to choose our canonical position variable to be the *node flux*, defined as

$$\phi(t) = \int_{-\infty}^{t} V(\tau)d\tau. \tag{7.30}$$

In your homework, you then showed that the canonical momentum that was conjugate to the node flux was equal to

$$Q = C\dot{\phi}. \tag{7.31}$$

Using this result and that the current could be expressed as

$$I = \phi/L, \tag{7.32}$$

we can rewrite our Hamiltonian of (7.29) into canonically conjugate variables as

$$H = \frac{1}{2C}Q^2 + \frac{1}{2L}\phi^2. \tag{7.33}$$

We now had multiple ways to find the equations of motion for this system. We could evaluate Hamilton's equations using the canonically conjugate variables ϕ and Q, or we could evaluate the Poisson bracket of our functions with the Hamiltonian. For our current purposes, it will be better to follow the Poisson bracket route. In terms of our conjugate variables ϕ and Q, the Poisson bracket is

$$\{f, g\} = \frac{\partial f}{\partial \phi}\frac{\partial g}{\partial Q} - \frac{\partial f}{\partial Q}\frac{\partial g}{\partial \phi}. \tag{7.34}$$

We then find our equation of motion for ϕ is

$$\frac{d}{dt}\phi = \{\phi, H\} = \left(\frac{\partial}{\partial \phi}\phi\right)\frac{\partial H}{\partial Q} - 0 = Q/C, \tag{7.35}$$

and our equation of motion for Q is

$$\frac{d}{dt}Q = \{Q, H\} = 0 - \left(\frac{\partial}{\partial Q}Q\right)\frac{\partial H}{\partial \phi} = -\phi/L. \tag{7.36}$$

We can consolidate these into a single equation by taking the time derivative of (7.35) and substituting in the result from (7.36) to get

$$\frac{d^2}{dt^2}\phi + \omega_0^2\phi = 0, \tag{7.37}$$

where $\omega_0 = 1/\sqrt{LC}$.

We can now *quantize* our description of this LC oscillator in a number of ways. One way is to suggest that because our Hamiltonian in (7.33) "looks" like that of a simple harmonic oscillator, we could simply rewrite our variables until this correspondence is more explicit and then just substitute conjugate variables for quantum "position" and "momentum" operators that we have been working with previously. This will work, but the more general process that works in a broader context is to follow Paul Dirac's *canonical quantization* procedure. In this approach, we quantize our classical theory by rewriting the Poisson bracket between two variables as a commutator between two quantum operators. More explicitly, we rewrite the Poisson bracket between two variables A and B of the classical theory as

$$\{A, B\} \to \frac{1}{i\hbar}[\hat{A}, \hat{B}], \tag{7.38}$$

where $\hat{A}$ and $\hat{B}$ are now quantum operators within the general mathematical framework of quantum mechanics.

For our case here, we would evaluate the Poisson bracket between ϕ and Q to get

$$\{Q, \phi\} = -1. \tag{7.39}$$

Under canonical quantization, we then *promote* ϕ and Q to be quantum operators with commutation relation

$$\frac{1}{i\hbar}[\hat{Q}, \hat{\phi}] = -1. \tag{7.40}$$

It is conventional to rewrite this commutation relation as

$$[\hat{Q}, \hat{\phi}] = -i\hbar. \tag{7.41}$$

It is common to refer to a commutation relation of this form as a *canonical commutation relation*, because all quantum operators that correspond to our canonical "position" and "momentum" variables from a classical theory will have this same commutation relation. In terms of these canonical operators, we can also write our quantum Hamiltonian operator by simply *promoting* our variables in the classical Hamiltonian to be quantum operators. For our case here, we get that (7.33) becomes

$$\hat{H} = \frac{1}{2C}\hat{Q}^2 + \frac{1}{2L}\hat{\phi}^2. \tag{7.42}$$

As mentioned previously, this Hamiltonian is of the same form as that of a quantum harmonic oscillator, and so similar strategies are applicable for working with this system. In particular, it is often very useful to introduce creation and annihilation operators for quantum harmonic oscillators. For this particular case, we will find that if we define

$$\hat{a} = \frac{1}{\sqrt{2L\hbar\omega_0}}\hat{\phi} + i\frac{1}{\sqrt{2C\hbar\omega_0}}\hat{Q}, \tag{7.43}$$

$$\hat{a}^\dagger = \frac{1}{\sqrt{2L\hbar\omega_0}}\hat{\phi} - i\frac{1}{\sqrt{2C\hbar\omega_0}}\hat{Q}, \tag{7.44}$$

we will be able to write our Hamiltonian in the desired form of

$$\hat{H} = \hbar\omega_0\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right). \tag{7.45}$$

We can also find that the introduced creation and annihilation operators satisfy the standard bosonic commutation relation of

$$[\hat{a}, \hat{a}^\dagger] = 1. \tag{7.46}$$

When we write our Hamiltonian like in (7.45), it becomes convenient to express the states of our system in a Fock-state representation that counts the number of quanta in our "quantum harmonic oscillator". We can then use the convenient properties of the Fock-state representation to learn important details about the quantum aspects of our system.

Before using the Fock-state representation, we should recall that when we first discussed the quantum harmonic oscillator we were able to develop an explicit solution in the "coordinate space" representation. In that representation, our wavefunction followed the form of a Gaussian times Hermite polynomials. Since our current problem is completely analogous to the quantum harmonic oscillator, we should still be able to have these same wavefunctions still describe our LC oscillator in some way. However, in this case we no longer have a "coordinate space" representation exactly, but instead have the variable that the Gaussian and Hermite polynomials are defined in terms of being our canonical "position" of nodal flux ϕ.

Hence, we interpret our wavefunction ψ as a function of ϕ with the quantity $|\psi(\phi)|^2$ giving us the probability distribution of the amplitude of the nodal flux in our LC oscillator.

Now, it is instructive to look at evaluating some of the statistical properties of our quantized $\hat{\phi}$ and $\hat{Q}$ in terms of Fock states to learn more about our quantum LC oscillator. To do this, it is useful to "invert" the expressions given in (7.43) and (7.44) to find that

$$\hat{\phi} = \phi_{\text{ZPF}}(\hat{a} + \hat{a}^\dagger) \tag{7.47}$$

$$\hat{Q} = -iQ_{\text{ZPF}}(\hat{a} - \hat{a}^\dagger), \tag{7.48}$$

where

$$\phi_{\text{ZPF}} = \sqrt{\frac{L\hbar\omega_0}{2}} \tag{7.49}$$

$$Q_{\text{ZPF}} = \sqrt{\frac{C\hbar\omega_0}{2}}. \tag{7.50}$$

The subscript ZPF stands for *zero-point fluctuations* for reasons that we will elucidate shortly.

We can now evaluate statistical properties of our canonical conjugate operators quite readily for different Fock states of our quantum LC oscillator. For instance, if we look at the expectation value of the ground state we can readily find that

$$\langle 0|\hat{\phi}|0\rangle = \langle 0|\phi_{\text{ZPF}}(\hat{a} + \hat{a}^\dagger)|0\rangle = 0, \tag{7.51}$$

with a similar result easily shown for $\langle 0|\hat{Q}|0\rangle = 0$. This seems to match our classical intuition, since in the ground state we have no quanta in the oscillator and we would think that it should then not be "moving". However, if we instead evaluate a higher-order statistical moment like the variance, we will find a perhaps somewhat surprising result. We have already seen that the expectation value of $\hat{\phi}$ and $\hat{Q}$ for the the ground state are 0, so we do not need to "remove" these from our operators to calculate the variance. Instead, we can just evaluate

$$\langle 0|\hat{\phi}^2|0\rangle = \phi_{\text{ZPF}}^2\langle 0|(\hat{a} + \hat{a}^\dagger)(\hat{a} + \hat{a}^\dagger)|0\rangle. \tag{7.52}$$

For practice, let's be very explicit with evaluating the operation of these combinations of ladder operators on the ground state. We can expand this result as

$$\langle 0|\hat{\phi}^2|0\rangle = \phi_{\text{ZPF}}^2\Big[\langle 0|\hat{a}\hat{a}|0\rangle + \langle 0|\hat{a}^\dagger\hat{a}|0\rangle + \langle 0|\hat{a}\hat{a}^\dagger|0\rangle + \langle 0|\hat{a}^\dagger\hat{a}^\dagger|0\rangle\Big]. \tag{7.53}$$

We know that $\hat{a}|0\rangle = 0$ by definition, so the first two terms on the right-hand side vanish. We can then apply the first creation operators in the remaining two terms to get

$$\langle 0|\hat{\phi}^2|0\rangle = \phi_{\text{ZPF}}^2\Big[\langle 0|\hat{a}|1\rangle + \langle 0|\hat{a}^\dagger|1\rangle\Big]. \tag{7.54}$$

Applying the next creation or annihilation operator we then get

$$\langle 0|\hat{\phi}^2|0\rangle = \phi_{\text{ZPF}}^2\bigg[\langle 0|0\rangle + \sqrt{2}\langle 0|2\rangle\bigg] = \phi_{\text{ZPF}}^2. \tag{7.55}$$

The final simplification occurs because the Fock states are orthonormal so that $\langle 0|0\rangle = 1$ and $\langle 0|2\rangle = 0$. We now see the surprising result that in the ground state of our oscillator the variance of our nodal flux is non-zero! This means that even when there are no quanta in the oscillator the nodal flux is still fluctuating around the mean value of 0 that we originally expected the system should be in. These fluctuations are referred to as *zero-point fluctuations* because they occur at the zero-point of our system. We also now see that the normalization of our operators in (7.47) and (7.48) are physically meaningful, as they represent the fundamental uncertainty of the value of these operators when in the ground state of the oscillator. We also see that we can in some sense "control" the amount of uncertainty in the various components of our oscillator by changing the values of L and C. In particular, if we design an oscillator with fixed frequency ω_0 and increase C, we will increase the size of fluctuations in $\hat{Q}$ while decreasing the fluctuations in $\hat{\phi}$. This kind of interplay in the uncertainty or sensitivity to different zero-point fluctuations can be an important consideration in designing quantum systems, as we will see in a different context when considering the design of "artificial atoms".

Before moving on, it is important to unpack this result of non-zero zero-point fluctuations. At face value, this result seems rather counter-intuitive. How is it possible that our oscillator is in some sense "moving" when we have put no quanta/energy into the system? The answer lies in the Hamiltonian, which we rewrite here for convenience as

$$\hat{H} = \hbar\omega_0\bigg(\hat{a}^\dagger\hat{a} + \frac{1}{2}\bigg). \tag{7.56}$$

If we evaluate the mean value of the energy in the ground state of our oscillator we can quickly find that

$$\langle 0|\hat{H}|0\rangle = \frac{1}{2}\hbar\omega_0, \tag{7.57}$$

which is non-zero! This value of $\hbar\omega_0/2$ is referred to as the *zero-point energy* of the system. It shows that even when we are in the ground state of the system, there is some amount of energy still existing in the oscillator. It is this zero-point energy that allows the zero-point fluctuations to occur in some sense. Although the zero-point energy of the system is an extremely important physical result, it turns out that a constant term like this in our Hamiltonian does not affect the equations of motion of a quantum system. As a result, it is extremely common in physics to simply leave off the zero-point energy from our Hamiltonian to be more concise. However, the zero-point energy is still "there", we just choose not to write it in contexts where it does not matter. In these cases, you will often see the Hamiltonian of a quantum harmonic oscillator written as

$$\hat{H} = \hbar\omega_0\hat{a}^\dagger\hat{a}. \tag{7.58}$$

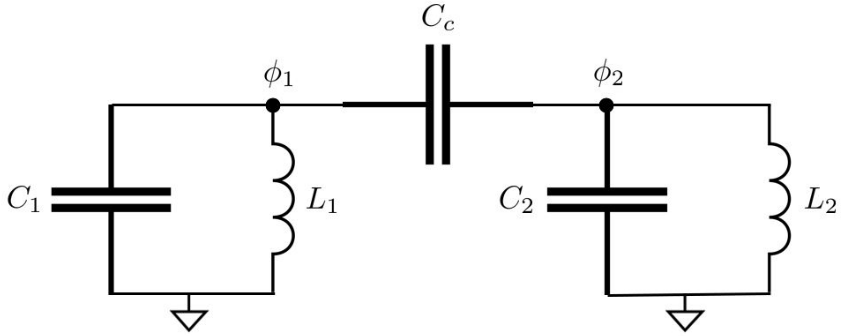


Figure 7.2: Circuit schematic of two capacitively-coupled LC oscillators.

One final note on terminology. For historical reasons, it is also common to refer to the ground state $|0\rangle$ of a quantum harmonic oscillator to be referred to as the *vacuum state* of the system (the terminology comes from the electromagnetic "vacuum" of space). Due to this, it will also be common to refer to zero-point fluctuations as *vacuum fluctuations*. Vacuum fluctuations play a very important role in quantum physics, and are central to explaining different kinds of interactions between different quantum systems. For instance, vacuum fluctuations are the main way to explain how spontaneous emission of photons occurs from atoms according to the theory of quantum electrodynamics, as well as more exotic kinds of quantum forces like the *Casimir effect* (which leads to an attractive or repulsive force between neutrally-charged objects when placed at very small distances from one another).

7.3 Capacitively-Coupled LC Oscillators

Now that we have some experience working with a single quantum LC oscillator, it is time to consider a more complicated circuit. In particular, we will consider the case of two LC oscillators coupled to each other through a capacitor, as shown in Fig. 7.2. We will illustrate two different approaches for solving this problem that constitute some of the main ways that more complicated practical problems are also tackled with.

The starting point for both approaches is to first determine the Lagrangian of our classical circuit. Once this is determined, we can follow through with quantizing the circuit in various ways. For the circuit shown in Fig. 7.2, we have already identified our generalized position variables as the nodal fluxes at the nodes of the two LC oscillators. In terms of these quantities, the Lagrangian is relatively simple to write down by simply subtracting the total inductive energy from the total capacitive energy. This gives us

$$\mathcal{L} = \frac{1}{2}C_1\dot{\phi}_1^2 + \frac{1}{2}C_2\dot{\phi}_2^2 + \frac{1}{2}C_c\left(\dot{\phi}_1 - \dot{\phi}_2\right)^2 - \frac{1}{2L_1}\phi_1^2 - \frac{1}{2L_2}\phi_2^2. \tag{7.59}$$

At this point, we will now approach the remainder of the quantization process following two different routes.

7.3.1 Find the Hamiltonian and then Perform Canonical Quantization

In our first approach, we will find our classical Hamiltonian in terms of canonical conjugate variables and then follow a canonical quantization process to introduce quantum operators and commutation relations. Considering this, our first step will be to convert our Lagrangian description of (7.59) into the equivalent Hamiltonian format. This begins by determining the conjugate momenta variables for our canonical position variables ϕ_1 and ϕ_2. Recall that the conjugate momenta Q_i is defined by

$$Q_i = \frac{\partial \mathcal{L}}{\partial \dot{\phi}_i}. \tag{7.60}$$

Evaluating the necessary derivatives, we find that

$$Q_1 = C_1\dot{\phi}_1 + C_c(\dot{\phi}_1 - \dot{\phi}_2), \tag{7.61}$$

$$Q_2 = C_2\dot{\phi}_2 - C_c(\dot{\phi}_1 - \dot{\phi}_2). \tag{7.62}$$

Note that the generalized velocities $\dot{\phi}_i$ get *mixed* in the definitions of the conjugate momenta. This is characteristic of coupling between systems, where the conjugate momenta have to be adjusted within the Hamiltonian formalism to arrive at the correct Hamiltonian and equations of motion in the end.

We now need to determine the Hamiltonian in terms of ϕ_i and Q_i. By definition, we find the Hamiltonian as

$$H = \sum_i Q_i\dot{\phi}_i - \mathcal{L}. \tag{7.63}$$

To evaluate this expression, we first need to write the $\dot{\phi}_i$ in terms of the Q_i. We can do this by inverting the relationships shown in (7.61) and (7.62). In this case, the process is relatively easy because we can consider (7.61) and (7.62) to define a 2x2 matrix equation that is straightforward to invert. Considering this, we will adopt a matrix notation for (7.61) and (7.62) as

$$\{Q\} = [C]\{\dot{\phi}\}, \tag{7.64}$$

where

$$[C] = \begin{bmatrix} C_1 + C_c & -C_c \\ -C_c & C_2 + C_c \end{bmatrix}. \tag{7.65}$$

We then have that

$$\{\dot{\phi}\} = [C]^{-1}\{Q\}, \tag{7.66}$$

with

$$[C]^{-1} = \frac{1}{C_1C_2 + C_cC_1 + C_cC_2}\begin{bmatrix} C_2 + C_c & C_c \\ C_c & C_1 + C_c \end{bmatrix}. \tag{7.67}$$

If we now define an inverse inductance matrix as

$$[L]^{-1} = \begin{bmatrix} L_1^{-1} & 0 \\ 0 & L_2^{-1} \end{bmatrix}, \tag{7.68}$$

we can write the original Lagrangian in matrix form as

$$\mathcal{L} = \frac{1}{2}\{\dot{\phi}\}^T[C]\{\dot{\phi}\} - \frac{1}{2}\{\phi\}^T[L]^{-1}\{\phi\}. \tag{7.69}$$

In terms of our conjugate momenta this becomes

$$\mathcal{L} = \frac{1}{2}\{Q\}^T[C]^{-1}\{Q\} - \frac{1}{2}\{\phi\}^T[L]^{-1}\{\phi\}. \tag{7.70}$$

The Hamiltonian can then be evaluated as

$$H = \{Q\}^T\{\dot{\phi}\} - \mathcal{L} = \frac{1}{2}\{Q\}^T[C]^{-1}\{Q\} + \frac{1}{2}\{\phi\}^T[L]^{-1}\{\phi\}. \tag{7.71}$$

At this point, it can be useful to "unpack" our matrix notation. If we expand everything out, we find that our Hamiltonian can be expressed as

$$H = \left(\frac{1}{2}[C]^{-1}_{11}Q_1^2 + \frac{1}{2L_1}\phi_1^2\right) + \left(\frac{1}{2}[C]^{-1}_{22}Q_2^2 + \frac{1}{2L_2}\phi_2^2\right) + \frac{1}{2}\big([C]^{-1}_{12} + [C]^{-1}_{21}\big)Q_1Q_2. \tag{7.72}$$

The first expressions in large parentheses look like the Hamiltonian we had for an LC oscillator, but now with frequency

$$\omega_i^2 = \frac{1}{L_i}[C]^{-1}_{ii} \tag{7.73}$$

instead of the frequencies of the *bare LC oscillators* (i.e., uncoupled oscillators with frequency $\omega_i^2 = 1/L_iC_i$). We can find our equations of motion for the different conjugate variables by taking the Poisson bracket with the Hamiltonian. For ϕ_1 and Q_1, we find that

$$\frac{d}{dt}\phi_1 = [C]^{-1}_{11}Q_1 + \frac{1}{2}\big([C]^{-1}_{12} + [C]^{-1}_{21}\big)Q_2, \tag{7.74}$$

$$\frac{d}{dt}Q_1 = -\frac{1}{L_1}\phi_1, \tag{7.75}$$

which can be combined to give

$$\frac{d^2}{dt^2}\phi_1 + \omega_1^2\phi_1 = \frac{1}{2}\big([C]^{-1}_{12} + [C]^{-1}_{21}\big)Q_2. \tag{7.76}$$

We see that this looks like the normal equation of motion for an LC oscillator, but there is now an additional driving term due to the capacitive coupling with the other oscillator. As mentioned previously, this coupling has shifted the resonant frequency of the oscillator, and we now see that it also allows for the oscillators to "drive" one another's motion. A similar equation can be derived for the other variables, with the same conclusions drawn.

We are also now in a position where we can follow the standard canonical quantization procedure that we established previously. First, we elevate our canonically conjugate variables to be quantum operators with canonical commutation relations of

$$[\hat{Q}_m, \hat{\phi}_n] = -i\hbar\delta_{mn}. \tag{7.77}$$

We can then write our Hamiltonian operator in terms of these operators. Alternatively, we can introduce creation and annihilation operators like before, where we will now have

$$\hat{a}_m = \frac{1}{\sqrt{2L_m\hbar\omega_m}}\hat{\phi}_m + i\sqrt{\frac{[C]^{-1}_{mm}}{2\hbar\omega_m}}\hat{Q}_m \tag{7.78}$$

$$\hat{a}^\dagger_m = \frac{1}{\sqrt{2L_m\hbar\omega_m}}\hat{\phi}_m - i\sqrt{\frac{[C]^{-1}_{mm}}{2\hbar\omega_m}}\hat{Q}_m, \tag{7.79}$$

with commutation relations of

$$[\hat{a}_m, \hat{a}^\dagger_n] = \delta_{mn}. \tag{7.80}$$

In terms of these ladder operators, our Hamiltonian now becomes

$$\hat{H} = \hbar\omega_1\left(\hat{a}^\dagger_1\hat{a}_1 + \frac{1}{2}\right) + \hbar\omega_2\left(\hat{a}^\dagger_2\hat{a}_2 + \frac{1}{2}\right) - \frac{1}{2}\hbar\sqrt{\omega_1\omega_2}\beta(\hat{a}_1 - \hat{a}^\dagger_1)(\hat{a}_2 - \hat{a}^\dagger_2), \tag{7.81}$$

where

$$\beta = \frac{C_c}{\sqrt{(C_1 + C_c)(C_2 + C_c)}}. \tag{7.82}$$

In certain situations, it is useful to analyze (7.81) directly. However, it is also very common to instead define a new basis that *diagonalizes* the Hamiltonian. The goal is to find a new basis such that our system behaves like two uncoupled quantum LC oscillators with overall Hamiltonian

$$\hat{H} = \sum_{m=1}^{2}\hbar\Omega_m\left(\hat{b}^\dagger_m\hat{b}_m + \frac{1}{2}\right). \tag{7.83}$$

Note that we have specifically written the oscillator frequencies and creation and annihilation operators with different symbols than those used in (7.81). This is because the new basis that allows us to treat the system as *uncoupled* LC oscillators will have different frequencies and operators than those in the coupled perspective of (7.81).

At this point it is instructive to think about how to express the quantum state of our coupled LC oscillators. We have already written our Hamiltonian in terms of ladder operators, which suggests that our Fock-state representation is likely to still be useful. This is the case, but we do have to consider a slight generalization of our description to properly describe the system with two coupled oscillators. The mathematical "tool" that is used to achieve this is known as a *tensor product.* Tensors are a generalization of matrices to be able to extend to higher dimensional sets of data. We can think of a tensor as having a number of indices that can run over the entire range of the Hilbert space that makes up each dimension of the tensor. For this particular situation, we would have the Hilbert spaces of oscillator 1 and 2, which we denote as $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively. The total Hilbert space that our system is analyzed in then becomes the tensor product of these two spaces, denoted as

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2. \tag{7.84}$$

A state within the overall Hilbert space then becomes the tensor product of states from the individual Hilbert spaces. As an example, if we have two quanta in oscillator 1 and 4 quanta in oscillator 2, we could write the state as $|2\rangle_1 \otimes |4\rangle_2$. This notation becomes rather cumbersome, especially if we have many coupled oscillators, and so simpler notations are often adopted. For instance, we could write this same state equivalently as

$$|2\rangle_1 \otimes |4\rangle_2 = |2\rangle_1|4\rangle_2 = |2\rangle|4\rangle = |2,4\rangle. \tag{7.85}$$

All of these notations are entirely equivalent, it is just a matter of preference for picking between one or another. Another point to note is that because each set of Fock states represents an orthonormal basis for that particular oscillator, our tensor product of these different states can be used as an orthonormal basis in the larger Hilbert space $\mathcal{H}$. One important consequence of this is that all of the states are mutually orthonormal, i.e.,

$$\langle m', n'|m, n\rangle = \delta_{mm'}\delta_{nn'}. \tag{7.86}$$

It should also be emphasized, that each state can exist in a quantum superposition with the same rules as before applied in regards to normalization of coefficients and statistical interpretation. For example, we could have the state

$$|\psi\rangle = \frac{1}{\sqrt{2}}|1,5\rangle - \frac{1}{\sqrt{2}}|3,2\rangle \tag{7.87}$$

as a valid quantum state. More generally, we can have

$$|\psi\rangle = \sum_{m,n} c_{mn}|m,n\rangle, \quad \sum_{m,n}|c_{mn}|^2 = 1. \tag{7.88}$$

Since we only have two oscillators, we can see that our tensor here can be easily represented using a matrix of the coefficients c_{mn}. However, if we had three oscillators coupled together, we would then need an additional index to traverse over in describing the state, and so a matrix no longer becomes a convenient way of imagining organizing the values in. This is where the more general concept of a tensor can become useful.

Now that we have a grasp on how we will express our states, we need to think about how our operators are defined within the combined Hilbert space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. It turns out that this is relatively straightforward, and just involves us taking our operators from the individual Hilbert spaces and taking the tensor product of them with the identity operator from the other Hilbert spaces. As an example, when we write $\hat{a}_1$ in (7.81), we actually mean

$$\hat{a}_1 = \hat{a}_1 \otimes \hat{I}_2, \tag{7.89}$$

where $\hat{I}_2$ is the identity operator in $\mathcal{H}_2$. To keep the notation simple, we almost never write our operators in this fully expanded tensor product notation. Instead, we just write $\hat{a}_1$ and understand what we actually mean by this in the larger Hilbert space. Before moving on, it is useful to point out that because our operators from the individual Hilbert spaces always have a tensor product with the identity operator in the other Hilbert spaces, we find that operators from different Hilbert spaces commute. As an example, there is no difference between $\hat{a}_1\hat{a}_2^\dagger$ and $\hat{a}_2^\dagger\hat{a}_1$. However, there is a difference between $\hat{a}_1\hat{a}_1^\dagger$ and $\hat{a}_1^\dagger\hat{a}_1$, as evidenced by the standard bosonic commutation relation.

7.3.2 Diagonalize Lagrangian Before Finding Hamiltonian

The other very common way to quantize a system like our coupled LC oscillators in Fig. 7.2 is to start by diagonalizing the Lagrangian. This results in a Lagrangian that looks like two uncoupled LC oscillators, for which the quantization process is extremely straightforward. More explicitly, we will find that the Hamiltonian will simply look like the summation of two uncoupled LC oscillators which can then be quantized immediately using the expressions for individual LC oscillators.

To see how this process works, we return to our matrix notation of the Lagrangian in (7.59), which was

$$\mathcal{L} = \frac{1}{2}\{\dot{\phi}\}^T[C]\{\dot{\phi}\} - \frac{1}{2}\{\phi\}^T[L]^{-1}\{\phi\}. \tag{7.90}$$

We cannot directly diagonalize this Lagrangian as it is currently written because the matrices $[C]$ and $[L]^{-1}$ do not commute, and so are not simultaneously diagonalizable. However, the extremely simple form of $[L]^{-1}$ does allow us to rescale our generalized position coordinates so that $[L]^{-1} \to [I]$. In particular, we define new coordinates

$$\Phi_i = \frac{1}{\sqrt{L_i}}\phi_i \tag{7.91}$$

so that the Lagrangian becomes

$$\mathcal{L} = \frac{1}{2}\{\dot{\Phi}\}^T[C']\{\dot{\Phi}\} - \frac{1}{2}\{\Phi\}^T[I]\{\Phi\}, \tag{7.92}$$

where

$$[C'] = \begin{bmatrix} L_1(C_1 + C_c) & -\sqrt{L_1L_2}C_c \\ -\sqrt{L_1L_2}C_c & L_2(C_2 + C_c) \end{bmatrix}. \tag{7.93}$$

We can now simultaneously diagonalize our system because $[C']$ commutes with the identity matrix $[I]$. This can be accomplished via an orthogonal transformation, defined by a matrix $[R]$ such that $[R]^T = [R]^{-1}$ and

$$[\Omega] = \begin{bmatrix} \Omega_1^{-2} & 0 \\ 0 & \Omega_2^{-2} \end{bmatrix} = [R][C'][R]^T. \tag{7.94}$$

Finding the explicit expressions for the Ω_i's and $[R]$ is rather tedious by hand, but is possible. It turns out that there are some useful tricks for solving this that we will not review at this time, but we will simply note that if we define $\alpha_i^{-2} = L_i(C_i + C_c)$ then we can write down the diagonal matrix elements as

$$\frac{1}{\Omega_1^2} = \frac{1}{2}\left(\frac{1}{\alpha_1^2} + \frac{1}{\alpha_2^2}\right) - \sqrt{\left[\frac{1}{2}\left(\frac{1}{\alpha_1^2} - \frac{1}{\alpha_2^2}\right)\right]^2 + \left(\frac{\beta}{\alpha_1\alpha_2}\right)^2} \tag{7.95}$$

$$\frac{1}{\Omega_2^2} = \frac{1}{2}\left(\frac{1}{\alpha_1^2} + \frac{1}{\alpha_2^2}\right) + \sqrt{\left[\frac{1}{2}\left(\frac{1}{\alpha_1^2} - \frac{1}{\alpha_2^2}\right)\right]^2 + \left(\frac{\beta}{\alpha_1\alpha_2}\right)^2}, \tag{7.96}$$

where β is the same as that given in (7.82). It does not take much imagination to guess that the explicit expressions for the matrix elements of $[R]$ can get rather nasty. Considering this, we will skip writing them down explicitly. Although this may seem somewhat lazy, it turns out that to analyze most quantum interactions it will be very important to know the frequencies given above, while having the explicit expressions for the transformation of our generalized coordinates into the diagonalized basis is less essential. Before moving on, it is important to note one of the characteristics of the expressions given above for the frequencies, as this general format is ubiquitous in the analysis of coupled systems. In particular, we see that the inverses of the two new frequencies get displaced above and below the average of the inverses of our "starting frequencies" α_1 and α_2.

Moving on, we can introduce the $[R]$ matrices into (7.92) in the following manner:

$$\begin{aligned}\mathcal{L} &= \frac{1}{2}\{\dot{\Phi}\}^T[I][C'][I]\{\dot{\Phi}\} - \frac{1}{2}\{\Phi\}^T[I][I][I]\{\Phi\}\\ &= \frac{1}{2}\{\dot{\Phi}\}^T[R]^T[R][C'][R]^T[R]\{\dot{\Phi}\} - \frac{1}{2}\{\Phi\}^T[R]^T[R][I][R]^T[R]\{\Phi\}.\end{aligned} \tag{7.97}$$

Defining new coordinates as

$$\{\Phi'\} = [R]\{\Phi\}, \tag{7.98}$$

we can simplify our above Lagrangian to be

$$\mathcal{L} = \frac{1}{2}\{\dot{\Phi}'\}^T[\Omega]\{\dot{\Phi}'\} - \frac{1}{2}\{\Phi'\}^T\{\Phi'\}. \tag{7.99}$$

At this point, the matrix notation is no longer really needed, and we can expand the Lagrangian as

$$\mathcal{L} = \sum_{m=1}^{2}\left[\frac{1}{2\Omega_m^2}\left(\dot{\Phi}'_m\right)^2 - \frac{1}{2}\left(\Phi'_m\right)^2\right]. \tag{7.100}$$

The Hamiltonian is found easily to be

$$H = \sum_{m=1}^{2}\left[\frac{1}{2}\Omega_m^2\left(Q'_m\right)^2 + \frac{1}{2}\left(\Phi'_m\right)^2\right], \tag{7.101}$$

where the conjugate momenta of each coordinate is

$$Q'_m = \frac{1}{\Omega_m^2}\dot{\Phi}'_m. \tag{7.102}$$

Note that because we diagonalized the Lagrangian first, we no longer have any direct mixing of variables from different degrees of freedom in the expressions for the conjugate momenta. However, the frequencies Ω_m still certainly have terms that are influenced by the coupling between the oscillators.

As promised this system now looks like two uncoupled LC oscillators. We can quantize the system by enforcing the canonical commutation relations as

$$[\hat{Q}'_m, \hat{\Phi}'_n] = -i\hbar\delta_{mn}. \tag{7.103}$$

The Hamiltonian operator then becomes

$$\hat{H} = \sum_{m=1}^{2}\left[\frac{1}{2}\Omega_m^2\big(\hat{Q}'_m\big)^2 + \frac{1}{2}\big(\hat{\Phi}'_m\big)^2\right]. \tag{7.104}$$

We can now introduce creation and annihilation operators as

$$\hat{b}_m = \frac{1}{\sqrt{2\hbar\Omega_m}}\hat{\Phi}'_m + i\sqrt{\frac{\Omega_m}{2\hbar}}\hat{Q}'_m \tag{7.105}$$

$$\hat{b}_m^\dagger = \frac{1}{\sqrt{2\hbar\Omega_m}}\hat{\Phi}'_m - i\sqrt{\frac{\Omega_m}{2\hbar}}\hat{Q}'_m \tag{7.106}$$

to transform our Hamiltonian into the standard form as

$$\hat{H} = \sum_{m=1}^{2}\hbar\Omega_m\left(\hat{b}_m^\dagger\hat{b}_m + \frac{1}{2}\right). \tag{7.107}$$

Although these results look rather different from those in Section 7.3.1, it is important to emphasize that they are completely equivalent. They simply amount to a different way of describing the physical processes involved in the capacitively-coupled LC oscillators, but they will be able to produce the same physical predictions. Whether one approach is more useful than the other depends on the context and goals of a particular analysis. However, in many cases, a description following the approach in this section is followed because the diagonalized Hamiltonian is easier to work with. Despite this, there are still situations where two rather different systems are coupled together (as we will discuss later in this course) and it is no longer as straightforward to arrive at a diagonalized description of the system like that shown in (7.107). Then, a Hamiltonian like that developed in Section 7.3.1 often will need to be used.

7.4 Schrödinger and Heisenberg Pictures

Up to this point, we have been considering quantum mechanics in what is often referred to as the *Schrödinger picture* for quantum dynamics. In this picture, we have the time-dependent Schrödinger equation

$$\hat{H}|\psi(t)\rangle = i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle \tag{7.108}$$

that we can solve to determine how $|\psi(t)\rangle$ evolves in time. Previously, we noted how we could develop a somewhat formal result for how $|\psi\rangle$ advances in time through the construction of the time evolution operator defined by

$$\hat{U}(t_f, t_i) = e^{-i\hat{H}(t_f - t_i)/\hbar}. \tag{7.109}$$

In terms of this operator, we could then write that

$$\hat{U}(t_f, t_i)|\psi(t_i)\rangle = |\psi(t_f)\rangle. \tag{7.110}$$

The Schrödinger picture is very useful, especially when working with systems where we have a convenient explicit form for the wavefunction, like in solving problems involving an electron moving through a potential energy landscape. However, when we have more abstract operators, like the creation and annihilation operators of a quantum harmonic oscillator, it can often be convenient to describe the dynamics of the quantum system in an equivalent, but different manner.

One popular alternative is the *Heisenberg picture*, where instead of having wavefunctions change in time we have our operators evolve in time. We can use the time evolution operator to express how the Heisenberg picture operator $\hat{X}_H$ changes in time as

$$\hat{X}_H(t) = \hat{U}^\dagger(t,0)\hat{X}\hat{U}(t,0), \tag{7.111}$$

where $\hat{X}$ is the Schrödinger picture form of the operator (i.e., the one that does not change in time). From (7.111), we see that the initial conditions for our Heisenberg picture operators are that they equal the Schrödinger picture version of the operators. This is necessary to ensure that the two pictures give equivalent results.

For instance, we can evaluate the expectation value of a particular operator in the Heisenberg picture in the following manner. If we have our system initially prepared in state $|\psi_H\rangle = |\psi(0)\rangle$, then the expectation value with operator $\hat{X}_H$ at time t is given in the Heisenberg picture as

$$\langle\psi_H|\hat{X}_H(t)|\psi_H\rangle. \tag{7.112}$$

For this picture to be consistent with the Schrödinger picture, we need to have

$$\langle\psi_H|\hat{X}_H(t)|\psi_H\rangle = \langle\psi(t)|\hat{X}|\psi(t)\rangle, \tag{7.113}$$

where the right-hand side of this equation is written in the Schrödinger picture. We can see that this is the case by expanding our Heisenberg picture operator in terms of the time evolution operators like in (7.111) to get

$$\langle\psi_H|\hat{X}_H(t)|\psi_H\rangle = \langle\psi_H|\hat{U}^\dagger(t,0)\hat{X}\hat{U}(t,0)|\psi_H\rangle. \tag{7.114}$$

We can now regroup our operators as

$$\langle\psi_H|\hat{X}_H(t)|\psi_H\rangle = \Big(\langle\psi_H|\hat{U}^\dagger(t,0)\Big)\hat{X}\Big(\hat{U}(t,0)|\psi_H\rangle\Big). \tag{7.115}$$

Recognizing that $|\psi_H\rangle = |\psi(0)\rangle$, we have that

$$\hat{U}(t,0)|\psi_H\rangle = \hat{U}(t,0)|\psi(0)\rangle = |\psi(t)\rangle \tag{7.116}$$

and

$$\langle\psi_H|\hat{U}^\dagger(t,0) = \Big(\hat{U}(t,0)|\psi_H\rangle\Big)^\dagger = \Big(|\psi(t)\rangle\Big)^\dagger = \langle\psi(t)|. \tag{7.117}$$

Using these results in (7.115), we arrive at the desired result that

$$\langle\psi_H|\hat{X}_H(t)|\psi_H\rangle = \langle\psi(t)|\hat{X}|\psi(t)\rangle, \tag{7.118}$$

showing that the Heisenberg and Schrödinger pictures can be used to make equivalent physical predictions.

One useful aspect of the Heisenberg picture is that it helps highlight the connection between classical and quantum mechanics in another way. In particular, this correspondence arises if we look at what the equation of motion will be for one of our quantum operators. We can determine the equation of motion by taking the time derivative of (7.111). We can evaluate this as

$$\frac{d}{dt}\hat{X}_H(t) = \frac{d}{dt}\Big(\hat{U}^\dagger(t,0)\hat{X}\hat{U}(t,0)\Big) = \Big(\frac{d}{dt}\hat{U}^\dagger(t,0)\Big)\hat{X}\hat{U}(t,0) + \hat{U}^\dagger(t,0)\hat{X}\Big(\frac{d}{dt}\hat{U}(t,0)\Big). \tag{7.119}$$

Considering the explicit form of the time evolution operator, we have that

$$\frac{d}{dt}\hat{U}(t,0) = \frac{d}{dt}e^{-i\hat{H}t/\hbar} = \frac{1}{i\hbar}\hat{H}\hat{U}(t,0) = \frac{1}{i\hbar}\hat{U}(t,0)\hat{H}. \tag{7.120}$$

The final equality comes from the fact that $[\hat{H},\hat{U}(t,0)] = 0$ for this case because $\hat{U}$ is only a function of the operator $\hat{H}$. We can also find for the adjoint operator $\hat{U}^\dagger$ that

$$\frac{d}{dt}\hat{U}^\dagger(t,0) = -\frac{1}{i\hbar}\hat{H}\hat{U}^\dagger(t,0). \tag{7.121}$$

Plugging these results into (7.119), we find that

$$\begin{aligned}\frac{d}{dt}\hat{X}_H(t) &= -\frac{1}{i\hbar}\Big(\hat{H}\hat{U}^\dagger(t,0)\hat{X}\hat{U}(t,0) - \hat{U}^\dagger(t,0)\hat{X}\hat{U}(t,0)\hat{H}\Big)\\ &= -\frac{1}{i\hbar}\Big(\hat{H}\hat{X}_H(t) - \hat{X}_H(t)\hat{H}\Big)\\ &= \frac{1}{i\hbar}[\hat{X}_H(t),\hat{H}].\end{aligned} \tag{7.122}$$

Recalling Dirac's rule for canonical quantization, we can recognize that the Heisenberg equation of motion given as

$$\frac{d}{dt}\hat{X}_H(t) = \frac{1}{i\hbar}[\hat{X}_H(t), \hat{H}] \tag{7.123}$$

will essentially match the equation of motion that we find for our classical variable through the Poisson bracket as

$$\frac{d}{dt}X = \{X, H\}. \tag{7.124}$$

Hence, in many cases we find that our quantum operators will have equations of motion that match our classical equations of motion. Although this correspondence can at first glance seem deceptively simple, it is important to use the actual Heisenberg picture formula to determine equations of motion for our operators to make sure that we do not make a simplification with functions that would not be acceptable with non-commuting quantum operators. However, this result is still very satisfying, and highlights that our classical intuition can still be useful for understanding how certain aspects of quantum problems will evolve in time.

7.5 Driven Quantum LC Oscillator and Coherent States

Now that we have some familiarity with the Heisenberg picture, it is interesting to look at the situation of a driven quantum LC oscillator. In particular, we will consider the case of a current source driving a parallel LC circuit, as shown in Fig. 7.3. Later, we will consider what happens when we drive our quantum LC oscillator with a *classical* current. This situation is particularly relevant to determining whether we can use classical control signals (e.g., a large current source in the lab) to control individual states of our quantum LC oscillator, which is a necessary requirement for building many quantum technologies of interest (e.g., quantum computers). As we will see, this classical drive leads to the quantum LC oscillator evolving in time according to a *coherent state*, which is a special superposition of all Fock states that behaves much like what we expect from a classical LC oscillator. As a result, we will need to utilize some other kind of quantum system than a quantum LC oscillator if we wish to control individual states in the way needed for many quantum technologies. Despite this difficulty, we will find that quantum LC oscillators (and other analogous physical systems) still form a very important piece of many quantum technologies.

7.5.1 Classical Case

Earlier in the course, we briefly discussed how a driving force could be included in the Lagrangian description of a mechanical oscillator by incorporating the additional potential energy due to the driving force. In that case, we had as a Lagrangian

$$\mathcal{L} = \frac{1}{2}m\dot{q}^2 - \frac{1}{2}kq^2 + qF(t), \tag{7.125}$$

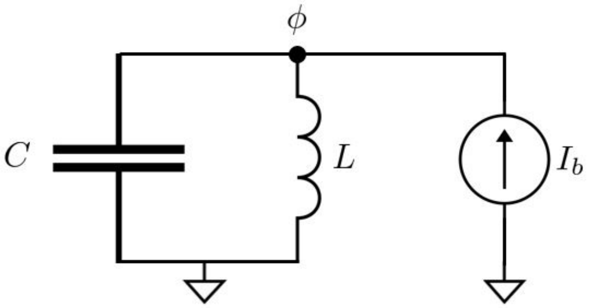


Figure 7.3: LC oscillator driven by a current source.

where $F(t)$ is the driving force term. In the case of our current-driven parallel LC circuit, the current source acts like the driving force acting on our "position variable" ϕ (i.e., the nodal flux). As a result, our Lagrangian becomes

$$\mathcal{L} = \frac{1}{2}C\dot{\phi}^2 - \frac{1}{2L}\phi^2 + \phi I_b(t). \tag{7.126}$$

We can follow our standard process to determine our Hamiltonian, which is

$$H = \frac{1}{2C}Q^2 + \frac{1}{2L}\phi^2 - \phi I_b(t). \tag{7.127}$$

From here, we can readily derive equations of motion for ϕ and Q.

However, for our current purposes, it will be more advantageous to increase the correspondence between the classical and quantum equations of motion that we will look at. Along these lines, we can introduce the complex *number*

$$a = \frac{1}{\sqrt{2L\omega_0}}\phi + i\frac{1}{\sqrt{2C\omega_0}}Q, \tag{7.128}$$

where $\omega_0 = 1/\sqrt{LC}$. This complex number will behave much like the annihilation operator $\hat{a}$ from the quantum mechanical analysis that we will look at soon. Now, in terms of (7.128), our Hamiltonian can be rewritten as

$$H = \omega_0 a^* a - \sqrt{\frac{L\omega_0}{2}}(a + a^*)I_b(t), \tag{7.129}$$

where a^* is simply the complex conjugate of a. We see that this looks similar to the Hamiltonian we would expect for a quantum harmonic oscillator if we make the substitution $a \to \hat{a},\ a^* \to \hat{a}^\dagger$; however, it does not match exactly. This situation can happen frequently when we try and derive a classical equation to "look" like a quantum one. The issue is that in the classical situation the terms a^*a and aa^* are the same, since complex numbers can commute. However, in the quantum case, we know that $\hat{a}^\dagger\hat{a} \neq \hat{a}\hat{a}^\dagger$. To help avoid these kinds of issues, it is common to "symmetrize" the classical equation by writing (7.129) in the equivalent form

$$H = \frac{1}{2}\omega_0\big(a^*a + aa^*\big) - \sqrt{\frac{L\omega_0}{2}}(a + a^*)I_b(t). \tag{7.130}$$

If we now make the substitution $a \to \hat{a},\ a^* \to \hat{a}^\dagger$ we will end up with the correct quantum Hamiltonian (up to some factors of $\hbar$ that are easy to insert where needed).

At this point, we can use a Poisson bracket to derive the equation of motion for a to find that (as you will show in a homework assignment)

$$\frac{d}{dt}a = -i\omega_0 a + i\sqrt{\frac{L\omega_0}{2}}I_b(t). \tag{7.131}$$

Solving this equation for an arbitrary $I_b(t)$ will typically require a numerical technique, but if we assume a simple form for I_b of

$$I_b(t) = \cos(\omega_d t), \tag{7.132}$$

then the solution becomes possible in closed form [24]. In particular, if we have as initial condition that $a(0) = \alpha$, where α is an arbitrary complex number, we find that the solution to (7.131) is

$$a(t) = \alpha e^{-i\omega_0 t} + a_p(t), \tag{7.133}$$

where $a_p(t)$ is the particular solution that takes a relatively complicated form that we will not need to reproduce here [24].

7.5.2 Coherent States

We have already discussed coherent states briefly earlier in the course. The particularly interesting property that we saw was that as the coherent state evolved in time the wavepacket formed by all the Fock states that make it up did not change its shape, but that the mean position of the wavepacket oscillated back and forth along the classical trajectory of the particle. We will now look at a few of these properties in a little more detail.

Mathematically, coherent states can be found to be eigenfunctions of the annihilation operator. That is, for the coherent state $|\alpha\rangle$ we have that

$$\hat{a}|\alpha\rangle = \alpha|\alpha\rangle, \tag{7.134}$$

where α is an arbitrary complex number. Although, $|\alpha\rangle$ is *not* an eigenfunction of $\hat{a}^\dagger$, we do still have the adjoint expression to (7.134) that

$$\langle\alpha|\hat{a}^\dagger = \alpha^*\langle\alpha|. \tag{7.135}$$

From this, we can readily determine that the expectation value of the number operator for a coherent state is

$$\langle\alpha|\hat{a}^\dagger\hat{a}|\alpha\rangle = |\alpha|^2\langle\alpha|\alpha\rangle = |\alpha|^2. \tag{7.136}$$

Considering this, we see that the parameter $|\alpha|$ is related to the mean number of quanta in the quantum harmonic oscillator.

We can also look at the expected value of the quantum LC oscillator $\hat{\phi}$ and $\hat{Q}$ by recalling that these operators can be expanded in terms of $\hat{a}$ and $\hat{a}^\dagger$ as

$$\hat{\phi} = \phi_{\mathrm{ZPF}}\big(\hat{a} + \hat{a}^\dagger\big), \quad \hat{Q} = -iQ_{\mathrm{ZPF}}\big(\hat{a} - \hat{a}^\dagger\big), \tag{7.137}$$

where ϕ_{ZPF} and Q_{ZPF} are constants that characterize the magnitude of zero-point fluctuations in ϕ and Q. We can then quickly find that

$$\langle\alpha|\hat{\phi}|\alpha\rangle = 2\phi_{\mathrm{ZPF}}\,\mathrm{Re}\{\alpha\}, \;\; \langle\alpha|\hat{Q}|\alpha\rangle = 2Q_{\mathrm{ZPF}}\,\mathrm{Im}\{\alpha\}. \tag{7.138}$$

The variances of these operators can also be found, with the result being independent of α. One can also find that these variances "saturate" the uncertainty principle for these operators, meaning that the product of the variances equals the minimal value possible according to quantum mechanics. As a result, it is common to refer to a coherent state as a *minimum-uncertainty wavepacket*. One important aspect of these results is that as $|\alpha|^2$ becomes very large the resulting quantum fluctuations in ϕ and Q become extremely small relative to their mean values. This is one of the properties needed for the coherent state to behave like a classical state.

The other property needed of the coherent state to behave like a classical state is that we expect it to somehow follow the trajectory of a classical LC oscillator (i.e., like a simple harmonic oscillator). To show this, it is helpful to expand the coherent state in terms of Fock states as

$$|\alpha\rangle = e^{-(1/2)|\alpha|^2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}|n\rangle. \tag{7.139}$$

This result can be established in a number of ways, but one simple way is by using the eigenvalue equation (7.134) and making use of the known properties of $\hat{a}$ acting on Fock states. We now wish to compute the time evolution of $|\alpha\rangle$ by applying the time evolution operator to it. For the free quantum harmonic oscillator, we need to compute the result of

$$e^{-i\hat{H}t/\hbar}|\alpha\rangle = e^{-i\omega_0(\hat{a}^\dagger\hat{a} + 1/2)t}|\alpha\rangle. \tag{7.140}$$

By expanding $|\alpha\rangle$ according to (7.139), we find that

$$\begin{aligned} e^{-i\hat{H}t/\hbar}|\alpha\rangle &= e^{-i\omega_0 t/2}e^{-(1/2)|\alpha|^2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}e^{-i\omega_0 t\,\hat{a}^\dagger\hat{a}}|n\rangle \\ &= e^{-i\omega_0 t/2}e^{-(1/2)|\alpha|^2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}e^{-i\omega_0 nt}|n\rangle \\ &= e^{-i\omega_0 t/2}e^{-(1/2)|\alpha|^2}\sum_{n=0}^{\infty}\frac{(\alpha e^{-i\omega_0 t})^n}{\sqrt{n!}}|n\rangle \\ &= e^{-i\omega_0 t/2}|e^{-i\omega_0 t}\alpha\rangle. \end{aligned} \tag{7.141}$$

That is, up to a global phase of $\exp[-i\omega_0 t/2]$, we see that the coherent state simply evolves into a new coherent state oscillating at the frequency of the free quantum harmonic oscillator.

If we consider (7.133) without the driving term that leads to $a_p(t)$, then we find that the complex number $\alpha(t)$ that characterizes our coherent state matches the classical trajectory given by

$$a(t) = \alpha e^{-i\omega_0 t}. \tag{7.142}$$

We can further find that

$$\langle e^{-i\omega_0 t}\alpha|\hat{\phi}|e^{-i\omega_0 t}\alpha\rangle = 2\phi_{\text{ZPF}} \operatorname{Re}\{\alpha e^{-i\omega_0 t}\}, \;\; \langle e^{-i\omega_0 t}\alpha|\hat{Q}|e^{-i\omega_0 t}\alpha\rangle = 2Q_{\text{ZPF}} \operatorname{Im}\{\alpha e^{-i\omega_0 t}\}, \tag{7.143}$$

which shows that the expected values of ϕ and Q match the classical oscillation cycle of the LC oscillator.

Before moving on, one more useful property of coherent states is the result of their inner product. In particular, for coherent states $|\alpha\rangle$ and $|\beta\rangle$, we have that

$$\langle\beta|\alpha\rangle = e^{-(1/2)|\beta-\alpha|^2} e^{i\text{Im}\{\beta^*\alpha\}}. \tag{7.144}$$

Hence, we see that coherent states are *not orthogonal*, however, if α and β are significantly different from each other we see that the first term on the right-hand side will significantly suppress the inner product. As a result, two coherent states with $|\beta-\alpha|^2$ being large will be nearly orthogonal.

7.5.3 Quantum Case

We will now quickly investigate how the quantum LC oscillator behaves when driven by a classical current source. To do this, we will look at the Heisenberg equation of motion for $\hat{a}$. We can derive this by noting that our quantum Hamiltonian will be

$$\hat{H} = \hbar\omega_0\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right) - \phi_{\text{ZPF}}\left(\hat{a} + \hat{a}^\dagger\right)\cos(\omega_d t), \tag{7.145}$$

where we have explicitly replaced the classical $I_b(t)$ by the assumed simple form discussed earlier in Section 7.5.1. The desired Heisenberg equation of motion is found from

$$\frac{d}{dt}\hat{a}_H = \frac{1}{i\hbar}[\hat{a}_H, \hat{H}], \tag{7.146}$$

which can be evaluated to give (as you will show in your homework)

$$\frac{d}{dt}\hat{a}_H = -i\omega_0\hat{a}_H + i\frac{\phi_{\text{ZPF}}}{\hbar}\cos(\omega_d t). \tag{7.147}$$

How do we solve an operator equation like this? In general, this can be a tricky process, but the typical strategy is to represent our operator in matrix form and then solve the corresponding ordinary differential equation defined by the matrix system. Here, we just want to look at the special case of what happens if our initial condition is that the system starts in a particular coherent state $|\alpha\rangle$. We can then look at how the expected value of our system changes by evaluating $\langle\alpha|\hat{a}_H(t)|\alpha\rangle = \langle a_H(t)\rangle_\alpha$ to gain our desired insight into this system.

To determine $\langle a_H(t)\rangle_\alpha$, we simply take the expectation value of (7.147) with $|\alpha\rangle$ to find that

$$\frac{d}{dt}\langle a_H(t)\rangle_\alpha = -i\omega_0\langle a_H(t)\rangle_\alpha + i\frac{\phi_{\text{ZPF}}}{\hbar}\cos(\omega_d t). \tag{7.148}$$

As initial condition, we must have that $\langle a_H(0)\rangle_\alpha = \langle a\rangle_\alpha$, that is, that it matches the Schrödinger picture expectation value initially. From the properties of coherent states, we can readily find that

$$\langle a_H(0)\rangle_\alpha = \langle a\rangle_\alpha = \alpha \tag{7.149}$$

because $|\alpha\rangle$ is an eigenstate of $\hat{a}$. With this initial condition, we can now solve (7.148) using standard methods to find that

$$\langle a_H(t)\rangle_\alpha = \alpha e^{-i\omega_0 t} + a_p(t), \tag{7.150}$$

where $a_p(t)$ is the same particular solution that we had in the classical case of (7.133) (up to a multiplicative factor in the scaling of the current source by $1/\hbar$). Hence, we see that so long as our classically-driven quantum LC oscillator starts in a coherent state, it will remain in a coherent state and evolve in time exactly following the classical trajectory of a driven LC oscillator [24]. Noting that the ground state $|0\rangle$ is itself a coherent state, we see that if we simply subject a quantum LC oscillator to a classical drive we will essentially only observe very "classical" behavior. Considering this, if we wish to design a technology that truly harnesses quantum behavior we will need to look for a different approach than using classical drives on quantum LC oscillators (or their other physical analogues). We will begin to study physical systems that satisfy this goal in the coming weeks.

7.6 Practice Problems

1. In this problem, you will prove the commutation relation

$$[\hat{a}, \hat{a}^\dagger] = 1 \tag{7.151}$$

for the creation and annihilation operators using two different methods.

(a) Use the definitions that

$$\hat{a}^\dagger = \frac{1}{\sqrt{2}}\left(-\frac{d}{d\xi} + \xi\right), \quad \hat{a} = \frac{1}{\sqrt{2}}\left(\frac{d}{d\xi} + \xi\right), \tag{7.152}$$

to prove (7.151).

Note: To help with getting the calculus correct, it is useful to think of applying the commutation relation to a function $f(\xi)$. You should then show that $[\hat{a}, \hat{a}^\dagger]f = f$. Since f is arbitrary we can conclude that $[\hat{a}, \hat{a}^\dagger] = 1$, as desired.

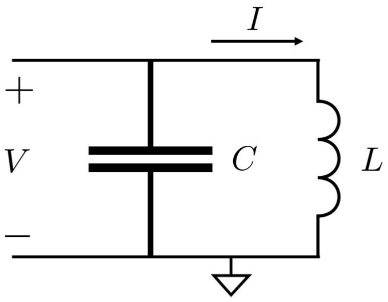


Figure 7.4: Circuit schematic of a parallel LC oscillator.

(b) Use the operator form of the creation and annihilation operators as

$$\hat{a}^\dagger = \frac{1}{\sqrt{2}}\left(-i\hat{\pi} + \hat{\xi}\right), \quad \hat{a} = \frac{1}{\sqrt{2}}\left(i\hat{\pi} + \hat{\xi}\right) \tag{7.153}$$

to prove (7.151).

Note: You will also want to use the canonical commutation relation $[\hat{\pi}, \hat{\xi}] = -i$ in your proof.

2. Using the commutation relation from (7.151), show that the Hamiltonian

$$\hat{H} = \frac{\hbar\omega}{2}\left(\hat{a}^\dagger\hat{a} + \hat{a}\hat{a}^\dagger\right) \tag{7.154}$$

is equivalent to the Hamiltonian

$$\hat{H} = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right) \tag{7.155}$$

that we had in the notes for a quantum harmonic oscillator.

3. For this problem, you will be working with a parallel LC oscillator, as shown in Fig. 7.4.

   (a) Find the variance of the charge Q for the oscillator in its ground state $|0\rangle$.

   (b) Using the result of (a) and the result for the variance of the nodal flux ϕ given in class, show that the state $|0\rangle$ is a *minimum uncertainty state*. This means that it has an uncertainty principle relationship between Q and ϕ that is the minimum possible according to quantum mechanics, i.e.,

   $$\sigma_\phi \sigma_Q = \frac{\hbar}{2}, \tag{7.156}$$

   where σ_X is the standard deviation of observable X.

4. For this problem, you will be working with a parallel LC oscillator, as shown in Fig. 7.4.

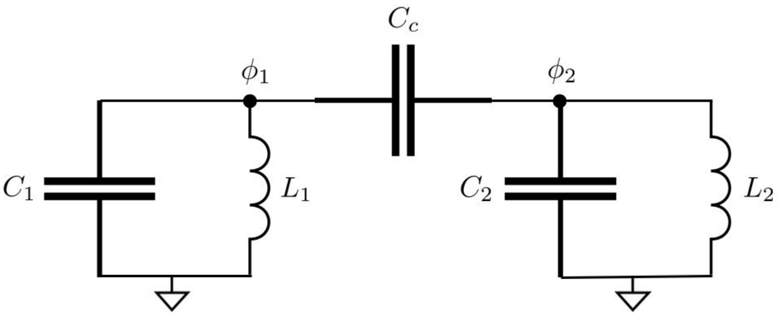


Figure 7.5: Circuit schematic of two capacitively-coupled LC oscillators.

(a) Given that for this circuit $V = Q/C$ and $I = \phi/L$, use the rules of canonical quantization to determine that the commutator between the voltage operator $\hat{V}$ and current operator $\hat{I}$ is

$$[\hat{V}, \hat{I}] = -i\hbar\omega_0^2, \tag{7.157}$$

where ω_0 is the resonant frequency of the LC oscillator.

(b) Using the expressions for $\hat{\phi}$ and $\hat{Q}$ given in class, determine expressions for $\hat{V}$ and $\hat{I}$ in terms of the ladder operators for this circuit.

(c) Using your result from (b) and the bosonic commutation relation given in (7.151), show that

$$[\hat{V}, \hat{I}] = -i\hbar\omega_0^2. \tag{7.158}$$

5. For this problem, assume that you are working with a set of capacitively coupled LC oscillators. The system has been quantized as discussed in the notes, and the resulting Hamiltonian is

$$\hat{H} = \hbar\omega_1\left(\hat{a}_1^\dagger\hat{a}_1 + \frac{1}{2}\right) + \hbar\omega_2\left(\hat{a}_2^\dagger\hat{a}_2 + \frac{1}{2}\right) - \frac{1}{2}\hbar\sqrt{\omega_1\omega_2}\beta(\hat{a}_1 - \hat{a}_1^\dagger)(\hat{a}_2 - \hat{a}_2^\dagger). \tag{7.159}$$

If we represent Fock states of this coupled system with a ket as $|m, n\rangle$ and a bra as $\langle m, n|$, where m corresponds to the number of quanta in oscillator 1 and n corresponds to the number of quanta in oscillator 2, evaluate the following expressions:

(a) $\hat{a}_1^\dagger\hat{a}_2^\dagger|3, 5\rangle$,
(b) $\langle 1, 2|\hat{a}_1\hat{a}_2^\dagger|2, 1\rangle$,
(c) $\langle 1, 2|\hat{a}_1^\dagger\hat{a}_2|2, 1\rangle$,
(d) $\langle 0, 1|\hat{a}_1\hat{a}_2^\dagger|0, 0\rangle$.

6. This problem considers two coupled quantum LC oscillators defined in the Hilbert space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, where $\mathcal{H}_j$ is the original Hilbert space associated with the jth oscillator. Creation and annihilation operators from $\mathcal{H}_j$ are denoted in $\mathcal{H}$ as $\hat{a}_j^\dagger$ and $\hat{a}_j$ respectively. The state $|m, n\rangle$ in $\mathcal{H}$ corresponds to a Fock state with m quanta in oscillator 1 and n quanta in oscillator 2.

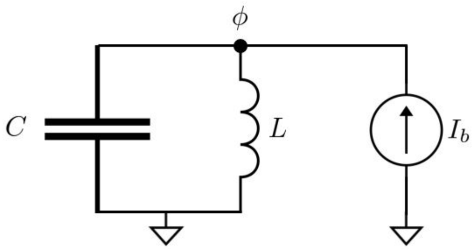


Figure 7.6: LC oscillator driven by a current source.

(a) Evaluate $\langle m, n|\big(\hat{a}_1^\dagger\hat{a}_2^\dagger\hat{a}_1\hat{a}_2 + \hat{a}_1^\dagger\hat{a}_1\hat{a}_2^\dagger\hat{a}_2\big)|m, n\rangle$.

(b) Evaluate $\langle 10, 3|\big(\hat{a}_1^\dagger\hat{a}_2 + \hat{a}_1\hat{a}_2^\dagger\big)|9, 4\rangle$.

(c) Evaluate $[\hat{a}_1, [\hat{a}_1, \hat{a}_2^\dagger]] + [\hat{a}_1, [\hat{a}_1, \hat{a}_1^\dagger]] + [\hat{a}_1^\dagger[\hat{a}_2, \hat{a}_2^\dagger], [\hat{a}_2^\dagger, \hat{a}_2]\hat{a}_1]$

7. In this problem, you will work with a parallel LC oscillator that is driven by a current source, as shown in Fig. 7.6.

(a) In class, we said that the classical Hamiltonian for this circuit was equal to

$$H = \frac{1}{2C}Q^2 + \frac{1}{2L}\phi^2 - \phi I_b(t). \tag{7.160}$$

By defining the complex number

$$a = \frac{1}{\sqrt{2L\omega_0}}\phi + i\frac{1}{\sqrt{2C\omega_0}}Q, \tag{7.161}$$

show that the classical Hamiltonian can be written as

$$H = \frac{1}{2}\omega_0\big(a^*a + aa^*\big) - \sqrt{\frac{L\omega_0}{2}}\big(a + a^*\big)I_b(t). \tag{7.162}$$

Hint: Express ϕ and Q in terms of a and a^*. Then, plug these expressions for ϕ and Q into (7.160) and simplify.

(b) Use a Poisson bracket to show that the equation of motion for a is

$$\frac{d}{dt}a = -i\omega_0 a + i\sqrt{\frac{L\omega_0}{2}}I_b(t) \tag{7.163}$$

Hint: You may want to revisit a property you proved about Poisson brackets applied to products of functions in Chapter 2 Practice Problems to simplify the process of evaluating the Poisson brackets for this problem.

(c) We will now move to the quantum case of Fig. 7.6, but the drive $I_b(t)$ will still be classical. In this case, we had that the quantum Hamiltonian of the circuit was

$$\hat{H} = \frac{\hbar\omega_0}{2}\left(\hat{a}^\dagger\hat{a} + \hat{a}\hat{a}^\dagger\right) - \phi_{\text{ZPF}}\big(\hat{a} + \hat{a}^\dagger\big)I_b(t). \tag{7.164}$$

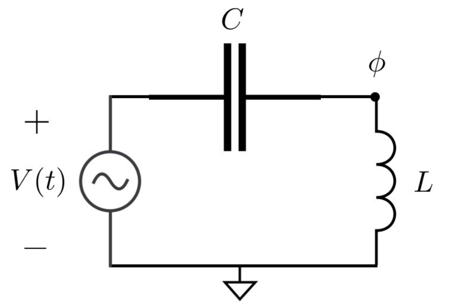


Figure 7.7: Circuit schematic of a voltage-driven series LC resonator.

Show that the Heisenberg equation of motion for $\hat{a}_H$ is

$$\frac{d}{dt}\hat{a}_H = -i\omega_0\hat{a}_H + i\sqrt{\frac{L\omega_0}{2\hbar}}I_b(t). \tag{7.165}$$

Note that besides the factor of $\hbar$, this equation has the same form as the classical equation given in (7.163).

8. In this problem, you will analyze the circuit shown in Fig. 7.7. You may assume that the voltage $V_b(t)$ is a known quantity corresponding to a classical drive for this circuit. You may also use previous results from a classical analysis of this circuit without fully re-deriving them, but be clear when defining them in this exam to ensure your work and notation can be followed easily.

   (a) Following a canonical quantization procedure, find the quantum Hamiltonian for this circuit written in terms of the node flux operator $\hat{\phi}$, its canonical conjugate operator $\hat{Q}$, and the classical drive V_b.

   (b) Determine appropriate expressions for the ladder operators $\hat{a}$ and $\hat{a}^\dagger$ for this circuit in terms of $\hat{\phi}$ and $\hat{Q}$. Show that your ladder operators satisfy the correct commutation relation of $[\hat{a}, \hat{a}^\dagger] = 1$.

   (c) Using the ladder operators determined in (b), show that the Hamiltonian you found in (a) can be written as

   $$\hat{H} = \hbar\omega_0\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right) - i\sqrt{\frac{C\hbar\omega_0}{2}}(\hat{a} - \hat{a}^\dagger)V_b(t),$$

   where ω_0 is the resonant frequency of the circuit.

   (d) Find the Heisenberg equation of motion for the operator $\hat{a}_H(t)$ and simplify it down as much as possible.

9. Show that the variance of an arbitrary quantum operator $\hat{X}_H(t)$ computed in the Heisenberg picture will match the result for the corresponding Schrödinger picture operator $\hat{X}$. Note that the variance in either picture can be computed as

$$\mathrm{var}(X(t)) = \langle X^2(t)\rangle - \langle X(t)\rangle^2, \tag{7.166}$$

where

$$\langle X(t)\rangle = \langle\psi_H|\hat{X}_H(t)|\psi_H\rangle \tag{7.167}$$

in the Heisenberg picture and

$$\langle X(t)\rangle = \langle\psi(t)|\hat{X}|\psi(t)\rangle \tag{7.168}$$

in the Schrödinger picture.

10. You are given an arbitrary Heisenberg picture operator $\hat{X}_H(t)$ that commutes with the Hamiltonian. Show that the expectation value of this operator for an arbitrary state $|\psi\rangle$ does not depend on time.

11. In this problem, you will work with the coherent state $|\alpha\rangle$ of a parallel LC oscillator.

    (a) Show that

    $$\langle\alpha|\hat{\phi}|\alpha\rangle = 2\phi_{\mathrm{ZPF}}\mathrm{Re}\{\alpha\}. \tag{7.169}$$

    (b) Show that

    $$\langle\alpha|\hat{Q}|\alpha\rangle = 2Q_{\mathrm{ZPF}}\mathrm{Im}\{\alpha\}. \tag{7.170}$$

12. Show that a coherent state $|\beta\rangle$ of a quantum parallel LC oscillator is a minimum uncertainty state. This means it has an uncertainty principle relationship between Q and ϕ that is the minimum possible according to quantum mechanics, i.e.,

    $$\sigma_\phi\sigma_Q = \frac{\hbar}{2}, \tag{7.171}$$

    where σ_X is the standard deviation of observable X.

Chapter 8

Perturbation Theory

Previously, we saw that driving a quantum LC oscillator with a "classical" source led to the quantum LC oscillator being placed into a highly classical state, namely, a coherent state. Considering this, we noted that if we wish to control quantum systems in such a way that they exhibit highly non-classical behavior that may be advantageous for designing new technologies we would need to search for a physical system other than a quantum LC oscillator (or similar physical analogues). Before discussing details about the other kinds of physical systems we will often need to utilize in designing revolutionary quantum technologies, it will be helpful to gain a better understanding of how quantum systems "transition" between different energy eigenstates. To establish an intuitive view of this process, it is common to utilize what is often referred to in quantum mechanics as *time-dependent perturbation theory.* However, before discussing this, we will discuss *time-independent perturbation theory* to understand some foundational concepts of perturbation theory. This section can be omitted if desired, as Section 8.2 is independent of Section 8.1.

8.1 Time-Independent Perturbation Theory

Time-independent perturbation theory is a useful technique for successively generating refined approximations to the solution of the time-independent Schrödinger equation when the Hamiltonian slightly differs from one that we have already solved. As a simple example, we can imagine needing to solve the time-independent Schrödinger equation for an infinitely-deep potential well that has a small "bump" in the potential energy within the potential well rather than just having a flat well, as shown in Fig. 8.1. The basic idea of time-independent perturbation theory is to try and develop an algorithmic process for using solutions we have already developed for an "unperturbed problem" to try and solve this new problem.

To begin illustrating the process, let's assume that we have already solved the time-independent Schrödinger equation

$$\hat{H}^{(0)}|\psi_n^{(0)}\rangle = E_n^{(0)}|\psi_n^{(0)}\rangle, \tag{8.1}$$

where the superscript (0) is used to help denote that this is the unperturbed problem. We now wish to seek the solution of the new problem

$$\hat{H}|\psi_n\rangle = E_n|\psi_n\rangle, \tag{8.2}$$

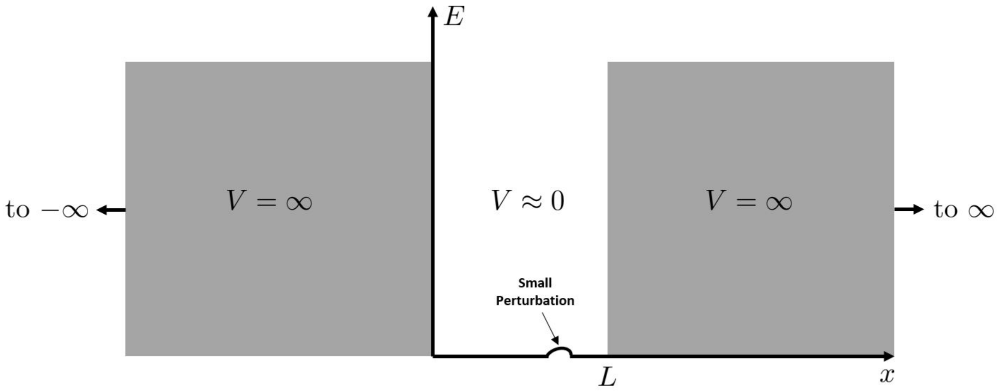


Figure 8.1: Infinitely-deep potential well with a small perturbation in the potential energy of the well.

where we can write the new Hamiltonian as

$$\hat{H} = \hat{H}^{(0)} + \lambda \hat{H}'. \tag{8.3}$$

We will often refer to $\hat{H}'$ as the *perturbing Hamiltonian.* The parameter λ is simply a number that we will use to help with bookkeeping in developing our perturbation theory solution. In reality, we will eventually set $\lambda = 1$ to arrive at our actual solution, but in the meantime it is helpful to include it. The reason for this is that we are going to use as an *ansatz solution* to the perturbed problem a power series in λ as

$$|\psi_n\rangle = |\psi_n^{(0)}\rangle + \lambda|\psi_n^{(1)}\rangle + \lambda^2|\psi_n^{(2)}\rangle + \ldots, \tag{8.4}$$

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \ldots. \tag{8.5}$$

Within this power series, we would refer to $E_n^{(1)}$ as the *first-order correction* to the energy and $E_n^{(2)}$ as the *second-order correction* to the energy. A similar terminology applies to the eigenfunctions $|\psi_n^{(1)}\rangle$ and $|\psi_n^{(2)}\rangle$ as well.

We now go about solving for our different corrections by substituting our power series expansions of (8.4) and (8.5) into (8.2). Our result can be originally written as

$$\begin{aligned}\left(\hat{H}^{(0)} + \lambda\hat{H}'\right)&\left[|\psi_n^{(0)}\rangle + \lambda|\psi_n^{(1)}\rangle + \lambda^2|\psi_n^{(2)}\rangle + \ldots\right] \\ &= \left(E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \ldots\right)\left[|\psi_n^{(0)}\rangle + \lambda|\psi_n^{(1)}\rangle + \lambda^2|\psi_n^{(2)}\rangle + \ldots\right].\end{aligned} \tag{8.6}$$

This is a bit messy at this point, but our strategy now is to group our terms according to the power of the bookkeeping parameter λ that terms are multiplied by. This gives us

$$\begin{aligned}\hat{H}^{(0)}|\psi_n^{(0)}\rangle + \lambda\left(\hat{H}^{(0)}|\psi_n^{(1)}\rangle + \hat{H}'|\psi_n^{(0)}\rangle\right) + \lambda^2\left(\hat{H}^{(0)}|\psi_n^{(2)}\rangle + \hat{H}'|\psi_n^{(1)}\rangle\right) + \ldots = E_n^{(0)}|\psi_n^{(0)}\rangle \\ + \lambda\left(E_n^{(0)}|\psi_n^{(1)}\rangle + E_n^{(1)}|\psi_n^{(0)}\rangle\right) + \lambda^2\left(E_n^{(0)}|\psi_n^{(2)}\rangle + E_n^{(1)}|\psi_n^{(1)}\rangle + E_n^{(2)}|\psi_n^{(0)}\rangle\right) + \ldots.\end{aligned} \tag{8.7}$$

The reason for arranging terms in this way is that we now see that on each side of our equation we have a power series in terms of λ. It is a general mathematical fact that power series expansions are *unique*, which in this case means that because our two power series equal each other we actually have the stronger property that each *coefficient* of each order of the power series equal one another.

We can now go order-by-order in λ and see what our new equations tell us. If we start at the zeroth-order, we find that

$$\hat{H}^{(0)}|\psi_n^{(0)}\rangle = E_n^{(0)}|\psi_n^{(0)}\rangle, \tag{8.8}$$

which is just our unperturbed equation, and so does not provide us with any new information. Our first-order equation *is* new, and takes the form

$$\hat{H}^{(0)}|\psi_n^{(1)}\rangle + \hat{H}'|\psi_n^{(0)}\rangle = E_n^{(0)}|\psi_n^{(1)}\rangle + E_n^{(1)}|\psi_n^{(0)}\rangle. \tag{8.9}$$

We can now be clever to isolate our first-order correction to the energy by taking the inner product of this equation with $|\psi_n^{(0)}\rangle$. This gives us

$$\langle\psi_n^{(0)}|\hat{H}^{(0)}|\psi_n^{(1)}\rangle + \langle\psi_n^{(0)}|\hat{H}'|\psi_n^{(0)}\rangle = E_n^{(0)}\langle\psi_n^{(0)}|\psi_n^{(1)}\rangle + E_n^{(1)}\langle\psi_n^{(0)}|\psi_n^{(0)}\rangle. \tag{8.10}$$

We can simplify this by noting that because $\hat{H}^{(0)}$ is Hermitian, we can actually apply it to the "left" in the first term in (8.10). The result is $\langle\psi_n^{(0)}|\hat{H}^{(0)} = E_n^{(0)}\langle\psi_n^{(0)}|$, so that our equation becomes

$$E_n^{(0)}\langle\psi_n^{(0)}|\psi_n^{(1)}\rangle + \langle\psi_n^{(0)}|\hat{H}'|\psi_n^{(0)}\rangle = E_n^{(0)}\langle\psi_n^{(0)}|\psi_n^{(1)}\rangle + E_n^{(1)}. \tag{8.11}$$

Canceling identical terms, we find that our first-order correction to the energy is

$$E_n^{(1)} = \langle\psi_n^{(0)}|\hat{H}'|\psi_n^{(0)}\rangle. \tag{8.12}$$

Finding the first-order correction to the wavefunction $|\psi_n^{(1)}\rangle$ is a little more involved. To begin, we rewrite our first-order equation as

$$\left(\hat{H}^{(0)} - E_n^{(0)}\right)|\psi_n^{(1)}\rangle = -\left(\hat{H}' - E_n^{(1)}\right)|\psi_n^{(0)}\rangle. \tag{8.13}$$

Since we have just found out how to calculate $E_n^{(1)}$, we can assume that this is now a known parameter so that we can consider the entire right-hand side of this equation to simply be some known function. As a result, we see that this becomes an inhomogeneous differential equation for our first-order correction $|\psi_n^{(1)}\rangle$. We can go about solving for $|\psi_n^{(1)}\rangle$ by first expanding it in terms of our complete set of functions from the unperturbed problem, i.e., the $|\psi_n^{(0)}\rangle$. This gives us

$$|\psi_n^{(1)}\rangle = \sum_m c_{n,m}^{(1)}|\psi_m^{(0)}\rangle, \tag{8.14}$$

where $c_{n,m}^{(1)}$ is the mth expansion coefficient for our first-order correction to the nth eigenstate. We have now reduced the solution of the inhomogeneous differential equation into solving for the expansion coefficients $c_{n,m}^{(1)}$. We can now substitute this expansion into (8.13) to get

$$\sum_m \left(\hat{H}^{(0)} - E_n^{(0)}\right) c_{n,m}^{(1)} |\psi_m^{(0)}\rangle = -\left(\hat{H}' - E_n^{(1)}\right) |\psi_n^{(0)}\rangle. \tag{8.15}$$

Recognizing that the $|\psi_m^{(0)}\rangle$ can have the unperturbed Hamiltonian easily applied to them, we find that

$$\sum_m \left(E_m^{(0)} - E_n^{(0)}\right) c_{n,m}^{(1)} |\psi_m^{(0)}\rangle = -\left(\hat{H}' - E_n^{(1)}\right) |\psi_n^{(0)}\rangle. \tag{8.16}$$

We now take the inner product with $|\psi_l^{(0)}\rangle$ to get

$$\sum_m \left(E_m^{(0)} - E_n^{(0)}\right) c_{n,m}^{(1)} \langle\psi_l^{(0)}|\psi_m^{(0)}\rangle = -\langle\psi_l^{(0)}|\hat{H}'|\psi_n^{(0)}\rangle + E_n^{(1)}\langle\psi_l^{(0)}|\psi_n^{(0)}\rangle. \tag{8.17}$$

There are now a few special cases that we need to consider more carefully. The first is what happens when $l = n$. In this case, the left-hand side of (8.17) becomes 0 and we learn no information for solving for one of our expansion coefficients. The other special case is if $m = n$. This likewise gives the left-hand side of (8.17) to be 0, so we have no constraint to apply to finding the expansion coefficient $c_{n,n}^{(1)}$. It turns out that this allows us to freely choose $c_{n,n}^{(1)}$ to be whatever we believe will help simplify our analysis. In most cases, it is best to simply choose $c_{n,n}^{(1)} = 0$ to keep the algebra easier, which also has the effect of making the first-order correction $|\psi_n^{(1)}\rangle$ to be orthogonal to the unperturbed nth eigenstate $|\psi_n^{(0)}\rangle$.

Now, considering these simplifications we can restrict our expansion to only cover values of $m \neq n$. We can return to (8.17) and see that we can simplify it to be

$$\left(E_l^{(0)} - E_n^{(0)}\right) c_{n,l}^{(1)} = -\langle\psi_l^{(0)}|\hat{H}'|\psi_n^{(0)}\rangle. \tag{8.18}$$

We can now isolate $c_{n,l}^{(1)}$ to find that

$$c_{n,l}^{(1)} = \frac{\langle\psi_l^{(0)}|\hat{H}'|\psi_n^{(0)}\rangle}{E_n^{(0)} - E_l^{(0)}}. \tag{8.19}$$

Returning to our original expansion of $|\psi_n^{(1)}\rangle$ in (8.14), we see that

$$|\psi_n^{(1)}\rangle = \sum_{m\neq n} \frac{\langle\psi_m^{(0)}|\hat{H}'|\psi_n^{(0)}\rangle}{E_n^{(0)} - E_m^{(0)}} |\psi_m^{(0)}\rangle. \tag{8.20}$$

It is important to note that we have implicitly assumed in writing (8.20) that we never have a situation where $E_n^{(0)} = E_m^{(0)}$ for $m \neq n$. Unfortunately, in practical problems it is often the case that the eigenvalues of a system have some *degeneracy* so that it *is* possible for

$E_n^{(0)} = E_m^{(0)}$ for $m \neq n$. In these situations one must use a different form of time-independent perturbation theory. To help differentiate between approaches, what we have discussed here is typically referred to as *nondegenerate time-independent perturbation theory*, while the approach that can handle degeneracy is referred to as *degenerate perturbation theory*.

Before moving on, we will briefly consider the result for the second-order correction to the energy $E_n^{(2)}$. Using similar methods to what we have illustrated here, it is possible to take the second-order equation from (8.7) and isolate $E_n^{(2)}$. We begin by taking the inner product with $|\psi_n^{(0)}\rangle$ to get

$$\langle\psi_n^{(0)}|\hat{H}^{(0)}|\psi_n^{(2)}\rangle + \langle\psi_n^{(0)}|\hat{H}'|\psi_n^{(1)}\rangle = E_n^{(0)}\langle\psi_n^{(0)}|\psi_n^{(2)}\rangle + E_n^{(1)}\langle\psi_n^{(0)}|\psi_n^{(1)}\rangle + E_n^{(2)}\langle\psi_n^{(0)}|\psi_n^{(0)}\rangle. \tag{8.21}$$

We can find that the first term on the left-hand side cancels the first term on the right-hand side, so after simplifying we find that

$$E_n^{(2)} = \langle\psi_n^{(0)}|\hat{H}'|\psi_n^{(1)}\rangle - E_n^{(1)}\langle\psi_n^{(0)}|\psi_n^{(1)}\rangle. \tag{8.22}$$

We can further simplify this by recognizing that by design we made $|\psi_n^{(1)}\rangle$ orthogonal to the unperturbed nth eigenstate $|\psi_n^{(0)}\rangle$, so that the second term on the right-hand side is 0. Expanding out our solution for $|\psi_n^{(1)}\rangle$ from (8.20), we see that

$$E_n^{(2)} = \langle\psi_n^{(0)}|\hat{H}'\left(\sum_{m\neq n}\frac{\langle\psi_m^{(0)}|\hat{H}'|\psi_n^{(0)}\rangle}{E_n^{(0)} - E_m^{(0)}}|\psi_m^{(0)}\rangle\right), \tag{8.23}$$

which we can rewrite more compactly as

$$E_n^{(2)} = \sum_{m\neq n}\frac{|\langle\psi_m^{(0)}|\hat{H}'|\psi_n^{(0)}\rangle|^2}{E_n^{(0)} - E_m^{(0)}}. \tag{8.24}$$

A few final remarks are in order for nondegenerate time-independent perturbation theory. First, it should hopefully now be more clear how this approach is a method of successive approximations. We started by solving for our first-order corrections in terms of the zeroth-order ones. We then proceeded by finding the second-order corrections using our first-order corrections. This pattern continues up to however high of an order is desired. In many cases, it is not useful to apply this kind of perturbation theory to high orders since the results do not converge very quickly. Instead, this approach is mostly useful for gaining some quick insight into how a small perturbation will approximately modify the energies and wavefunctions of a system. For instance, from formulas like (8.24), we see that energy levels that are far apart from one another will not contribute much to corrections between each other. Likewise, if we have two states that do not have a significant interaction through matrix elements like $\langle\psi_m^{(0)}|\hat{H}'|\psi_n^{(0)}\rangle$ we see that they will not cause much change to each other.

8.2 Time-Dependent Perturbation Theory

Previously, we saw that driving a quantum LC oscillator with a "classical" source led to the quantum LC oscillator being placed into a highly classical state, namely, a coherent state. Considering this, we noted that if we wish to control quantum systems in such a way that they exhibit highly non-classical behavior that may be advantageous for designing new technologies we would need to search for a physical system other than a quantum LC oscillator (or similar physical analogues). Before discussing details about the other kinds of physical systems we will often need to utilize in designing revolutionary quantum technologies, it will be helpful to gain a better understanding of how quantum systems "transition" between different energy eigenstates. To establish an intuitive view of this process, it is common to utilize what is often referred to in quantum mechanics as *time-dependent perturbation theory.*

To begin our discussion on time-dependent perturbation theory, recall that for the time-dependent Schrödinger equation we have that

$$\hat{H}_o|\Psi\rangle = i\hbar\frac{\partial}{\partial t}|\Psi\rangle, \tag{8.25}$$

and that we can solve this via separation of variables using eigenstates of $\hat{H}_o$ from the time-independent Schrödinger equation as

$$|\Psi\rangle = \sum_n c_n e^{-iE_n t/\hbar}|\psi_n\rangle. \tag{8.26}$$

Although we can find some interesting time-dependent behavior from (8.26), we note that our probability density $\langle\Psi|\Psi\rangle$ does not really exhibit time dependence because all the exponential factors in (8.26) cancel when we evaluate $\langle\Psi|\Psi\rangle$. Hence, if we wish to see *transitions* between eigenstates of our system (i.e., cause the c_n's to change as a function of time) we will need to introduce some additional time-dependent influence that makes $\hat{H}_o$ no longer fully characterize the system. If this "additional influence" is small, we can pursue a time-dependent perturbation theory solution to the problem to gain intuition into the physical processes involved.

More explicitly, we will consider a system where the total Hamiltonian can be written as

$$\hat{H} = \hat{H}_o + \hat{H}_p(t), \tag{8.27}$$

where $\hat{H}_p(t)$ is the time-dependent perturbation. Our time-dependent Schrödinger equation becomes

$$\hat{H}|\Psi\rangle = i\hbar\frac{\partial}{\partial t}|\Psi\rangle, \tag{8.28}$$

where we can now think of expanding $|\Psi\rangle$ in terms of the eigenstates of $\hat{H}_o$ as

$$|\Psi\rangle = \sum_n c_n(t) e^{-iE_n t/\hbar}|\psi_n\rangle. \tag{8.29}$$

Note that the only real change that has occurred in our expansion is that we now must write our expansion coefficients as being time dependent. Considering this, we could lump our $\exp[-iE_n t/\hbar]$ factor into the definition of $c_n(t)$. However, in many cases it is much better to keep this explicit time dependence separate from $c_n(t)$ because it will generally be found that $\exp[-iE_nt/\hbar]$ describes much of the time dependence for this particular state. In these situations, we will find that $\exp[-iE_nt/\hbar]$ oscillates very quickly in comparison to $c_n(t)$, so by explicitly accounting for this variation in our expressions we can simplify what goes "into" the solution of $c_n(t)$ and also gain more insight into our system.

We can now take our expansion in (8.29) and substitute this into the time-dependent Schrödinger equation. To simplify some of our notation, we will use a common shorthand for time derivatives that

$$\dot{c}_n = \frac{\partial c_n}{\partial t}, \tag{8.30}$$

so that (8.28) becomes (after evaluating the time derivative)

$$\sum_n c_n\bigg(\hat{H}_o + \hat{H}_p(t)\bigg)e^{-iE_nt/\hbar}|\psi_n\rangle = \sum_n \bigg(i\hbar\dot{c}_n + c_nE_n\bigg)e^{-iE_nt/\hbar}|\psi_n\rangle. \tag{8.31}$$

We can simplify this further by recognizing that on the left-hand side that

$$\hat{H}_o|\psi_n\rangle = E_n|\psi_n\rangle, \tag{8.32}$$

which then cancels the second term on the right-hand side of (8.31). Considering this, we have after some slight rearranging that

$$\sum_n i\hbar\dot{c}_n e^{-iE_nt/\hbar}|\psi_n\rangle = \sum_n c_n e^{-iE_nt/\hbar}\hat{H}_p(t)|\psi_n\rangle. \tag{8.33}$$

We can find an ordinary differential equation for a particular $\dot{c}_q$ by taking the inner product of (8.33) with $|\psi_q\rangle$ to get

$$i\hbar\dot{c}_q e^{-iE_qt/\hbar} = \sum_n c_n e^{-iE_nt/\hbar}\langle\psi_q|\hat{H}_p(t)|\psi_n\rangle. \tag{8.34}$$

To simplify our notation, we introduce the frequency

$$\omega_{qn} = (E_q - E_n)/\hbar \tag{8.35}$$

so that (8.34) can be rewritten as

$$i\hbar\dot{c}_q(t) = \sum_n c_n(t)e^{i\omega_{qn}t}\langle\psi_q|\hat{H}_p(t)|\psi_n\rangle \tag{8.36}$$

after moving $\exp[-iE_qt/\hbar]$ to the right-hand side.

Up to this point, we have made no approximations whatsoever. However, if the perturbation $\hat{H}_p(t)$ is "small", we can write the full Hamiltonian as $\hat{H} = \hat{H}_o + \lambda\hat{H}_p(t)$ where the parameter λ is going to be a bookkeeping tool that we can envision setting to 1 at the end of a calculation. In terms of λ, we can now propose as an ansatz that each expansion coefficient can be represented by a power series of the form

$$c_n = c_n^{(0)} + \lambda c_n^{(1)} + \lambda^2 c_n^{(2)} + \ldots. \tag{8.37}$$

Substituting this into (8.36) and recalling that we must now write $\hat{H}_p(t) \to \lambda \hat{H}_p(t)$, we get

$$\begin{aligned}&\left(\dot{c}_q^{(0)} + \lambda \dot{c}_q^{(1)} + \lambda^2 \dot{c}_q^{(2)} + \dots\right) \\ &\qquad = \frac{1}{i\hbar} \sum_n \left(c_n^{(0)} + \lambda c_n^{(1)} + \lambda^2 c_n^{(2)} + \dots\right) e^{i\omega_{qn}t} \langle \psi_q | \lambda \hat{H}_p(t) | \psi_n \rangle. \quad (8.38)\end{aligned}$$

It is a general mathematical fact that power series expansions are *unique*, which in this case means that because our two power series equal each other we actually have the stronger property that each *coefficient* of each order of the power series equal one another. Due to this, we can equate the coefficients of terms of different orders in λ within the power series on either side of our equation.

If we look at our zeroth-order equation, we find that

$$\dot{c}_q^{(0)}(t) = 0, \tag{8.39}$$

which simply tells us that the coefficient does not change in time, and so contains no new information. However, if we proceed to finding our first-order equation we see that it is

$$\dot{c}_q^{(1)}(t) = \frac{1}{i\hbar} \sum_n c_n^{(0)} e^{i\omega_{qn}t} \langle \psi_q | \hat{H}_p(t) | \psi_n \rangle. \tag{8.40}$$

It should be noted that from (8.39) we have found that all the $c_n^{(0)}$ are independent of time, so it is only the exponential and the perturbing Hamiltonian that actually vary as a function of time on the right-hand side of our equation. Since these are all known quantities, we can solve this equation using standard techniques for integrating an ordinary differential equation. If we stop with our perturbation approach here, we find that

$$c_q(t) \approx c_q^{(0)} + c_q^{(1)}(t). \tag{8.41}$$

We can use this to approximately determine the state of our system at time t and to then evaluate different quantities of interest. More generally, we can find that

$$\dot{c}_q^{(d+1)}(t) = \frac{1}{i\hbar} \sum_n c_n^{(d)}(t) e^{i\omega_{qn}t} \langle \psi_q | \hat{H}_p(t) | \psi_n \rangle \tag{8.42}$$

to calculate our perturbative solution to higher orders. We see from this result that time-dependent perturbation theory is an approach of *successive approximations*, with the result from the previous order being used in calculating the result at the next higher order in the expansion.

It should be noted that it is typically only the first one or two orders of time-dependent perturbation theory that are useful in many situations (after that point, it is usually better to invest time in forming a more accurate solution than continuing the perturbation theory). Despite the seeming simplicity of these results, it turns out that even first-order perturbation theory can be used to derive extremely useful results that are applicable in many situations. We will now consider the development of one such result to see how it helps us better understand the probability of transitions occurring between different states in a system.

8.2.1 Oscillating Perturbations

We will now consider one of the most important results of first-order time-dependent perturbation by considering a harmonically oscillating perturbation. As a simple example of a case where this would be relevant, we can consider again the Hamiltonian of a driven quantum LC oscillator, which was

$$\hat{H} = \frac{1}{2C}\hat{Q}^2 + \frac{1}{2L}\hat{\phi}^2 - \hat{\phi} I_b(t), \tag{8.43}$$

where $I_b(t)$ was a classical current source. We can then envision partitioning our Hamiltonian into the unperturbed part

$$\hat{H}_o = \frac{1}{2C}\hat{Q}^2 + \frac{1}{2L}\hat{\phi}^2 \tag{8.44}$$

and the time-dependent perturbation

$$\hat{H}_p(t) = -\hat{\phi} I_b(t). \tag{8.45}$$

We will specifically consider sinusoidal perturbations, so it will be convenient to write $I_b(t)$ as

$$I_b(t) = 2I_0 \cos(\omega t) = I_0\big(e^{-i\omega t} + e^{i\omega t}\big), \tag{8.46}$$

where we have used Euler's identity to rewrite the cosine function into complex exponentials, as this will help simplify later manipulations significantly.

Although we could proceed directly with this example, we will instead want to consider this problem in a more general form. Considering this, we will not utilize any special properties from a specific form of $\hat{H}_o$ like in (8.44). Likewise, we will write our perturbing Hamiltonian in a more generic form as

$$\hat{H}_p(t) = \begin{cases} \hat{H}_{po}\big(e^{-i\omega t} + e^{i\omega t}\big), & 0 < t < t_0 \\ 0, & \text{elsewhere}, \end{cases} \tag{8.47}$$

where we have assumed the perturbation is turned on for only a finite time t_0 and where $\hat{H}_{po}$ is the time-independent part of the perturbing Hamiltonian. In the example above, we would take $\hat{H}_{po} = -\hat{\phi} I_0$.

Now, we wish to establish a useful result that tells us the probability that a transition has occurred between two states after the perturbation has been turned off. To do this, we will assume that for $t < 0$ our system has been prepared in a specific energy eigenstate $|\psi_m\rangle$. As a result, all of our zeroth-order expansion coefficients for our time-dependent perturbation theory will be 0 except for $c_m^{(0)} = 1$. With these assumptions, we can simplify (8.40) to be

$$\dot{c}_q^{(1)}(t) = \frac{1}{i\hbar} c_m^{(0)} e^{i\omega_{qm} t} \langle \psi_q | \hat{H}_p(t) | \psi_m \rangle. \tag{8.48}$$

If we now wish to determine $c_q^{(1)}$ for times after the perturbation has been turned off, we can integrate (8.48) to get

$$c_q^{(1)}(t > t_0) = \frac{1}{i\hbar} \int_0^{t_0} e^{i\omega_{qm}\tau} \langle \psi_q | \hat{H}_p(\tau) | \psi_m \rangle d\tau. \tag{8.49}$$

Evaluating this integral in general situations would likely require a numerical method, but for our simple oscillating perturbation we can handle this analytically.

To begin, we can substitute in the specific form of our perturbation from (8.47) to get

$$c_q^{(1)}(t > t_0) = \frac{1}{i\hbar}\langle\psi_q|\hat{H}_{po}|\psi_m\rangle \int_0^{t_0} \left(e^{i(\omega_{qm}-\omega)\tau} + e^{i(\omega_{qm}+\omega)\tau}\right)d\tau. \tag{8.50}$$

Evaluating the integral then gives

$$c_q^{(1)}(t > t_0) = -\frac{1}{\hbar}\langle\psi_q|\hat{H}_{po}|\psi_m\rangle \left[\frac{e^{i(\omega_{qm}-\omega)t_0} - 1}{\omega_{qm} - \omega} + \frac{e^{i(\omega_{qm}+\omega)t_0} - 1}{\omega_{qm} + \omega}\right], \tag{8.51}$$

which is the correct answer, but is a bit "messy". We can tidy this expression up by factoring out part of the exponential factors in each term inside the square brackets so that the remaining part can be rewritten as a sine function using Euler's function. This kind of simplification happens quite frequently in electrical engineering, and is a good strategy to take note of for other situations. Here, it allows us to simplify our result to be

$$\begin{aligned} c_q^{(1)}(t > t_0) = \frac{t_0}{i\hbar}&\langle\psi_q|\hat{H}_{po}|\psi_m\rangle\times \\ &\left[e^{i(\omega_{qm}-\omega)t_0/2}\,\frac{\sin[(\omega_{qm}-\omega)t_0/2]}{(\omega_{qm}-\omega)t_0/2} + e^{i(\omega_{qm}+\omega)t_0/2}\,\frac{\sin[(\omega_{qm}+\omega)t_0/2]}{(\omega_{qm}+\omega)t_0/2}\right], \end{aligned} \tag{8.52}$$

where we have also factored a t_0 out of the expression so that the terms involving the sine functions take a special form. In particular, although the argument is somewhat complex, we see that these take the form of a sinc function, which is defined by

$$\mathrm{sinc}(x) = \frac{\sin x}{x}. \tag{8.53}$$

Considering this, we write our final expression for $c_q^{(1)}(t > t_0)$ as

$$\begin{aligned} c_q^{(1)}(t > t_0) = \frac{t_0}{i\hbar}&\langle\psi_q|\hat{H}_{po}|\psi_m\rangle\times \\ &\left(e^{i(\omega_{qm}-\omega)t_0/2}\,\mathrm{sinc}[(\omega_{qm}-\omega)t_0/2] + e^{i(\omega_{qm}+\omega)t_0/2}\,\mathrm{sinc}[(\omega_{qm}+\omega)t_0/2]\right). \end{aligned} \tag{8.54}$$

To unpack this expression a bit, it is helpful to recall the basic shape of a sinc function. One can take the limit as $x \to 0$ to find that sinc functions peak to a value of 1 at their origin. The result then falls off quickly as a function of x due to the denominator, however, the sine function does lead to oscillatory behavior as this happens. This is illustrated in Fig. 8.2, where see that after only a few periods the sinc function has reduced to a rather small value in comparison to its peak. We will go into more detail about this shortly, but if we take a quick look at (8.54) we see that a particular expansion coefficient will be heavily suppressed if t_0 is "long enough" relative to $\omega_{qm} \pm \omega$. This can be counteracted if our drive frequency ω is close to ω_{qm}, which suggests that to have an appreciable interaction between two energy eigenstates we need to make our drive frequency *resonant* with the frequency defined by the transition between the relevant energy eigenstates.

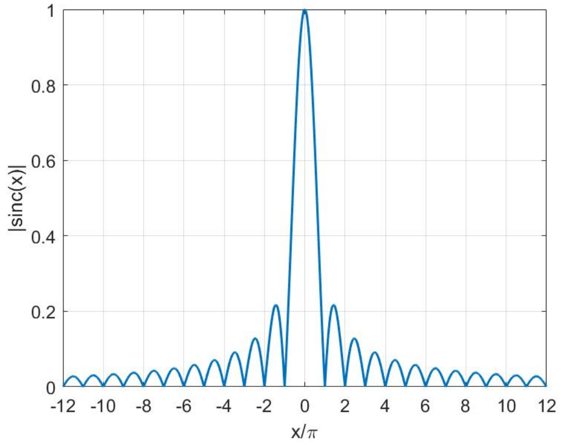


Figure 8.2: Absolute value of a sinc function over ±6 periods of oscillations. We see that the value drops off rapidly as x increases.

To see this idea in more detail, it is helpful to compute the probabilities of transitions between different energy eigenstates. To do this, we will first note that for $t > t_0$ we can write our complete quantum state to first order as

$$|\Psi\rangle \approx e^{-iE_m t/\hbar}|\psi_m\rangle + \sum_q c_q^{(1)}(t > t_0)e^{-iE_q t/\hbar}|\psi_q\rangle. \tag{8.55}$$

Recalling our earlier discussion of time-dependent perturbation theory, we recognize the first term in our superposition is the result if there were no perturbation and the remaining terms correspond to the effect of the perturbation to first order. Looking at this expression, we also can readily see a small "issue" with our state – it is no longer normalized! This is a common issue with perturbation theory, however, the system is still close to normalized because the perturbation is assumed to be small. The practical consequence of this is that when we evaluate "probabilities" with our state given in (8.55) they are no longer quite true probabilities because they will not all sum to 1. However, because the error is relatively small, we will still proceed with interpreting our results as probabilities. Along these lines, we can now see that the probability (to first order) that the system has transitioned from the initial state $|\psi_m\rangle$ to some other state $|\psi_q\rangle$ will be

$$P(q) \approx |c_q^{(1)}(t > t_0)|^2. \tag{8.56}$$

We can expand this using our expression for $c_q^{(1)}(t > t_0)$ given in (8.54) to find that

$$\begin{aligned} P(q) \approx \frac{t_0^2}{\hbar^2}|\langle\psi_q|\hat{H}_{po}|\psi_m\rangle|^2\bigg(&\mathrm{sinc}^2[(\omega_{qm}-\omega)t_0/2] + \mathrm{sinc}^2[(\omega_{qm}+\omega)t_0/2] \\ &+ 2\cos(\omega t_0)\mathrm{sinc}[(\omega_{qm}-\omega)t_0/2]\,\mathrm{sinc}[(\omega_{qm}+\omega)t_0/2]\bigg). \end{aligned} \tag{8.57}$$

We can simplify this expression a bit more by making a useful approximation. Namely, we recognize that if we have ω be "near" the transition frequency $|\omega_{qm}|$ and let t_0 be large enough the final term in (8.57) will have a minor impact on the overall probability, and can then be safely ignored. The reason for this is that if $\omega \approx |\omega_{qm}|$ then the two sinc functions will be spaced far apart from one another so that in the multiplication one of the sinc functions will always be very small (the large t_0 also has the effect of further "sharpening" the sinc functions as a function of frequency, which also supports minimizing the overlap of the functions). With this simplification, we have the probability of transition being

$$P(q) \approx \frac{t_0^2}{\hbar^2}|\langle\psi_q|\hat{H}_{po}|\psi_m\rangle|^2\bigg(\text{sinc}^2[(\omega_{qm}-\omega)t_0/2]+\text{sinc}^2[(\omega_{qm}+\omega)t_0/2]\bigg). \qquad (8.58)$$

We can now look at the different terms and see what kind of physics they describe. The first term to consider is $|\langle\psi_q|\hat{H}_{po}|\psi_m\rangle|^2$, which we see is related to the matrix element of the "static" part of the perturbation. If we recall our example of a current driving an LC oscillator, we see that this term will be proportional to I_0^2. Recalling basic circuit theory, we see that this is itself proportional to the power being supplied by the source. Hence, we get the intuitive behavior that if we drive our system with a higher intensity, transitions between a given set of energy eigenstates will be more likely to occur.

Although this is intuitive, the more important behavior is captured in our two sinc functions. In particular, we see that the first sinc function will be dominant if we set our drive frequency ω to be close to ω_{qm}. More explicitly, we have this term being dominant if $\omega \approx (E_q - E_m)/\hbar$. Since by convention we only consider driving frequencies like ω to be able to be positive numbers, we see that we can only have the condition $\omega \approx (E_q - E_m)/\hbar$ if $E_q > E_m$. Considering this, we see that the first sinc function in (8.58) is describing an *absorption* process that leads to the system transitioning to a higher energy state. This process is much more likely to occur if we match our drive frequency ω to the frequency defined by the transition ω_{qm}, which is a very important consideration that must be taken into account when attempting to control quantum systems, as we will discuss later in this course.

If the first sinc function describes absorption, what does the second one represent? We see that this term will only be large if $\omega_{qm} \approx -\omega$. Recalling that ω must be positive, we see that this condition can only occur if $E_q < E_m$. That is, this kind of transition involves moving from a higher energy state to a lower one. Although this is not explicitly shown in this analysis, this transition to a lower energy level must be accompanied by another physical process to allow the energy to "leave" the particular quantum system we are studying. In this case, we can recognize this as an *emission* process that leads to our system emitting a quanta of energy, with energy equal to $|\hbar\omega_{qm}|$. In the case of a quantized circuit, we would often refer to the emitted quanta of energy as a *photon* due to its similar characteristics to photons from the broader theory of quantum electrodynamics. This emission result is also extremely important, since it tells us that we can *increase* the occurrence of emission by driving our quantum system at the correct frequency. This physical process is often referred to as *stimulated emission*, which is a key component of why lasers work (in case you are not aware, the term laser is actually an acronym for "light amplification by stimulated emission of radiation").

We are now in a position that we can briefly reflect again on the failure of using a quantum LC oscillator with a classical drive to achieve highly quantum mechanical effects. As we will discuss later, it is often desirable to control quantum systems by selectively driving transitions between two *specific* energy levels of the system. We see that this is not possible with a quantum LC oscillator (and its physical analogues) because all of the energy levels are evenly spaced. As a result, once the system has been driven between two particular levels there is nothing "stopping it" from continuing to transition to higher and higher levels in the presence of a drive. To avoid this issue occurring, we will need to search for systems that have energy levels that are *not* evenly spaced. We will begin to discuss some general properties of these systems and look at an important specific example next week.

8.2.2 Fermi's Golden Rule

Although the analysis leading to (8.58) was very useful for gaining physical insight into important absorption and emission processes, it does have some drawbacks from a practical use standpoint due to the various "hand-waving" explanations we have relied on. For instance, our analysis required leaving t_0 on for long enough to sharpen the sinc functions appropriately, but also breaks down if we leave t_0 on for "too" long. It is possible to alleviate some of these caveats in specific cases, which turn out to happen quite frequently in practical situations. The end result of resolving some of these issues is what is often referred to as *Fermi's golden rule*, which is a very useful tool for estimating transition rates in quantum systems. Although Fermi's golden rule is still an approximate result, it is found to apply to many practical systems due to other effects not accounted for in our analysis. These other effects help to continually "reset" parameters around the approximations made in our analysis, so that the approximations that are made rarely become inaccurate.

The particular situation that can resolve some of these issues is if instead of only having a single transition with energy difference $\hbar\omega_{qm}$ to interact with there actually exist a dense set of such transitions all near the same frequency. This happens frequently in optical absorption problems, but can also occur for specific types of circuits as well. In particular, this type of dense level structure would be applicable to *transmission lines* that are designed to guide electromagnetic waves through systems. In this situation, we could imagine a quantum LC oscillator (or some other kind of quantum circuit) being coupled to the transmission line. The number of states that a photon in the LC oscillator could be emitted into the transmission line are so dense that they effectively form a *continuum* of states, rather than a discrete set like we have been considering up to this point.

For this situation, we can now think of adding up the probabilities of all transitions in a "neighborhood" of a particular frequency to determine the total transition probability. That is, if we only focus on the absorption process from the perspective of the dense set of states, we have the total probability being

$$P_{\text{tot}} \approx \sum_{q \text{ s.t. } \omega_{qm} \approx \omega} |c_q^{(1)}|^2. \tag{8.59}$$

In problems dealing with a nearly continuous set of states, it is a common strategy in quantum mechanics to approximate a summation over the dense set as an integration over a continuous set of states because integrations are generally much easier to deal with mathematically than abstract summations. To perform this conversion, we must introduce what is often referred to as a *density of states*. In essence, the density of states helps us convert from a summation index that always increments by a value of 1 to an integral measure that captures the actual change in a quantity of interest as you increment between “distinct states”. This is a somewhat nuanced subject that we will not need to explore in complete depth for this course, but if you find yourself needing to understand it more fully in the future it is a subject that is often covered in quantum mechanics textbooks, such as [8, 9, 25]. For now, we will simply say that we can transition from the summation in (8.59) to the integral

$$P_{\text{tot}} \approx \frac{t_0^2}{\hbar^2}|\langle\psi_q|\hat{H}_{po}|\psi_m\rangle|^2 \int \text{sinc}^2[(\omega_{qm}-\omega)t_0/2]g(\hbar\omega_{qm})d(\hbar\omega_{qm}), \tag{8.60}$$

where $g(\hbar\omega_{qm})$ is the density of states and we have focused specifically on the absorption process by only including the single sinc function.

To evaluate this integral in closed form, it is assumed that because the sinc function is quite narrow and the density of states is assumed to vary slowly, we can approximately pull it out of our integral as a constant evaluated at the center of the sinc function. This gives us

$$P_{\text{tot}} \approx \frac{t_0^2}{\hbar^2}|\langle\psi_q|\hat{H}_{po}|\psi_m\rangle|^2 g(\hbar\omega) \int \text{sinc}^2[(\omega_{qm}-\omega)t_0/2]d(\hbar\omega_{qm}). \tag{8.61}$$

This integral of the sinc function can be now be completed in closed form as

$$\int \text{sinc}^2(x)dx = \pi. \tag{8.62}$$

Using this result, we find that

$$P_{\text{tot}} \approx \frac{2\pi t_0}{\hbar}|\langle\psi_q|\hat{H}_{po}|\psi_m\rangle|^2\, g(\hbar\omega). \tag{8.63}$$

We can now define a *transition rate* by dividing by our somewhat arbitrary interaction time t_0 to get

$$W = \frac{2\pi}{\hbar}|\langle\psi_q|\hat{H}_{po}|\psi_m\rangle|^2\, g(\hbar\omega), \tag{8.64}$$

which is known as Fermi's golden rule. This is an important result that is often used in engineering quantum systems, where the goal is typically to enhance or suppress transitions at specific frequencies. This can often be accomplished by engineering the environment around the quantum system being studied, such as placing a nanoantenna near an atom to enhance emission in a desired direction. The importance of Fermi's golden rule is it helps guide us as to what aspects of the environment will be able to impact the transition rate by inspecting the static perturbing Hamiltonian $\hat{H}_{po}$ for the specific situation being considered.

It is common to also see Fermi's golden rule written in terms of Dirac delta functions. In particular, the transition rate between specific states $|\psi_m\rangle$ and $|\psi_q\rangle$ is expressed as

$$w_{qm} = \frac{2\pi}{\hbar}|\langle\psi_q|\hat{H}_{po}|\psi_m\rangle|^2\delta(E_{qm}-\hbar\omega). \tag{8.65}$$

The total transition rate in the “neighborhood” of a particular transition is then found as

$$W = \int w_{qm} g(\hbar\omega_{qm})d(\hbar\omega_{qm}). \tag{8.66}$$

8.3 Practice Problems

1. Problem 6.3.2 from D. A. B. Miller. You may use a calculator rather than explicitly evaluating the necessary integrals.
    (a) Solve the problem as written in D. A. B. Miller using first-order perturbation theory. For a potential well of width 10 nm and $u = 10^{-5}$, compare the transition energy for the unperturbed and perturbed systems.
    (b) Use a finite difference computer program to solve the problem explicitly and compare the solution to the one found in (a) for a 10 nm width when $u = 10^{-5}$ and when $u = 10^{-4}$.
2. Problem 7.2.1(i) from D. A. B. Miller.
3. You are given an infinitely-deep potential well prepared in its ground state. You wish to transition this system into its third stationary state by using a classical control mechanism with amplitude $A(t)$ that you can design arbitrarily. Assuming the Hamiltonian for this system is given by

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x) + A(t)\hat{p}$$

and the amplitude of $A(t)$ is sufficiently small, describe a possible procedure to achieve the desired transition. You should justify your explanation with relevant results from time-dependent perturbation theory.

Chapter 9

Artificial Atoms

9.1 Introduction to Artificial Atoms

Previously, we mentioned that it is common in many quantum technologies to want to be able to selectively drive transitions between two specific energy levels in a system. The reason for this is that these two energy levels can then potentially be used as a *quantum bit (qubit)*. A qubit can be viewed as a fundamental "building block" of quantum information, much like a classical bit is a fundamental piece of classical information in digital computers. Qubits must obey the laws of quantum mechanics, which allows them to be manipulated in interesting ways that is not possible with our classical versions of bits. It is these new and exciting properties of qubits that are fundamental to many emerging technologies, like quantum communication systems, quantum computers, and quantum sensors. We will learn more about these technologies later in the course. Before this, we will need to learn more about what kinds of quantum systems can potentially act as qubits and how we can control them.

When we discussed time-dependent perturbation theory, we saw that when a classical system is "driving" a quantum one the probability of a transition happening between two energy levels was heavily-dependent on the applied classical drive being *resonant* with the frequency defined by the transition. More explicitly, we saw that we needed the classical angular frequency $\omega \approx (E_m - E_n)/\hbar$ to drive transitions between states $|\psi_m\rangle$ and $|\psi_n\rangle$. At the time, we noted that this meant that a quantum LC oscillator (and other physical analogues) would not allow us to selectively drive transitions between different energy levels because all of the energy levels are evenly spaced. Hence, we will now need to turn our attention to considering quantum systems that have unevenly-spaced energy levels.

We have actually already encountered an example of this behavior toward the beginning of the course. In particular, when we analyzed an infinitely-deep square potential well we found that the eigenvalues of the system were

$$E_n = \frac{\hbar^2}{2m}\left(\frac{n\pi}{L}\right)^2, \quad n \in \mathbb{Z}, \tag{9.1}$$

which are clearly not evenly spaced. This particular problem was somewhat artificial (i.e., it was one-dimensional and had an infinite potential well), but it does give us useful intuition that if we can confine an electron (or electrons) within a small space we may find the behavior we are looking for from a qubit. One approach in this direction that is of current technological interest are *quantum dots*, which are nanoscale semiconductor structures that tightly confine electron wavefunctions in three dimensions. The resulting systems have energy level spacings that correspond to optical wavelengths, which is a favorable quality since there already exist many optical technologies that can then be repurposed for use in interacting with quantum dots. Quantum dots have also found broader application due to their light emission properties, and have already been commercialized into televisions (namely, QLED TVs). These systems have been found to be so important that some of the pioneers received the 2023 Nobel Prize in Chemistry for their work.

Although quantum dots have some very interesting properties, they are rather complex solid state systems made from many constituent "parts". Can we have simpler systems that still provide us with the desired behavior of unevenly-spaced energy levels? The answer is a definitive "yes", with an excellent example being one of the simplest quantum systems that you can likely think of – a hydrogen atom. Hydrogen atoms only contain a single electron, and as a result can be analyzed quite readily using the time-independent Schrödinger equation we have used throughout this course, albeit this does require some mathematical "maneuvering" to account for the fact that the proton in the nucleus is also "there". More details on these manipulations can be found in introductory quantum mechanics books like [8,9]. Once these manipulations are done, we can eventually write the Schrödinger equation as

$$\left(-\frac{\hbar^2}{2\mu}\nabla^2 - \frac{q^2}{4\pi\epsilon_0 r} \right)\psi(r,\theta,\phi) = E\psi(r,\theta,\phi), \tag{9.2}$$

where μ is the *reduced mass* (which is an effective mass partially accounting for the presence of the proton), q is the charge of an electron, ϵ_0 is the vacuum permittivity, and we have written everything in spherical coordinates due to the spherical symmetry of the Coulomb potential that *binds* the electron to stay part of the atom. This eigenvalue problem can be solved analytically, although it requires learning more about solving partial differential equations than will be useful for us to consider within this course. The main takeaway is that the energy levels are unevenly-spaced, as is the case in any atom, and so atoms form an obvious and natural candidate to potentially act as qubits in modern quantum technologies. Due to this, atoms (as well as ions) will sometimes be referred to as *nature's qubit.* Along these lines, a system that has been designed/manufactured to provide properties "like" an atom will often be referred to as an *artificial atom.* A quantum dot is an excellent example of an artificial atom, and we will take a detailed look at more examples shortly.

Considering these points, it should hopefully not surprise you at this point to learn that there are an *extremely* wide range of physical systems that will lead to unevenly-spaced energy levels. However, not all of these systems will have the properties we desire of a "good" artificial atom or qubit. Hence, we will need to take a deeper look at what properties are desirable for a physical system to potentially fulfill this role for us in designing new quantum technologies. Due to the diverse range of technologies being developed, there is no single list of all criteria needed for a qubit or artificial atom to be "good". However, within the context of processing *quantum information* (i.e., information stored in qubits), *DiVincenzo's criteria* represent an agreed upon set of conditions that must be "checked" for a physical system to be considered a viable candidate as a qubit. The DiVincenzo's criteria are [14]:

1. A scalable physical system with well-characterized qubits.
2. The ability to initialize the state of the qubits to a simple fiducial state.
3. Long relevant decoherence times.
4. A "universal" set of quantum gates.
5. A qubit-specific measurement capability.

We will not be able to investigate all of these properties in significant detail in this course. However, we will now consider the process of making sure we have a "well-characterized" physical system through a specific example of superconducting circuit artificial atoms. It is important to emphasize that these same kinds of considerations must be taken for any physical system when considering whether it can form the basis of a qubit in a broader quantum technology context.

9.2 Josephson Junctions

Superconducting circuit quantum devices are currently one of the most popular and mature approaches being pursued to develop a wide array of quantum technologies. As suggested by the name, these systems are built from superconducting materials that are arranged in different electrical circuit topologies to achieve desirable effects. Fortunately, a detailed knowledge of superconductivity is not required to understand many of the engineering considerations involved in designing these devices. However, for brevity, we will not be covering any of the superconducting physics in any detail in this course. We will simply introduce certain simple equations as "facts" when needed and move on with our discussion. Those interested in a more detailed presentation of some of these concepts can consult [5, 23, 26] and the references contained within. Currently, the most essential superconducting circuit component in building superconducting circuit artificial atoms is a *Josephson junction*, which we will now begin learning about.

9.2.1 Basic Properties

Josephson junctions are widely studied systems in their own right, but for our purposes we will only consider a relatively simple version that is formed via a superconductor-insulator-superconductor "sandwich". By keeping the insulator very thin (usually around 1 nm thick), a nonlinear I-V relationship is achieved when considering the junction as a circuit element. A basic illustration of a Josephson junction is shown in Fig. 9.1.

The nonlinear I-V relationship arises because the insulative gap in the junction acts as a barrier to the flow of charge carriers in the superconductor, which are known as *Cooper pairs*. These can roughly be thought of as 2 electrons that have "joined" together to form an effective *quasiparticle* that can move around the superconductor. By keeping the insulating

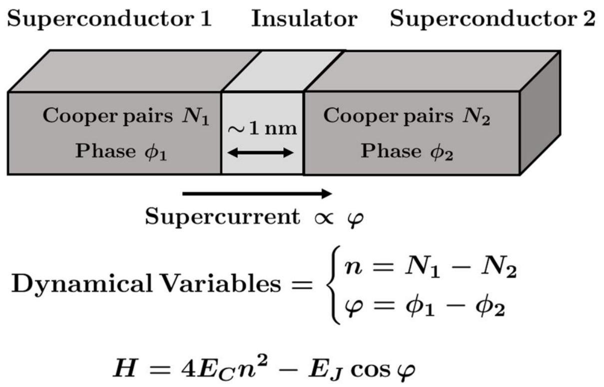


Figure 9.1: Illustration of a Josephson junction and the dynamical variables involved in expressing the Hamiltonian of the junction (image modified from [23]).

gap “thin”, a narrow potential barrier is produced that charge carriers can tunnel between in a manner similar in principle to the tunneling that occurs in a diode or transistor. This tunneling behavior is encoded at a macroscopic level in the *Josephson relations*, which can be used to derive the I-V relationship of the junction.

As with earlier circuit components we have considered, voltage and current are not suitable conjugate variables to describe the physics of a Josephson junction within a Hamiltonian mechanics framework. Instead, it is typical to use a dimensionless set of conjugate variables n and φ. Here, n is the Cooper pair density difference across the junction (effectively, the number of pairs that have tunneled through the junction) and φ is the *Josephson phase* (a parameter related to the phase difference in the wavefunctions in the two superconductors). Using these variables, the two Josephson relations are

$$I = I_c \sin\varphi, \tag{9.3}$$

$$\frac{\partial\varphi}{\partial t} = \frac{2q}{\hbar}V, \tag{9.4}$$

where I_c is known as the critical current that describes the maximum tunneling current possible (before certain assumptions break down) and q is the electron charge. Taking a closer look at (9.4), we see that the Josephson phase variable is closely related to the node flux we have been using in describing linear circuits. Considering this, we will think of φ as acting like a generalized position variable for the Josephson junction.

We can now find the I-V relationship of the Josephson junction by taking the time derivative of (9.3) and plugging in from (9.4) to get

$$\frac{\partial I}{\partial t} = \frac{\partial I}{\partial\varphi}\frac{\partial\varphi}{\partial t} = I_c\frac{2q}{\hbar}\cos(\varphi)V. \tag{9.5}$$

Considering that φ is a function of V, we see that this I-V relationship is that of a *nonlinear inductor*. From the perspective of a Hamiltonian description, the inductive energy of this system is characterized as

$$H_L = -E_J \cos\varphi, \tag{9.6}$$

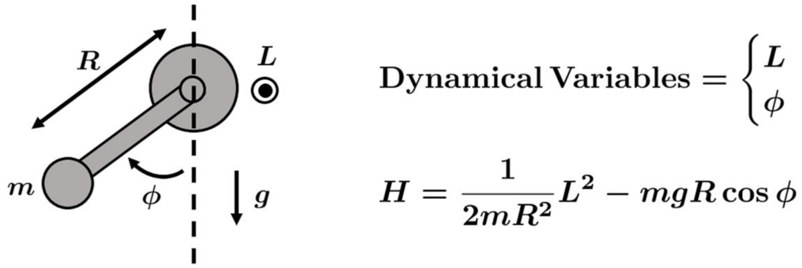


Figure 9.2: Illustration of a simple mechanical pendulum that exhibits the same dynamical behavior as a Josephson junction (image modified from [23]).

where $E_J = \hbar I_c/2q$ is the Josephson energy of the junction (which measures the energy of a Cooper pair tunneling through the junction).

Although the inductive energy is the "main" feature of the Josephson junction, it is not a complete description. In particular, due to the arrangement of two conductors near each other in forming the junction, there is also a capacitive energy in the system similar to a "parallel plate" capacitor. We can determine the capacitive energy for a total junction capacitance of C by noting that the total charge Q "stored" in the capacitance can be given by $2qn$, where the $2q$ is the charge of a single Cooper pair being "counted" by the dimensionless n. The charging energy of a single electron is $E_C = q^2/2C$, so the total capacitive energy $Q^2/2C$ can be included in the Hamiltonian as

$$H_C = 4E_C n^2. \tag{9.7}$$

Combining this with the earlier inductive energy given in (9.6), the full Hamiltonian of the Josephson junction can be given as

$$H = H_C + H_L = 4E_C n^2 - E_J \cos\varphi. \tag{9.8}$$

From a circuit theory perspective, this Hamiltonian corresponds to a linear capacitor in parallel with a nonlinear inductor. It is the nonlinearity of this inductance which is the key to making the corresponding quantized version of this system have unevenly-spaced energy levels, as desired for acting as a building block of an artificial atom.

This circuit theory perspective is useful, but you are likely not used to working with nonlinear inductors from your introductory circuits classes. Considering this, it can be useful to see if there is another system that behaves like our Josephson junction that is more physically intuitive. It turns out that such an analogy exists, and it is a simple mechanical pendulum with a fixed rotor of length R, as depicted in Fig. 9.2. We see that our Josephson phase variable φ naturally maps to the angular position ϕ of the rotor, and that the charge difference n corresponds to the angular momentum L of the rotor. When trying to understand the physics of the Josephson junction it can be useful to reflect back on this physical analogy, as we will consider later.

Before continuing, it is useful to discuss the relative scales of the energies in the Josephson junction given by E_C and E_J. For building next-generation quantum technologies, the Josephson junctions must often be kept physically small. As a result, the effective capacitance

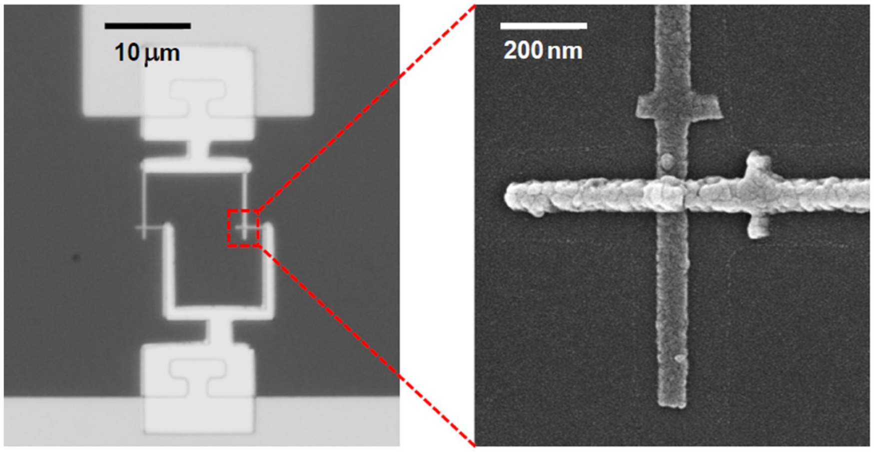


Figure 9.3: Physical layout of a (left) SQUID and (right) one of the Josephson junctions making up the SQUID (image modified from [23]).

of the junction is quite small, usually in the femtofarad to low picofarad range. Likewise, the tunneling energy that represents the effective inductance of the junction is also rather small, usually in the nanohenry to microhenry range. Importantly, because we are *designing* the Josephson junction we can tune these parameters so that the characteristic ratio E_J/E_C can be controlled to desirable operating regimes. Further, we can embed Josephson junctions into different circuit topologies to have even larger control over E_J/E_C. Overall, artificial atoms built with Josephson junctions have been realized with E_J/E_C ranging from much less than 0.1 to over 10^6.

This flexibility in design is one of the many factors that makes working with artificial atoms very attractive. If we use natural atoms/ions as qubits we do not have design control over their characteristics, we simply must work with whatever nature gave us. However, the tradeoffs between natural vs. artificial atoms are not this simple, and there are plenty of features of natural atoms that are superior to the artificial atoms that have been developed to date. As a result, work continues to be pursued on developing quantum technologies with both natural and artificial atoms.

9.2.2 Magnetic Flux-Tunable Josephson Junction

One additional advantage of artificial atoms is that their properties can often be tuned *in situ*. In the case of artificial atoms built with Josephson junctions, this is often accomplished by placing two Josephson junctions in parallel with each other. In this case, the two junctions form a superconducting loop, which is often referred to as a *superconducting quantum interference device (SQUID)*. A physical implementation of a SQUID within a superconducting artificial atom is shown in Fig. 9.3. The SQUID is useful because it makes the effective Josephson energy of the overall circuit tunable by applying an external magnetic flux through the loop formed by the SQUID. As we will see later, the Josephson energy helps establish the spacing between energy levels in the artificial atom so that tuning the effective E_J will allow the operating frequency of the artificial atom to be tuned.

To see why E_J becomes tunable, we must revisit the effective inductance Hamiltonian H_L

of the Josephson junction. With the second junction, this part of the Hamiltonian becomes

$$H_L = -E_J \cos\varphi_1 - E_J \cos\varphi_2, \tag{9.9}$$

where φ_i is the Josephson phase of the ith junction and we have assumed that the Josephson energy for both junctions are identical. In reality, the Josephson energies will be asymmetric either on purpose or due to manufacturing variability. However, junction asymmetries do not change the overall conceptual result, so we ignore them here for simplicity.

To proceed, we need another property of superconductors. Namely, that

$$\varphi_1 - \varphi_2 = 2\pi\ell + 2\pi\Phi/\Phi_0, \tag{9.10}$$

where ℓ is an integer, Φ is the total magnetic flux intersecting the loop formed by the SQUID, and $\Phi_0 = h/2q$ is the *superconducting flux quantum*. This equation tells us that the phase difference around the SQUID loop must be an integer multiple of 2π, which is necessary so that the wavefunctions in the superconductors remain single-valued. The final term in (9.10) reflects the fact that the applied magnetic flux will impact the phase differences in the superconductors. This is similar in concept to Faraday's law of induction, where a time-varying magnetic flux intersecting a loop can induce a voltage in the circuit.

Now, we can utilize standard trigonometric identities to rewrite (9.9) as

$$H_L = -2E_J \cos\left(\frac{\varphi_1 - \varphi_2}{2}\right)\cos\left(\frac{\varphi_1 + \varphi_2}{2}\right). \tag{9.11}$$

Defining a new effective Josephson phase variable as $\varphi = (\varphi_1 + \varphi_2)/2$ and using (9.10), we can simplify the above to be

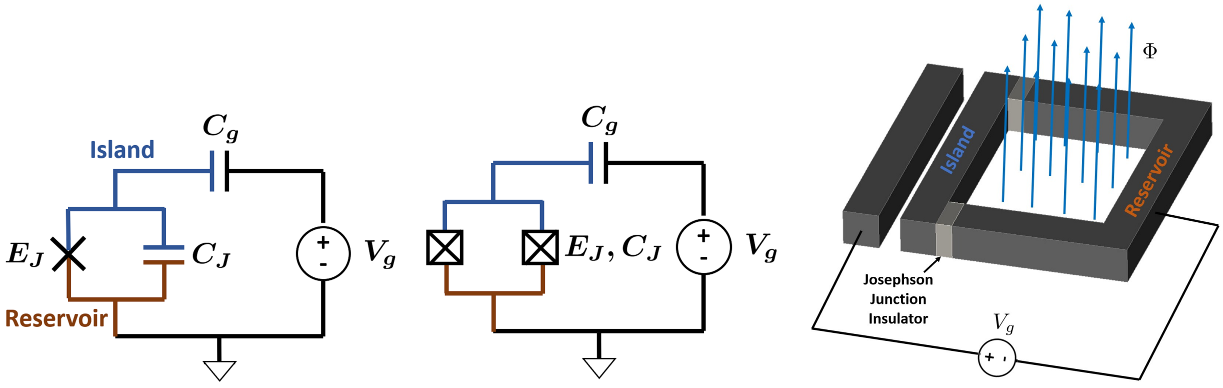
$$H_L = -E_{J\Phi} \cos\varphi, \tag{9.12}$$

where $E_{J\Phi} = 2E_J \cos(\pi\Phi/\Phi_0)$. From this, we see that the effective inductance of the SQUID has the same form as that of a single Josephson junction. The main change is that the effective Josephson energy $E_{J\Phi}$ is now tunable due to an applied magnetic flux.

Before continuing, it should also be mentioned that SQUIDs are a widely used technology in their own right. They have a long history of use in many areas of engineering. For instance, they can be used as precision magnetometers for many scientific and biomedical applications. They have also found use in metrology and standards applications, as well as in other sensors and electronic components.

9.3 Superconducting Circuit Artificial Atoms

To date, there have been a large number of artificial atoms designed using Josephson junctions embedded into superconducting circuits. Each artificial atom has its own set of characteristics that must be carefully analyzed to understand the advantages and disadvantages of a particular topology. Here, we will only focus on a particular subset of artificial atoms that are often referred to as *charge qubits*. We will consider a classic example in this area, known as a *Cooper pair box (CPB)* as a first example. Following this, we will discuss the transmon, which can be viewed as a kind of optimized CPB.



(a) (b) (c)

Figure 9.4: Circuit schematics for the (a) CPB and (b) split CPB. A pure Josephson junction tunneling element is represented schematically as an "X". Often, to simplify the schematics the small junction capacitance C_J is absorbed into the Josephson tunneling element symbol and is represented as a box with an "X" through it, as seen in (b). (c) Illustration of how conductors can be arranged to physically form a split CPB. (Images in (a) and (b) are modified from [23]).

9.3.1 Introduction to the Cooper Pair Box

The traditional CPB is formed by a Josephson junction that connects a superconducting "island" and "reservoir" [27]. For the CPB, the island is not directly connected to other circuitry, while the reservoir can be in contact with external circuit components (if desired). To control the operating point of the CPB system, a voltage source is capacitively coupled to the superconducting island. A circuit schematic of this setup is shown in Fig. 9.4(a). Although we will not focus on this in depth, a SQUID loop can be introduced into the CPB to make incorporate an additional degree of tunability, as shown in Fig. 9.4(b). This setup is sometimes referred to as a *split CPB*. An exaggerated illustration of how conductors can be arranged to physically form a split CPB is shown in Fig. 9.4(c).

We now need to determine how to incorporate the voltage source into our Hamiltonian description of the system. Although the circuit is changed somewhat compared to a simple Josephson junction, the resulting Hamiltonian will still have a very similar form to (9.8). Carrying out the algebraic manipulations to demonstrate this is tedious, so we will postpone this until a later time when we consider a more detailed analysis of quantizing this kind of system. For now, we simply state that the adjusted Hamiltonian can be written as

$$H = 4E_{C_\Sigma}(n - n_g)^2 - E_J \cos\varphi, \tag{9.13}$$

where E_{C_Σ} is the single electron charging energy for the total capacitance from the island to ground, which for this circuit is $C_\Sigma = C_J + C_g$. Further, we have that n_g is a dimensionless quantity typically called the *gate charge* or *offset charge* [5]. The gate charge incorporates the effect of the bias voltage V_g in adjusting the effective "operating point" of the voltage on the island. Explicitly, we have that $n_g = -C_g V_g / 2q$. To gain a little more intuition into this Hamiltonian, we can expand the quadratic and keep only the quantities that have n in them. We only need keep these quantities since any other values that are constant in n will not contribute to the equations of motion derived via Hamilton's equations. Performing this and rewriting terms into the forms with units that we have been considering elsewhere in our Hamiltonian treatments of circuits, we have that

$$4E_C(n-n_g)^2 \to \frac{1}{2C_\Sigma}Q^2 + \beta Q V_g, \tag{9.14}$$

where $\beta = C_g/C_\Sigma$ is a voltage divider that expresses how much of V_g is actually applied to the Josephson junction capacitance. Hence, we see that this potential energy has a form close to what we would expect for a voltage source biasing a circuit in terms of its effect on the charge in the system.

At this point, it is useful to quantize the system. As alluded to earlier, n and φ are canonically conjugate variables, although with a slight catch. In particular, these variables have been made dimensionless so that the canonical commutation relation is not directly applicable to them. Instead, we have for the quantum operators that

$$[\hat{n}, \hat{\varphi}] = -i. \tag{9.15}$$

In terms of these operators, the Hamiltonian operator becomes

$$\hat{H} = 4E_{C_\Sigma}(\hat{n}-n_g)^2 - E_J \cos\hat{\varphi}, \tag{9.16}$$

where it is important to note that n_g is still a real-valued classical variable.

For a typical CPB, $E_J/E_{C_\Sigma} \ll 1$, so we can see that the dominant term in (9.16) is the capacitive energy terms. Considering this, it can often be useful to expand the CPB Hamiltonian in terms of eigenstates of the charge operator. These eigenstates are often denoted as $|N\rangle$ and are somewhat similar to the Fock states we have encountered previously. In particular, $|N\rangle$ has eigenvalue N, which counts the number of Cooper pairs that have tunneled onto or off of the island in the CPB. It is part of the quantization of this system that when measuring this value, the result will always be an integer number of Cooper pairs. Now, in terms of charge states, we can write the capacitive part of the Hamiltonian as

$$\hat{H}_C = 4E_{C_\Sigma}(\hat{n}-n_g)^2 = 4E_{C_\Sigma}\sum_N (N-n_g)^2 |N\rangle\langle N|, \tag{9.17}$$

which can be derived by inserting resolutions of the identity operator in terms of charge states and utilizing their orthonormality properties. Although it requires more work to show, it is also possible to expand the inductive part of the Hamiltonian in terms of charge states. This gives

$$\hat{H}_L = -E_J\cos\hat{\varphi} = \frac{E_J}{2}\sum_N \Big(|N\rangle\langle N+1| + |N+1\rangle\langle N|\Big), \tag{9.18}$$

which clearly illuminates the role of this term in describing the tunneling of Cooper pairs onto or off of the island.

Within this basis, one may construct a matrix representation of the Hamiltonian and see that due to $E_J/E_{C_\Sigma} \ll 1$ the matrix is diagonally dominant. One important effect of this

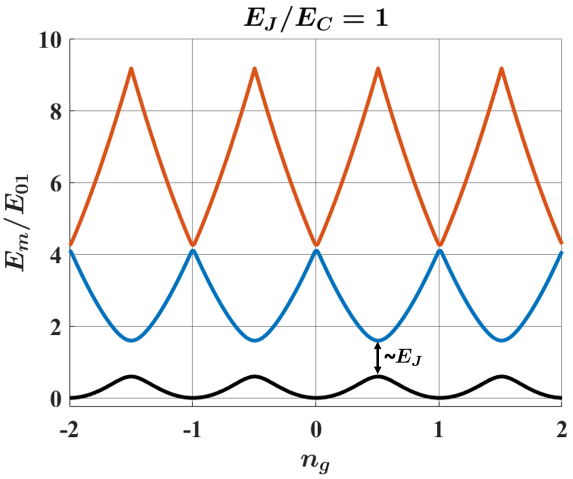


Figure 9.5: First three energy levels of the Hamiltonian given in (9.16) for $E_J/E_C = 1$. The energy levels are all normalized by the transition energy between the first two states evaluated at half-integer values of n_g (image from [23]).

is that if one wishes to find the eigenvalues and eigenvectors of the entire Hamiltonian then the results will be seen to be relatively minor perturbations to the existing charge states we were working with. Although it is instructive to go through this example, we will instead plot the first few energy levels as a function of n_g to further our discussion related to how the CPB could be used as a qubit. Considering this, we include the desired energy spectra in Fig. 9.5 for the case of $E_J/E_{C_\Sigma} = 1$. We consider a larger than normal E_J/E_{C_Σ} for a CPB in anticipation of discussing the transmon later and to make the plot visually easier to interpret.

The first observation to have about Fig. 9.5 is that the energy level spacing of the CPB is significantly dependent on the tunable control parameter n_g. For use as a qubit, the two main operating points we are interested in are when n_g is set to either integer or half-integer values.

Beginning with n_g having an integer value, it is seen that there is a large separation between the ground (i.e., the lowest energy eigenstate) and excited states. This is useful for initializing the system into a "simple fiducial state" (DiVincenzo's second criteria); namely, ensuring that the system is in the ground state prior to attempting any further operations. This operating point is useful for this purpose because quantum systems in higher energy states tend to "want" to decay back to their lower energy states, often by spontaneously emitting energy into another coupled system (e.g., emitting a photon into a circuit). If we wait long enough, we can be almost certain that our system will have emitted any energy keeping it in its higher excited state. However, just as our system can spontaneously emit energy, it can also spontaneously absorb energy from the environment (e.g., from thermal fluctuations in a material). The larger the energy spacing is between two levels are, the lower the probability is that our system will be able to absorb enough energy from the environment to spontaneously rise to a higher state. Hence, the large level separation at integer values of n_g provides us significant protection from this spontaneous absorption process, and as a result is a good operating point to allow our system to relax back into its ground state.

Although this operating point can establish a good ground state, it is not useful for qubit operations because the first two excited states have similar energies. Since this energy difference is relatively small, it becomes much easier for the system to spontaneously absorb enough energy from the environment to transition into the second excited state. Likewise, it is easy to emit this energy back into the environment to transition back and forth between the first two excited states. This causes a number of issues, but fundamentally we see that our system is no longer reliably in the ground or first excited state, so trying to think of it as a "qubit" that only has two important energy levels would not be appropriate.

Considering the energy level diagram of Fig. 9.5 again, it is seen that half-integer values of n_g are ideal for performing qubit operations. At these values, there is large separation between the energy levels of the first two excited states. Further, it is seen that the energy levels exhibit a high degree of *anharmonicity* here. That is, the energy levels are not evenly spaced like a harmonic oscillator or linear resonator. As a result of these two properties, the transition between the ground and first excited state can be selectively driven with fast classical drives (on the order one to tens of nanoseconds) without requiring significant filtering of the drive pulse. This allows for qubit operations to be performed very quickly, which is attractive for building many quantum technologies.

Another point to note about Fig. 9.5 is that at the half-integer values of n_g the energy spacing between the ground and first excited state is approximately equal to E_J. Considering this, if we use a SQUID in the CPB as opposed to a single junction we see that we will have direct control over the transition frequency of the qubit. This is a very advantageous property for many quantum technologies.

The final important point to note about Fig. 9.5 is that the two lowest energy levels are locally flat at half-integer values of n_g. This is important because it means that the operating frequency of the CPB will be less sensitive to *noise* in the n_g variable (often referred to as *charge noise*). For many quantum technologies, it is extremely important that the operating frequency of the qubit remains stable over time. If the operating frequency changes too much, it leads to various forms of *decoherence* in the qubit, which essentially makes the "phase" characterizing the state of our qubit become lost to the point that attempting to perform any more "operations" with the qubit will be effectively meaningless. The improved coherence of the CPB at half-integer values of n_g has lead to this operating point being known as the "sweet spot". Operating at this point resulted in CPBs with coherence times on the nanosecond to low microsecond scale [28]. Unfortunately, these systems were still far too sensitive to charge noise to work as desired for many quantum technologies. The transmon was introduced as a kind of "optimized CPB" to attempt to address this issue.

9.3.2 Introduction to the Transmon

The transmon qubit is best viewed as a CPB shunted by a capacitor that is large relative to the stray capacitance of the Josephson junction. Recalling that E_{C_Σ} is inversely proportional to capacitance, this results in the characteristic ratio E_J/E_{C_Σ} transitioning from $E_J/E_{C_\Sigma} \ll 1$ for a traditional CPB to $E_J/E_{C_\Sigma} \gg 1$ for a transmon. Originally, the large shunting capacitance was made using interdigital capacitors like that shown in Fig. 9.6. Now, the shunting capacitance is more commonly implemented as a large "plus" shape (or similar variants), as shown in Fig. 9.7, and is sometimes also referred to as an "Xmon qubit". This

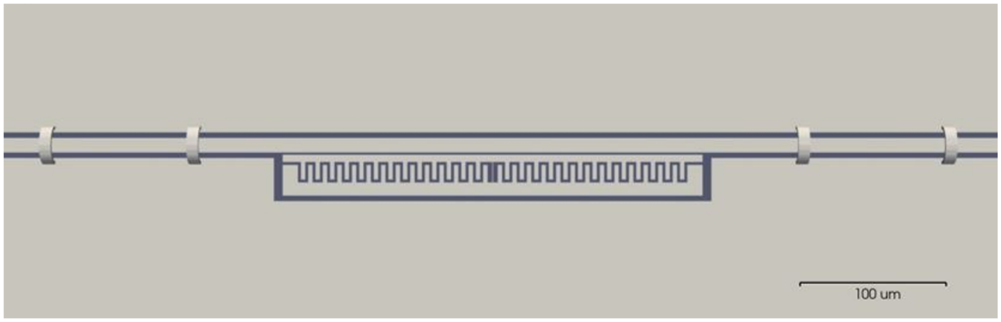


Figure 9.6: Example of a transmon qubit where the large capacitance is formed using an interdigital capacitor.

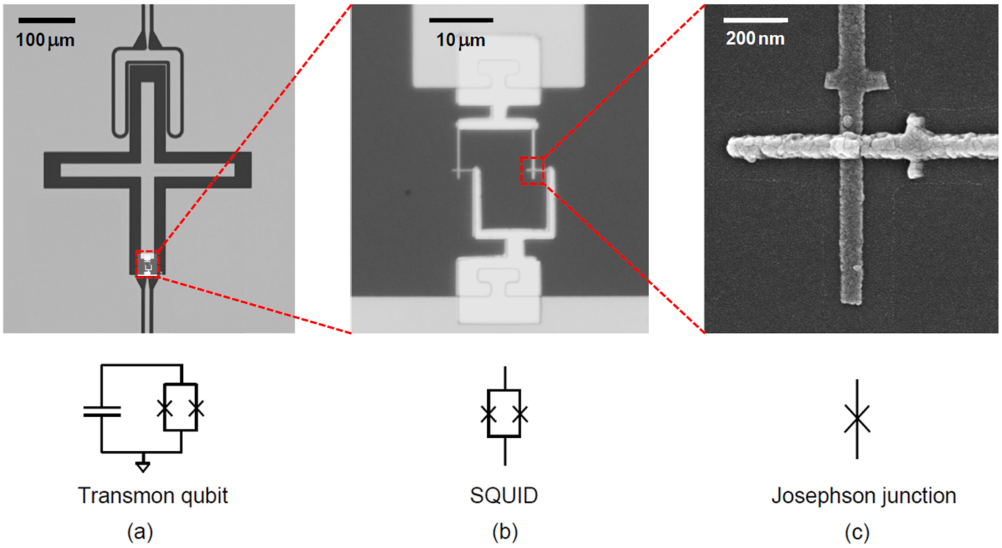
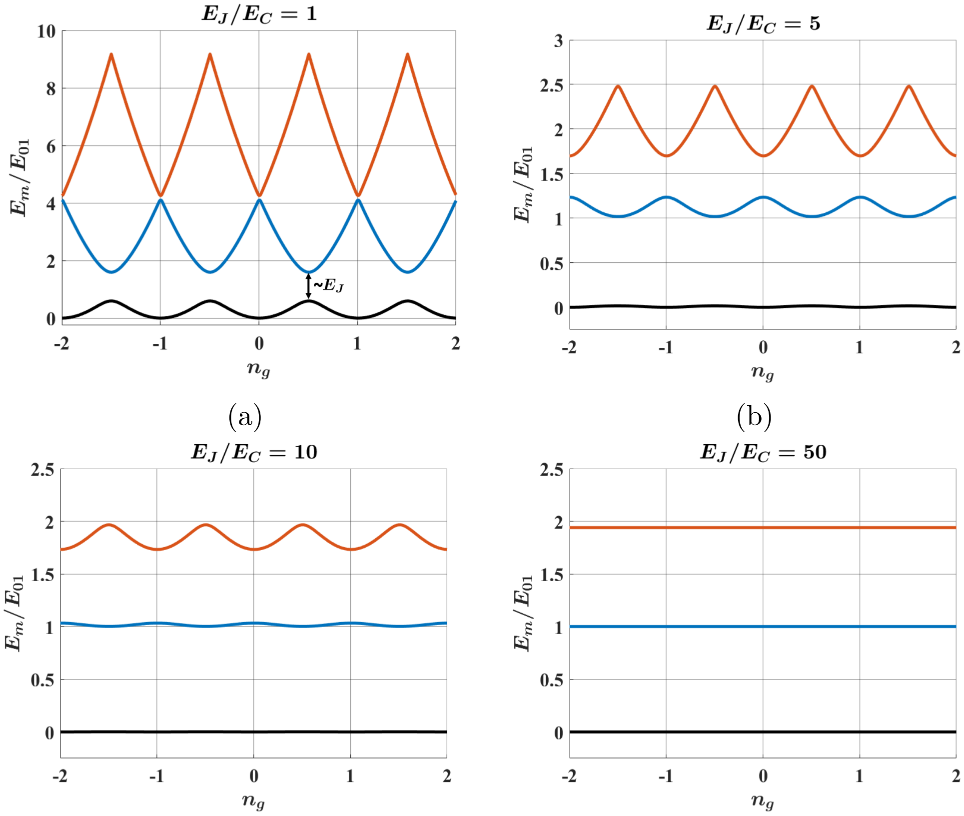


Figure 9.7: Image of a frequency-tunable transmon qubit: (a) the entire transmon consisting of a large capacitance (the "+" shape) in parallel with a SQUID to the ground, (b) a zoom-in of the SQUID, and (c) a single Josephson junction (images from [23]).

newer shape is typically favored as it provides better interconnectivity between different parts of a design. Although the physical implementation of the transmon qubit is different from more traditional CPB qubits, a Hamiltonian with the same form as (9.16) can still be used to describe its behavior. Similarly, a SQUID can be used to connect the different superconductors that make up the transmon (see Fig. 9.7), allowing the operating frequency of the transmon to be tuned via an applied magnetic flux.

It is instructive to see how the first few energy levels of (9.16) change as a function of E_J/E_{C_Σ} to understand the impact of adding the large shunting capacitance around the Josephson junction. This is shown for four different values of E_J/E_{C_Σ} in Fig. 9.8. These demonstrate the transition from a (close to) traditional CPB in Fig. 9.8(a) to a transmon in Fig. 9.8(d). From Fig. 9.8, it is seen that as E_J/E_{C_Σ} is increased, the energy levels flatten out as a function of n_g while simultaneously becoming more harmonic (closer to equally spaced like a quantum harmonic oscillator). Importantly, the sensitivity to charge noise reduces exponentially while the anharmonicity reduces by a weak power law. This leads to a qubit



(a) (b)

(c) (d)

Figure 9.8: First three energy levels of the qubit Hamiltonian given in (9.16) for values of E_J/E_C ranging from the CPB regime in (a) to the transmon regime in (d). Energy levels are normalized by the transition energy between the first two states evaluated at half-integer values of n_g (images from [23]).

that is insensitive to charge noise for practical purposes, but is still able to provide sufficient anharmonicity to effectively act as a qubit. For reference, the main operating frequencies of transmons range from a few GHz to 10 GHz, with anharmonicities of approximately 100 to 300 MHz. As a result, these qubits are controlled using microwave frequency transmission lines and related techniques, which represent a mature area of electrical engineering that has found use in many technologies over the years (e.g., communication systems, radars, etc., and of course... microwave ovens).

From the view of a mechanical equivalent, the transmon case is analogous to the rigid pendulum with an very long radius. This leads to the system having significant inertia, so that "noise" in the angular momentum won't be able to modify the position of the pendulum as much. For the transmon, this corresponds to insensitivity to charge noise in the n_g variable, making this only able to have a small role on the overall dynamics of the transmon. Due to this, an applied voltage source is no longer a useful control mechanism for a transmon, and so is omitted for practical designs.

Another effect that can be understood from the mechanical equivalent is that the pendulum stays close to its equilibrium point aligned with the gravitational field (i.e., ϕ is close

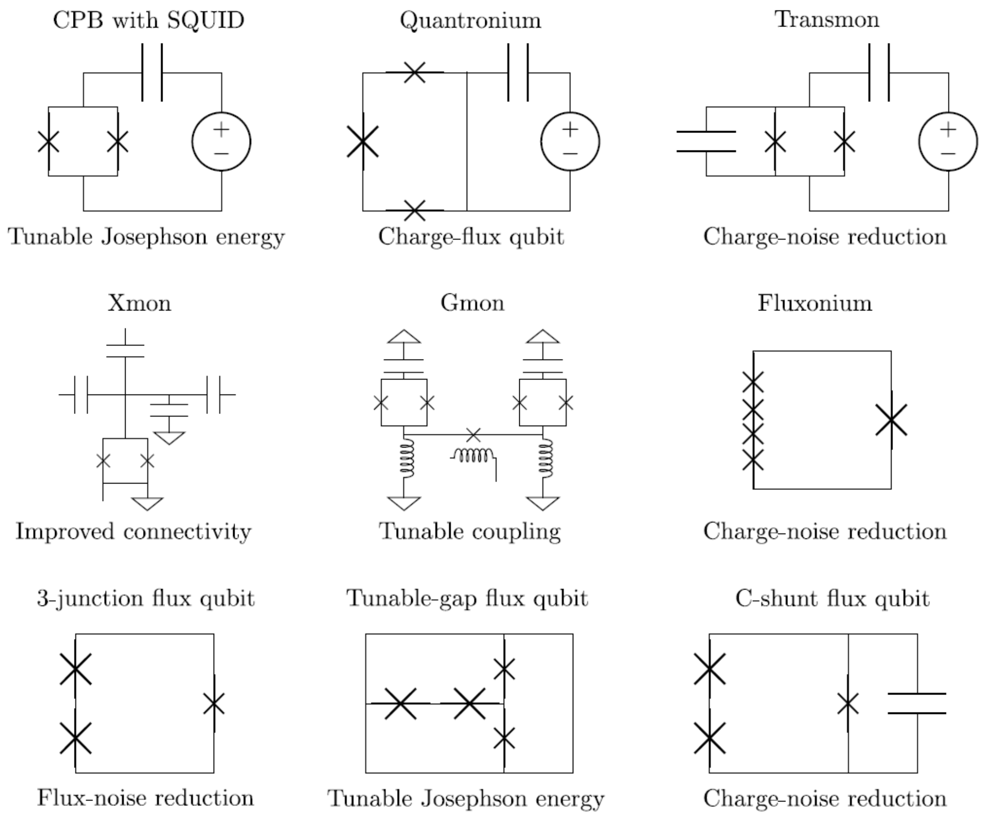


Figure 9.9: Examples of some of the main superconducting circuit artificial atoms that have been developed to date (image from [27]).

to 0). For this operating point, the dynamics correspond to a *weakly anharmonic oscillator*, which can usually be adequately described using a Taylor series including only a few terms for the cosine potential. Recalling that this Taylor series is

$$\cos x = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 + \dots, \tag{9.19}$$

it is common to see the cosine potential replaced with the x^2 and x^4 terms only to gain intuition for how the transmon will operate in certain complex scenarios (recall that the constant term is meaningless since it will not impact the dynamics of the system).

As a result of the transmon's resilience to charge noise (and other manufacturing and experimental advances), the coherence times of modern transmons are able to achieve values in the hundreds of microseconds range, representing an orders of magnitude improvement over the best CPB qubits. However, other sources of decoherence still persist, making the identification and reduction of decoherence still an important topic of research for this kind of artificial atom (broadly speaking, identifying and reducing decoherence is still an important topic of research for *all* artificial atoms). Despite these challenges, great accomplishments have been achieved with transmon qubits. For instance, these qubits have been central to many of the most successful quantum computing platforms built to date [16, 29].

9.3.3 Other Superconducting Circuit Artificial Atoms

Although transmons are one of the most popular artificial atoms currently being used, many other superconducting circuit artificial atoms have been developed. Each has their own unique advantages and disadvantages that must be taken into account when considering using them in a quantum technology. We will not be able to discuss all of these different artificial atoms, but to convey an idea of the variety we include a summary of some of the most significant ones in Fig. 9.9.

9.4 Two-Level Systems

Previously, we discussed two different superconducting circuit artificial atoms. We saw that they had very different properties and that their operating characteristics could often be further modified by changing design parameters or applying different kinds of biases. These particular artificial atoms were related to each other, so a relatively common language could be used to discuss their commonalities and differences. However, there are many other kinds of artificial or natural atoms that require substantially different physical equations to describe their properties. This variety provides us with a very rich "toolbox" for choosing/designing an (artificial) atom for a particular application, but it also becomes cumbersome when we wish to discuss certain general or "universal" characteristics of a particular kind of technology (e.g., general concepts of quantum computation irrespective of the exact hardware platform being used). The standard way to handle this difficulty in physics and engineering is to develop a kind of "abstraction" that we can use a unified language to discuss. It is then a matter of determining how a particular physical system maps into this generalized/abstracted framework to relate our "universal" considerations to a particular physical implementation. Whether it is better to use the abstract or actual physical description of a system depends on the context of what is being discussed or studied.

For atom-like systems, the most common abstract description used for them is that of a *two-level system*. A two-level system is the archetypal model of a "perfect" qubit – i.e., it contains exactly two levels that transitions can occur through via interactions with another external system. In reality, there are extremely few physical systems that actually only contain two levels. However, we typically consider a two-level description of an atom-like system to be reasonable if we can find two particular levels in the overall energy spectrum that are "well-separated" from other levels such that we can interact with them in a well-defined manner without worrying about the broader complexity of the system. Considering this, we will now discuss the common framework used for discussing two-level systems (and qubits, more generally).

Unfortunately, there are actually two more-or-less identical descriptions commonly used for two-level systems that still exist to this day. So, we will need to discuss both of these to be able to successfully interpret many expressions. Now, the first one we will discuss is the "spin representation" of a two-level system. We briefly discussed spin earlier in this course. As a reminder, spin is an intrinsic property of elementary particles that has many very important consequences on how said particles interact with one another and other systems. In the case of electrons, the value of spin can be either $+1/2$ or $-1/2$, and so these are referred to as "spin 1/2" particles.

Now, spin is best thought of as providing a kind of intrinsic "internal" angular momentum to the particle. This angular momentum can be oriented in a particular direction, and so it becomes necessary to think of spin as also having a geometric vector associated with it in 3D space. This angular momentum acts like a small magnetic dipole moment, and is the origin of many magnetic effects of materials. To actually describe the spin state of an electron, it is conventional to reference it to the z-axis of the coordinate system. We then write a general spin state $|s\rangle$ as

$$|s\rangle = c_{1/2}|1/2\rangle + c_{-1/2}|-1/2\rangle = c_{1/2}|\uparrow\rangle + c_{-1/2}|\downarrow\rangle = \begin{bmatrix} c_{1/2} \\ c_{-1/2} \end{bmatrix}, \tag{9.20}$$

where we have introduced a few commonly used notations. Most often, the $|\uparrow\rangle$ or $|\downarrow\rangle$ notation is used, where the arrows are meant to point "up" or "down" to correspond to the direction the spin angular momentum vector is oriented with respect to the z-axis. It is also very common to use a matrix notation when working with spin state vectors, where we see in the final equality of (9.20) that the "spin-up" state coefficient is the first coefficient in the vector.

From this description, we see that we have implied that we can express a general spin state that could be oriented in any direction in 3D space using only vectors "pointing" in the "up" or "down" directions. This may seem counter-intuitive at first, but it is important to remember that these are well-defined quantum states, and so our coefficients are complex-valued numbers subject to the normalization condition

$$\langle s|s\rangle = |c_{1/2}|^2 + |c_{-1/2}|^2 = 1. \tag{9.21}$$

It is these properties that help these two states describe any possible spin orientation. Actually determining what combinations of coefficients will cause the spin vector to be oriented in a particular direction is a little bit trickier to deduce. We will build up to this by first introducing the operators that help "measure" the spin angular momentum along a particular direction.

This is actually a somewhat nuanced issue to work through because spin angular momentum does not lend itself to a particular representation where it is easy to deduce the desired operator form. However, they can be determined through an analogy to "traditional" angular momentum operators in terms of the commutation relations that we require of the different operators. In terms of the more common "dimensionless" form, the commutation relations are

$$[\hat{\sigma}_x, \hat{\sigma}_y] = 2i\hat{\sigma}_z, \tag{9.22}$$

$$[\hat{\sigma}_y, \hat{\sigma}_z] = 2i\hat{\sigma}_x, \tag{9.23}$$

$$[\hat{\sigma}_z, \hat{\sigma}_x] = 2i\hat{\sigma}_y, \tag{9.24}$$

where the operator $\hat{\sigma}_x$ measures the spin along the x-axis, and similar for the other two operators.

There are many possible matrix representations of these operators that will be able to satisfy these commutation relations. However, the typical representation (which we are following) that uses the basis vectors

$$|{\uparrow}\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad |{\downarrow}\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{9.25}$$

corresponds to the following matrix representation:

$$\hat{\sigma}_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \tag{9.26}$$

$$\hat{\sigma}_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \tag{9.27}$$

$$\hat{\sigma}_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{9.28}$$

These matrices are often referred to as *Pauli spin matrices.* We can write these compactly as a vector that will be useful for "measuring" the spin angular momentum. In particular, we have

$$\hat{\boldsymbol{\sigma}} = \tilde{x}\hat{\sigma}_x + \tilde{y}\hat{\sigma}_y + \tilde{z}\hat{\sigma}_z, \tag{9.29}$$

where the $\tilde{x}$ notation is meant to denote a unit vector along the x-axis, with similar definitions for $\tilde{y}$ and $\tilde{z}$. We use this unconventional tilde notation to not confuse a unit vector with an operator.

Now, we can do some basic determination of coefficients to achieve a spin state oriented along a particular direction by looking at the eigenvectors and eigenvalues of the Pauli spin matrices. For instance, it is easy to see that the eigenvectors of $\hat{\sigma}_z$ are

$$|{\uparrow}\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad |{\downarrow}\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \tag{9.30}$$

with eigenvalues $+1$ and -1, respectively. Similarly, after a little matrix algebra, we can find that the eigenvectors of $\hat{\sigma}_x$ are

$$|\pm\rangle = \frac{1}{\sqrt{2}}\big(|{\uparrow}\rangle \pm |{\downarrow}\rangle\big) \tag{9.31}$$

and the eigenvalues are ± 1, respectively. From this, we see that the state

$$|+\rangle = \frac{1}{\sqrt{2}}\big(|{\uparrow}\rangle + |{\downarrow}\rangle\big) \tag{9.32}$$

will correspond to a spin angular momentum vector pointing in the $+x$-direction, and similar for $|-\rangle$ pointing in the $-x$-direction. We can also look at the eigenvectors of $\hat{\sigma}_y$, which are

$$|\pm i\rangle = \frac{1}{\sqrt{2}}\big(|{\uparrow}\rangle \pm i|{\downarrow}\rangle\big). \tag{9.33}$$

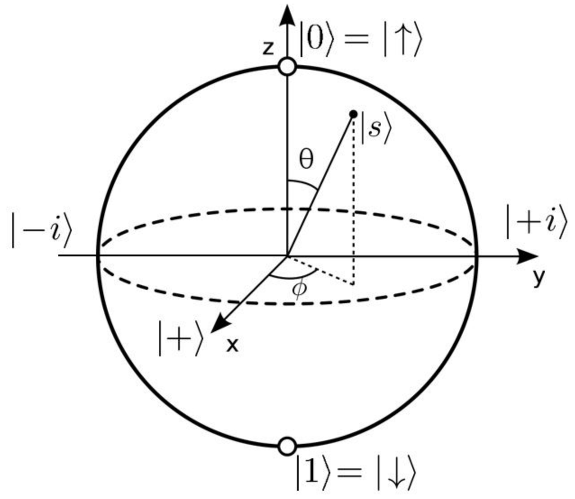


Figure 9.10: Illustration of the Bloch sphere that is commonly used to visualize a general linear combination of states in a two-level system (image modified from Wikipedia [30]).

These also have eigenvalues of ±1 and can be used to determine the combinations of $|\uparrow\rangle$ and $|\downarrow\rangle$ to have a spin angular momentum vector pointing along the $\pm y$-direction.

All of these different eigenvectors are useful in their own right, and play a special role in many quantum technologies (a point we will touch on more toward the end of the course). However, they only give us the answer to what combinations of $|\uparrow\rangle$ and $|\downarrow\rangle$ we need for a spin angular momentum vector to be pointing along one of our three coordinate axes. How do we represent a completely arbitrary direction? The most useful way to conceptualize this geometry is through the *Bloch sphere*, which is shown in Fig. 9.10. The Bloch sphere is a geometric representation of all of the valid linear combinations of $|\uparrow\rangle$ and $|\downarrow\rangle$. Because any state must be normalized, we know that all the linear combinations of $|\uparrow\rangle$ and $|\downarrow\rangle$ will have coefficients with the same "radius", equal to 1, as highlighted in (9.21). The shape that naturally describes this is a sphere, which we can then parameterize most easily in terms of spherical coordinates.

Since the radius is fixed to 1, the only parameters we need to consider are the angles θ and ϕ. Although it takes a little work to show it, we can choose a linear combination of $|\uparrow\rangle$ and $|\downarrow\rangle$ in terms of these two angles such that the angles point in the direction of spin angular momentum and also match the eigenvector definitions we had earlier in (9.30), (9.31), and (9.33). The combination that does this is

$$|s\rangle = \cos(\theta/2)|\uparrow\rangle + \sin(\theta/2)e^{i\phi}|\downarrow\rangle. \tag{9.34}$$

We can check to see if this definition matches what we were looking for in a few ways. One is to simply do "spot" checks along the different coordinate axes to make sure they match our eigenvectors given earlier. Another more general approach is to take the expectation value of $\hat{\boldsymbol{\sigma}}$ with the state $|s\rangle$. Performing the necessary algebra, we find that

$$\begin{aligned} \langle s|\hat{\boldsymbol{\sigma}}|s\rangle &= \tilde{x}\langle s|\hat{\sigma}_x|s\rangle + \tilde{y}\langle s|\hat{\sigma}_y|s\rangle + \tilde{z}\langle s|\hat{\sigma}_z|s\rangle \\ &= \tilde{x}\sin\theta\cos\phi + \tilde{y}\sin\theta\sin\phi + \tilde{z}\cos\theta, \end{aligned} \tag{9.35}$$

which does match our usual expressions for converting from a spherical coordinate representation of a point to a Cartesian representation. Hence, the angles θ and ϕ on the Bloch sphere map correctly to the angle the spin angular momentum is oriented toward.

Now, returning to our original task, when we use a "spin representation" for a two-level system we simply mean that we identify one of the two levels in the system with $|\uparrow\rangle$ and the other level with $|\downarrow\rangle$. This is done regardless of whether the states in our two-level system have anything to do with spin, and is simply done in this way due to historical reasons. Likewise, the Pauli spin matrices will still be used to compose different "operations" on the two-level system. For instance, any change to the state of the two-level system can be viewed as rotating our state vector on the Bloch sphere. These rotations can be described in terms of the action of the Pauli spin matrices on a state.

Our other representation for a two-level system can be thought of as a "qubit representation". In this case, we denote one level as the "0-state" and another as the "1-state". In Dirac notation, we express these as $|0\rangle$ and $|1\rangle$, respectively. These states are still typically mapped onto the Bloch sphere, and utilize the same linear combination as in (9.34) to interpret the location on the sphere. The particular correspondence is that $|\uparrow\rangle \rightarrow |0\rangle$ and $|\downarrow\rangle \rightarrow |1\rangle$. Unfortunately, this convention is not followed universally, and so some works will follow a notation of $|\uparrow\rangle \rightarrow |1\rangle$ and $|\downarrow\rangle \rightarrow |0\rangle$. As a result, one must be very careful in interpreting or combining expressions from various sources to ensure that a consistent choice has been made with notation.

Since it is common for a qubit to be formed by two different energy levels in an atom-like system, it is also common to refer to the $|0\rangle$ state as the "ground state" and the $|1\rangle$ state as the "excited state". This terminology also reflects that the $|0\rangle$ state is typically some kind of "stable" lower energy state of the system that would be occupied in situations where the atom-like system has not absorbed any additional energy from an external system.

In the "qubit representation", it can also sometimes be common to use a slightly different notation for the Pauli spin matrices. In particular, one may denote $\hat{\sigma}_x$ as $\hat{X}$, and likewise interpretations for $\hat{Y}$ and $\hat{Z}$. Regardless of the symbol used, the underlying matrices still match those defined in (9.26) to (9.28). This notation is very common in quantum computing, where one will talk about applying an "X gate" or a "Y gate", for instance.

9.5 Practice Problems

1. Using two "resolution of the identity operator" in terms of charge states, show how the capacitive energy part of the Cooper pair box Hamiltonian becomes

$$4E_{C_\Sigma}\big(\hat{n}-n_g\big)^2 \rightarrow 4E_{C_\Sigma}\sum_N \big(N-n_g\big)^2 |N\rangle\langle N|. \tag{9.36}$$

   Hint: To show this, you will need to put one identity operator before $\big(\hat{n}-n_g\big)^2$ and one identity operator after this. Don't forget that when you introduce these resolution of the identity operators that the dummy indices for the summations ***must*** be different!

2. The Hamiltonian of a Cooper pair box or a transmon can be expressed in the charge

basis as

$$\hat{H} = 4E_{C_\Sigma} \sum_N \left(N - n_g\right)^2 |N\rangle\langle N| - \frac{E_J}{2} \sum_N \Big(|N\rangle\langle N+1| + |N+1\rangle\langle N|\Big). \tag{9.37}$$

Using a basis of $N \in [-10, 10]$, write a computer program to form a matrix representation of this Hamiltonian operator.

(a) For $E_J/E_{C_\Sigma} = 0.1$ (Cooper pair box regime), compute the eigenvalues and eigenvectors of this Hamiltonian for $n_g = 0$ and $n_g = 3$. Plot the probability distribution of finding that N Cooper pairs have tunneled onto/off of the superconducting island for the first four eigenstates. That is, if the eigenstates are denoted by $|\psi_n\rangle$ you should plot $|\langle N|\psi_n\rangle|^2$ as a function of the discrete variable N. Comment on the effect of n_g on the eigenstates.

Note: Because N is a discrete variable, these plots typically look better when plotted as bar graphs.

(b) Repeat this analysis for a transmon with $E_J/E_{C_\Sigma} = 50$. Comment on how the eigenstates have changed in comparison to (a).

3. Implement a finite difference program to compute the eigenvalues and eigenvectors of the Cooper pair box/transmon Hamiltonian in the "phase basis". In this basis, the phase operator $\hat{\varphi}$ becomes a regular position variable φ and the charge operator $\hat{n}$ becomes similar to the momentum operator as $\hat{n} \to -i\frac{d}{d\varphi}$. The Hamiltonian is then

$$\hat{H} = -4E_{C_\Sigma} \frac{d^2}{d\varphi^2} - E_J \cos\varphi, \tag{9.38}$$

where we have set $n_g = 0$ for simplicity. The boundary condition applicable to wavefunctions for this system is $\psi(\varphi) = \psi(\varphi + 2\pi)$, which is also called a *periodic boundary condition*. To discretize this problem, consider the phase range $\varphi \in [-\pi, \pi]$ and use 201 discrete points over this interval. Plot the probability distribution of the first 3 eigenfunctions of this system in the phase basis for the following cases.

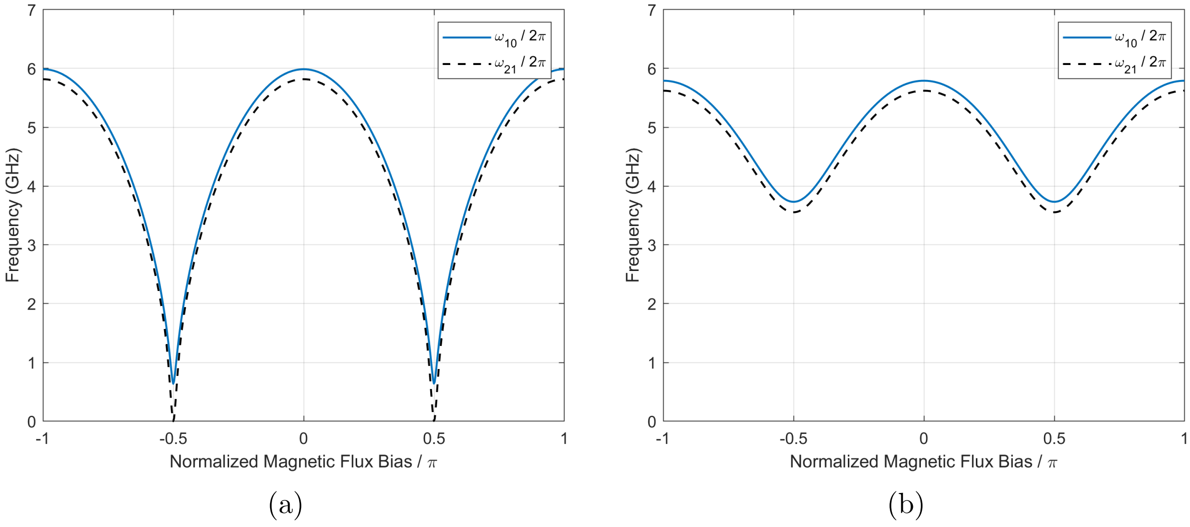
(a) For $E_J/E_{C_\Sigma} = 0.1$, the Cooper pair box regime.

(b) For $E_J/E_{C_\Sigma} = 50$, the transmon regime.

(c) Discuss why the wavefunctions in (a) are not localized with respect to φ, but they are for (b).

Note: Don't forget to normalize your wavefunctions!

4. This problem considers aspects of energy level diagrams for transmon qubits that have different kinds of SQUIDs. The case of a symmetric SQUID (both Josephson junctions have the same Josephson energy) is illustrated in Fig. 9.11(a), while the case of an asymmetric SQUID (the Josephson junctions have different Josephson energies) is shown in Fig. 9.11(b).



(a) (b)

Figure 9.11

(a) It is clear that the transition frequency tunability of a symmetric SQUID is, in principle, far superior to that of an asymmetric SQUID. Provide at least two reasons why this larger tuning range is not able to be realistically used when attempting to operate the transmon with a symmetric SQUID as a qubit.

(b) What magnetic flux bias points would lead to the best qubit performance for a transmon with an asymmetric SQUID? Which bias points would lead to the worst qubit performance?

5. This problem considers the possibility of transitioning an artificial atom described by the following Hamiltonian

$$\hat{H}_0 = 4E_C\big(\hat{n} - 0.5\big)^2 - E_J \cos\hat{\varphi} \tag{9.39}$$

into various energy eigenstates via an interaction with an external voltage source. The energy eigenstates are denoted by $|m\rangle$, where m is an integer counting from 0 to positive numbers that denotes the ordering of the energy eigenstates. The total Hamiltonian for this system is

$$\hat{H} = \hat{H}_0 + 2qV_b(t)\hat{n}, \tag{9.40}$$

where $V_b(t)$ is a bias voltage source. Relevant matrix elements of $\hat{n}$ are given in Fig. 9.12 as a function of the characteristic ratio E_J/E_C.

(a) Assume the artificial atom is operated roughly in the Cooper pair box regime where $E_J/E_C \approx 1$ and that the system starts in the ground state $|0\rangle$. Discuss the relative difficulty of using the voltage source $V_b(t)$ to transition the system into the $|2\rangle$ or $|3\rangle$ states, where it is acceptable to get to these final eigenstates through any number of intermediary states (i.e., a direct transition from $|0\rangle$ to $|2\rangle$ or $|3\rangle$ is *not* required).

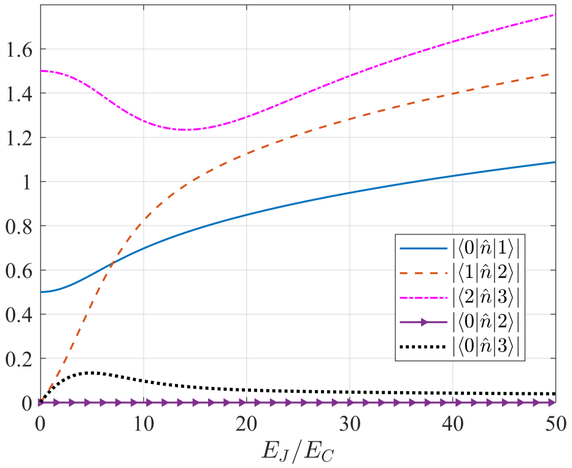


Figure 9.12

(b) The artificial atom is now operated as a transmon where $E_J/E_C \gg 1$ and the system begins in the ground state $|0\rangle$. Describe two procedures to transition the system into the $|3\rangle$ state. When describing your procedures, provide qualitative information about the pulses that will need to be applied (e.g., their sequence, the center frequency of the pulse, the relative amplitudes of the pulses).

6. For the superposition state of a two-level system in the spin representation given by

$$|s\rangle = \cos(\theta/2)|\uparrow\rangle + \sin(\theta/2)e^{i\phi}|\downarrow\rangle, \tag{9.41}$$

show that the expectation value of the spin angular momentum satisfies

$$\langle s|\hat{\boldsymbol{\sigma}}|s\rangle = \tilde{x}\sin\theta\cos\phi + \tilde{y}\sin\theta\sin\phi + \tilde{z}\cos\theta, \tag{9.42}$$

where $\tilde{x}$ denotes a unit vector along the x-direction of a usual 3D Cartesian coordinate system (and similar for $\tilde{y}$ and $\tilde{z}$).

Chapter 10

Circuit and Artificial Atom Interactions

10.1 Introduction to Formulating the Classical Hamiltonian

Now that we have an understanding of the generic description of a two-level system, let's look at an example of how we can "map" an actual physical system into this unified picture. Specifically, we will consider a superconducting circuit "charge qubit" (e.g., either a Cooper pair box or a transmon) driven by a voltage source capacitively coupled to our artificial atom. We already introduced the Hamiltonian for this situation previously, but simply stated it as a "fact". Later, we will want to consider how to formulate the quantized interactions between an LC oscillator and this type of artificial atom. Considering this, we will take a more careful look at how to formulate the Hamiltonian for our current case as a stepping stone toward our later goals.

Towards this end, we will consider a slightly more "phenomenological" model for our voltage source. In particular, instead of assuming it to be a perfectly ideal source from the beginning, we will assume that it is simply a very large capacitor. Our circuit schematic for this situation is shown in Fig. 10.1.

We can now go about formulating the Lagrangian and then the resulting Hamiltonian for this circuit. This follows closely to the example we considered previously of two capacitively-coupled LC oscillators. For our current situation, we can quickly determine that the Lagrangian should be

$$\mathcal{L} = \frac{1}{2}C_1\dot{\phi}_1^2 + \frac{1}{2}C_2\dot{\phi}_2^2 + \frac{1}{2}C_c\left(\dot{\phi}_1 - \dot{\phi}_2\right)^2 + E_J\cos\left(\frac{2q}{\hbar}\phi_1\right), \tag{10.1}$$

where $(2q/\hbar)\phi_1$ will be equal to the Josephson phase variable we previously considered for charge qubits. We will eventually write our Hamiltonian using that kind of dimensionless quantities, but for now we will keep everything in terms of ϕ_1 and ϕ_2. Also, note here that the capacitive part of this Lagrangian exactly matches what we had for the capacitively-coupled LC oscillators case considered previously. Considering this, we can follow a similar process

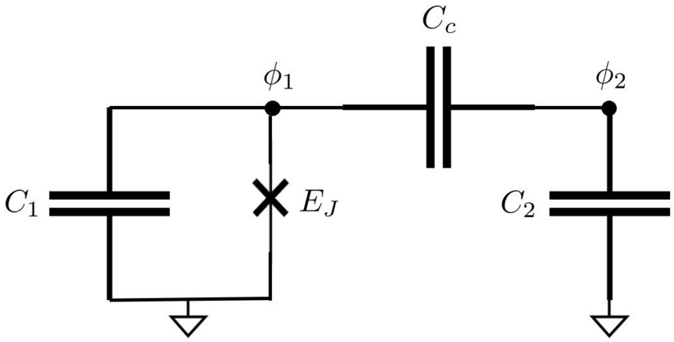


Figure 10.1: Circuit schematic for a large capacitance (C_2) capacitively-coupled to a charge qubit.

to find our Hamiltonian, which will require us to determine the inverse of the capacitance matrix encoded into (10.1). From our prior analysis, we can find that this will be

$$[C]^{-1} = \frac{1}{C_1C_2 + C_cC_1 + C_cC_2}\begin{bmatrix} C_2 + C_c & C_c \\ C_c & C_1 + C_c \end{bmatrix}. \tag{10.2}$$

This inverse capacitance matrix is a bit of a "mess", but we can simplify it by defining a few new capacitances. In particular, we can find that the diagonal elements of the inverse capacitance matrix are equivalent to the inverse of the total capacitance to ground seen by our nodes that ϕ_1 and ϕ_2 are defined at. Explicitly, we have that the total capacitance at nodes 1 and 2 are

$$C_{\Sigma_1} = C_1 + \frac{C_cC_2}{C_c + C_2} \tag{10.3}$$

$$C_{\Sigma_2} = C_2 + \frac{C_cC_1}{C_c + C_1}, \tag{10.4}$$

respectively. We can also define a voltage divider between the coupling capacitance C_c and the artificial atom capacitance C_1 as

$$\beta = \frac{C_c}{C_c + C_1}. \tag{10.5}$$

In terms of these new symbols, (10.2) becomes

$$[C]^{-1} = \begin{bmatrix} 1/C_{\Sigma_1} & \beta/C_{\Sigma_2} \\ \beta/C_{\Sigma_2} & 1/C_{\Sigma_2} \end{bmatrix}. \tag{10.6}$$

Defining our conjugate momenta Q_j in the usual manner, we can find that our Hamiltonian becomes

$$H = \frac{1}{2C_{\Sigma_1}}Q_1^2 - E_J\cos\left(\frac{2q}{\hbar}\phi_1\right) + \frac{1}{2C_{\Sigma_2}}Q_2^2 + \frac{\beta}{C_{\Sigma_2}}Q_1Q_2. \tag{10.7}$$

Our manipulations up to this point have been exact. However, to simplify our expression we will now make use of the assumption that C_2 is very large compared to the other capacitances. In this situation, we can assume that the bias voltage provided from node 2 is

$$V_b = \frac{1}{C_{\Sigma_2}} Q_2. \tag{10.8}$$

In terms of this, our Hamiltonian becomes

$$H = \frac{1}{2C_{\Sigma_1}} Q_1^2 - E_J \cos\left(\frac{2q}{\hbar}\phi_1\right) + \frac{1}{2} C_{\Sigma_2} V_b^2 + \beta V_b Q_1. \tag{10.9}$$

To be more consistent with the standard description of charge qubits, we will now rewrite Q_1 and ϕ_1 into their dimensionless forms as $Q_1 = 2qn$ and $\varphi = (2q/\hbar)\phi_1$. Our Hamiltonian is now

$$H = 4E_C n^2 - E_J \cos\varphi + \frac{1}{2} C_{\Sigma_2} V_b^2 + 2q\beta V_b n, \tag{10.10}$$

where $E_C = q^2/2C_{\Sigma_1}$.

Next, we will consider a *semiclassical* treatment of this system to continue our investigation. In this semiclassical picture, we will treat the Josephson junction part of the system quantum mechanically, but assume that our bias voltage V_b is classical.

10.2 Semiclassical Treatment

Previously, we derived a classical Hamiltonian describing the circuit in Fig. 10.1 to be (10.10). We will now consider a *semiclassical* treatment of this system to continue our investigation. In this semiclassical picture, we will treat the Josephson junction part of the system quantum mechanically, but assume that our bias voltage V_b is classical.

Our Hamiltonian in (10.10) has been written in terms of (dimensionless) conjugate variables, so we can readily quantize it as we did when originally discussing this kind of artificial atom. We will, however, leave our V_b as a classical variable. In this case, we have the commutation relation

$$[\hat{n}, \hat{\varphi}] = -i \tag{10.11}$$

and our Hamiltonian operator becomes

$$\hat{H} = 4E_C \hat{n}^2 - E_J \cos\hat{\varphi} + 2q\beta V_b \hat{n}, \tag{10.12}$$

where we have dropped the V_b^2 term since this is a constant in terms of our dynamical variables, and so, will not affect the dynamics of the system.

To describe this system in our unified spin or qubit representation, we need to determine which quantum states will compose our two-level system. For this kind of artificial atom, we will utilize the two lowest energy eigenstates of the Hamiltonian

$$\hat{H}_0 = 4E_C \hat{n}^2 - E_J \cos\hat{\varphi} \tag{10.13}$$

to act as the levels in our two-level system. Denoting our eigenstates as $|m\rangle$, we would find that they must satisfy the time-independent Schrödinger equation

$$\hat{H}_0|m\rangle = \hbar\omega_m|m\rangle, \tag{10.14}$$

where we have elected to write the eigenvalue ω_m in units of frequency. This is commonly done when considering interactions between two different quantum systems, since as we have seen previously, many interactions will be suppressed unless the frequencies of the systems are nearly resonant.

If we now make the assumption that our two lowest energy levels of $\hat{H}_0$ are isolated enough from any other energy levels of $\hat{H}_0$, then we can make our two-level approximation. In particular, we will assume that the only states in the system that matter are the $|0\rangle$ and $|1\rangle$ states of $\hat{H}_0$. We can then project the full Hamiltonian of (10.12) onto these two states. We will denote this projected Hamiltonian as $\hat{H}_{\text{TLS}}$, which we can compute as

$$\hat{H}_{\text{TLS}} = \Bigg(\sum_{m=0}^{1}|m\rangle\langle m|\Bigg)\hat{H}\Bigg(\sum_{\ell=0}^{1}|\ell\rangle\langle\ell|\Bigg). \tag{10.15}$$

We can expand our terms out to facilitate evaluating $\hat{H}_{\text{TLS}}$, which gives us

$$\begin{aligned}\hat{H}_{\text{TLS}} &= \Big(|0\rangle\langle 0| + |1\rangle\langle 1|\Big)\Big(\hat{H}_0 + 2q\beta V_b\hat{n}\Big)\Big(|0\rangle\langle 0| + |1\rangle\langle 1|\Big)\\ &= |0\rangle\langle 0|\hat{H}_0|0\rangle\langle 0| + |1\rangle\langle 1|\hat{H}_0|1\rangle\langle 1| + |0\rangle\langle 0|\hat{H}_0|1\rangle\langle 1| + |1\rangle\langle 1|\hat{H}_0|0\rangle\langle 0|\\ &\quad + 2q\beta V_b\Big(|0\rangle\langle 0|\hat{n}|0\rangle\langle 0| + |1\rangle\langle 1|\hat{n}|1\rangle\langle 1| + |0\rangle\langle 0|\hat{n}|1\rangle\langle 1| + |1\rangle\langle 1|\hat{n}|0\rangle\langle 0|\Big).\end{aligned} \tag{10.16}$$

We can simplify the terms involving $\hat{H}_0$ using properties of the eigenstates to get

$$\hat{H}_{\text{TLS}} = \hbar\omega_0|0\rangle\langle 0| + \hbar\omega_1|1\rangle\langle 1| + \hbar V_b\Big(g_{00}|0\rangle\langle 0| + g_{11}|1\rangle\langle 1| + g_{01}|0\rangle\langle 1| + g_{10}|1\rangle\langle 0|\Big), \tag{10.17}$$

where we have also introduced the coupling factor

$$g_{m\ell} = (2q\beta/\hbar)\langle m|\hat{n}|\ell\rangle. \tag{10.18}$$

Although we have simplified our Hamiltonian to a large extent, (10.17) can be simplified further by making a few assumptions about $g_{m\ell}$. In particular, for many artificial atoms (including our charge qubits), one can find that the diagonal elements like g_{00} and g_{11} will vanish. This often happens due to underlying symmetries in the Hamiltonian $\hat{H}_0$ that cause terms like $\langle m|\hat{n}|m\rangle$ to equal 0. One further simplification is that we can also see that, in general, $g_{m\ell} = g_{\ell m}^*$. Since there is always an arbitrary phase shift that we can apply to our eigenstates, we can choose to make this such that g_{01} is real. In this case, we can simplify our expressions by recognizing that $g_{01} = g_{10} \equiv g$. Applying these two simplifications to (10.17), we get that

$$\hat{H}_{\text{TLS}} = \hbar\omega_0|0\rangle\langle 0| + \hbar\omega_1|1\rangle\langle 1| + \hbar g V_b\big(|0\rangle\langle 1| + |1\rangle\langle 0|\big). \tag{10.19}$$

We can write this into a matrix notation in terms of the Pauli spin matrices by shifting the energy spectrum by a constant value of $-\hbar\omega_{10}/2$, where $\omega_{10} = \omega_1 - \omega_0$, to get

$$\hat{H}_{\text{TLS}} = -\frac{1}{2}\hbar\omega_{10}\hat{\sigma}_z + \hbar g V_b(\hat{\sigma}_+ + \hat{\sigma}_-), \tag{10.20}$$

where

$$\hat{\sigma}_+ = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad \hat{\sigma}_- = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}. \tag{10.21}$$

Note that we could have written $\hat{\sigma}_+ + \hat{\sigma}_- = \hat{\sigma}_x$, but we have left our Hamiltonian in terms of $\hat{\sigma}_+$ and $\hat{\sigma}_-$ for now in anticipation of some upcoming simplifications we will consider. Before continuing, it should be mentioned that the Hamiltonian in (10.20) is more often written with the convention that $|\uparrow\rangle \to |1\rangle$ and $|\downarrow\rangle \to |0\rangle$ so that the sign in front of $\hat{\sigma}_z$ will be positive. We have not followed that convention here to keep everything consistent with the Bloch sphere that we are using in this course.

To gain some insight into this system, it is instructive to look at the Heisenberg equations of motion for our various operators. Recall that these equations of motion can be computed as

$$\frac{d}{dt}\hat{X}(t) = \frac{1}{i\hbar}[\hat{X}(t), \hat{H}_{\text{TLS}}]. \tag{10.22}$$

Considering this, we will need to know the commutators between $\hat{\sigma}_z$, $\hat{\sigma}_+$, and $\hat{\sigma}_-$. These can be found very easily using the definitions of the various matrix representations, which we summarize as

$$[\hat{\sigma}_z, \hat{\sigma}_\pm] = \mp 2\hat{\sigma}_\pm, \quad [\hat{\sigma}_-, \hat{\sigma}_+] = \hat{\sigma}_z. \tag{10.23}$$

Using these commutators, we find that

$$\frac{d}{dt}\hat{\sigma}_z(t) = i2gV_b\big(\hat{\sigma}_+(t) - \hat{\sigma}_-(t)\big), \tag{10.24}$$

$$\frac{d}{dt}\hat{\sigma}_+(t) = i\omega_{10}\hat{\sigma}_+(t) + igV_b\hat{\sigma}_z(t), \tag{10.25}$$

$$\frac{d}{dt}\hat{\sigma}_-(t) = -i\omega_{10}\hat{\sigma}_-(t) - igV_b\hat{\sigma}_z(t). \tag{10.26}$$

If we temporarily ignore the presence of the bias voltage V_b, we can see that the underlying *free oscillations* of $\hat{\sigma}_\pm(t)$ will follow $\hat{\sigma}_\pm(0)\exp[\pm i\omega_{10}t]$ time dependence. We can utilize this result and some judicious approximations to now rewrite (10.20) into a particularly useful form for understanding how a classical drive can control the state of our two-level system.

To do this, we will first assume that we can control V_b as a function of time. We will further parameterize our bias voltage as

$$V_b(t) = \begin{cases} V_x(t)\cos(\omega_d t) + V_y(t)\sin(\omega_d t), & 0 < t < t_g \\ 0, & \text{otherwise}, \end{cases} \tag{10.27}$$

where ω_d is the drive frequency and $V_x(t)$ and $V_y(t)$ are *envelope functions* that are slowly-varying in comparison to ω_d. This kind of parametrization is very commonly used, and is known as *IQ modulation.* The "I" stands for "in phase" and the "Q" stands for "quadrature phase". These names are in reference to the two *channels* of our signal being multiplied by a cosine (in-phase component) and sine (quadrature-phase component). As we will see shortly, these two channels will act as independent control sources for our two-level system.

We can now substitute (10.27) into (10.20) to find that (assuming $0 < t < t_g$)

$$\hat{H}_{\text{TLS}} = -\frac{1}{2}\hbar\omega_{10}\hat{\sigma}_z(t) + \hbar g\Big(V_x(t)\cos(\omega_d t) + V_y(t)\sin(\omega_d t)\Big)\Big(\hat{\sigma}_+(t) + \hat{\sigma}_-(t)\Big). \tag{10.28}$$

It is now instructive to "factor out" the underlying known time variation of $\hat{\sigma}_+(t)$ and $\hat{\sigma}_-(t)$ to get

$$\hat{H}_{\text{TLS}} = -\frac{1}{2}\hbar\omega_{10}\hat{\sigma}_z(t) + \hbar g\Big(V_x(t)\cos(\omega_d t) + V_y(t)\sin(\omega_d t)\Big)\Big(\hat{\sigma}_+(t)e^{i\omega_{10}t} + \hat{\sigma}_-(t)e^{-i\omega_{10}t}\Big). \tag{10.29}$$

We can now expand $\cos(\omega_d t)$ and $\sin(\omega_d t)$ in terms of exponential functions and regroup terms to get

$$\begin{aligned}\hat{H}_{\text{TLS}} = -\frac{1}{2}\hbar\omega_{10}\hat{\sigma}_z(t) &+ \frac{1}{2}\hbar g V_x(t)\Big[\sigma_+(t)e^{i(\omega_{10}-\omega_d)t} + \sigma_-(t)e^{-i(\omega_{10}-\omega_d)t}\Big] \\ &+ \frac{i}{2}\hbar g V_y(t)\Big[\sigma_+(t)e^{i(\omega_{10}-\omega_d)t} - \sigma_-(t)e^{-i(\omega_{10}-\omega_d)t}\Big] \\ &+ \frac{1}{2}\hbar g V_x(t)\Big[\sigma_+(t)e^{i(\omega_{10}+\omega_d)t} + \sigma_-(t)e^{-i(\omega_{10}+\omega_d)t}\Big] \\ &- \frac{i}{2}\hbar g V_y(t)\Big[\sigma_+(t)e^{i(\omega_{10}+\omega_d)t} - \sigma_-(t)e^{-i(\omega_{10}+\omega_d)t}\Big].\end{aligned} \tag{10.30}$$

Although this expression is somewhat messy, the purpose of expanding terms in this way was to help us recognize that if we make $\omega_d \approx \omega_{10}$ then we can see that terms that are multiplied by $\exp[\pm i(\omega_{10}-\omega_d)t]$ will be very slowly-varying, while terms multiplied by $\exp[\pm i(\omega_{10}+\omega_d)t]$ will be oscillating extremely quickly in comparison.

We can now make a very common approximation, known as the *rotating wave approximation (RWA)*, and say that terms multiplied by $\exp[\pm i(\omega_{10}+\omega_d)t]$ oscillate so quickly that over any "appreciable" time scale that we observe the system over these terms will have oscillated so many times that their contribution effectively averages to 0. Due to this, we can choose to simply neglect these terms entirely and write our Hamiltonian as

$$\begin{aligned}\hat{H}_{\text{TLS,RWA}} = -\frac{1}{2}\hbar\omega_{10}\hat{\sigma}_z(t) &+ \frac{1}{2}\hbar g V_x(t)\Big[\sigma_+(t)e^{i(\omega_{10}-\omega_d)t} + \sigma_-(t)e^{-i(\omega_{10}-\omega_d)t}\Big] \\ &+ \frac{i}{2}\hbar g V_y(t)\Big[\sigma_+(t)e^{i(\omega_{10}-\omega_d)t} - \sigma_-(t)e^{-i(\omega_{10}-\omega_d)t}\Big].\end{aligned} \tag{10.31}$$

(If you are uncomfortable with this approximation, jump ahead to Fig. 10.2 to see the impact of these quickly-varying terms on the dynamics of the system for a practical scenario.) We can further move to a "rotating frame" to view the dynamics relative to our drive frequency. We will omit the details here for brevity, but this leads to the Hamiltonian

$$\hat{H}_{\text{TLS,RWA}} = -\frac{1}{2}\hbar\delta\hat{\sigma}_z(t) + \frac{1}{2}\hbar g V_x(t)\Big[\sigma_+(t) + \sigma_-(t)\Big] + \frac{i}{2}\hbar g V_y(t)\Big[\sigma_+(t) - \sigma_-(t)\Big], \quad (10.32)$$

where $\delta = \omega_{10} - \omega_d$ and we have embedded the underlying free oscillating time dependence back into $\sigma_\pm(t)$. We can further write this Hamiltonian purely in terms of our original Pauli spin matrices to see that

$$\hat{H}_{\text{TLS,RWA}} = -\frac{1}{2}\hbar\delta\hat{\sigma}_z(t) + \frac{1}{2}\hbar g V_x(t)\hat{\sigma}_x(t) + \frac{1}{2}\hbar g V_y(t)\hat{\sigma}_y(t). \quad (10.33)$$

As alluded to previously, we can now see that our V_x and V_y can be used to independently control the $\hat{\sigma}_x$ and $\hat{\sigma}_y$ terms in the Hamiltonian.

10.2.1 Driven Rabi Oscillations

We now wish to see what applying a drive on one of our channels will do to the state of the two-level system. We can determine this by solving the time-dependent Schrödinger equation for the Hamiltonian (10.33) written back in the Schrödinger picture. To simplify the analysis, we will assume that our drive is perfectly resonant with the transition frequency ω_{10} of the two-level system (i.e. $\delta = 0$) and that we only have a drive on the V_x channel. Our Hamiltonian then becomes

$$\hat{H} = \frac{1}{2}\hbar g V_x(t)\hat{\sigma}_x. \quad (10.34)$$

We can now expand the time-dependent Schrödinger equation in terms of the states of the two-level system. We will find in a matrix notation that the time-dependent Schrödinger equation becomes

$$i\hbar\frac{d}{dt}\begin{Bmatrix} c_0(t) \\ c_1(t) \end{Bmatrix} = \hat{H}\begin{Bmatrix} c_0(t) \\ c_1(t) \end{Bmatrix}. \quad (10.35)$$

From this, we see that our system corresponds to two ordinary differential equations for our two expansion coefficients as

$$\dot{c}_0(t) = -\frac{i}{2} g V_x(t) c_1(t) \quad (10.36)$$

$$\dot{c}_1(t) = -\frac{i}{2} g V_x(t) c_0(t). \quad (10.37)$$

Note that these equations take this simple form because of the off-diagonal nature of $\hat{\sigma}_x$.

To gain more insight into the basic behavior of these equations, it is instructive to assume that we leave our envelope function $V_x(t)$ to be equal to some constant for an "extended" period of time. If we assume this constant is $V_{x,0}$ then our equations simplify to

$$\dot{c}_0(t) = -\frac{i}{2} g V_{x,0} c_1(t), \quad (10.38)$$

$$\dot{c}_1(t) = -\frac{i}{2} g V_{x,0} c_0(t). \tag{10.39}$$

We can solve these equations by differentiating (10.38) and substituting in from (10.39) to get

$$\ddot{c}_0(t) + \left(\frac{g V_{x,0}}{2}\right)^2 c_0(t) = 0. \tag{10.40}$$

This equation describes simple sinusoidal oscillations with frequency $g V_{x,0}/2$. If we set as initial condition that $c_0(0) = 1$ and $c_1(0) = 0$, then we can determine that the solution is

$$c_0(t) = \cos\left(\frac{g V_{x,0}}{2} t\right), \tag{10.41}$$

$$c_1(t) = -i \sin\left(\frac{g V_{x,0}}{2} t\right). \tag{10.42}$$

There are a few points of interest that we can note from (10.41) and (10.42). First, if we look at the probability of the states being occupied we find that

$$|c_0(t)|^2 = \cos^2\left(\frac{g V_{x,0}}{2} t\right), \tag{10.43}$$

$$|c_1(t)|^2 = \sin^2\left(\frac{g V_{x,0}}{2} t\right). \tag{10.44}$$

We see that our system smoothly oscillates back and forth between being in the ground and excited states, with a frequency set by the coupling between our source and the two-level system and the amplitude of our drive. These oscillations are known as *driven Rabi oscillations*, and the parameter $g V_{x,0}$ is typically referred to as the *Rabi frequency* and is often denoted as Ω_R.

Another interesting point to note is the trajectory of these oscillations on the Bloch sphere. We can see that if we let $t = (1/\Omega_R)\,\pi/2$ then our system will be in the state $|-i\rangle = 1\sqrt{2}(|0\rangle - i|1\rangle)$. If we allow the oscillation to continue to $t = (1/\Omega_R)\,\pi$ we find we are in state $|1\rangle$ (up to an unimportant phase factor). We see that these oscillations appear to be causing a counterclockwise rotation around the $+x$-axis of our state vector on the Bloch sphere, which can be confirmed from a more detailed analysis. Hence, we see that driving our channel "connected" to $\hat{\sigma}_x$ leads to rotations around the $+x$-axis. A similar analysis can be performed to see that driving the channel "connected" to $\hat{\sigma}_y$ will drive counterclockwise rotations around the $+y$-axis.

For practical purposes, it is undesirable to try and control a two-level system by simply leaving a purely sinusoidal drive on for some amount of time to achieve the desired rotation on the Bloch sphere. If we allow the pulse envelope function to take on more arbitrary shapes, it becomes convenient to define the *pulse area* of the envelope as

$$\Theta(t) = g V_0 \int_0^t s(t') dt', \tag{10.45}$$

where V_0 is our amplitude and $s(t)$ defines the pulse shape. If we design our parameters such that the total area of the pulse gives us $\Theta = \pi$, then this pulse will drive a rotation on the Bloch sphere of π degrees around the axis the pulse was "connected" to (e.g., x- or y-axes). This is often referred to as a π-pulse, and can be used to take a starting state of $|0\rangle$ and transition it to $|1\rangle$ (or vice-versa). Similarly, a $\pi/2$-pulse is also commonly used, which would rotate the state $|0\rangle$ into one of the states $|\pm\rangle$ or $|\pm i\rangle$, depending on the specific settings of the pulse. These are equal superposition states that exist on the "equator" of the Bloch sphere, and play very important roles in many quantum computing algorithms.

10.2.2 A Brief Note on the Rotating Wave Approximation

Before concluding our discussion on this semiclassical treatment, it is worth revisiting the rotating wave approximation (RWA) briefly. The first time this is seen it often comes across in a somewhat mysterious way that can raise questions on the validity of it. In a general scenario, it can sometimes be difficult to determine whether the RWA is valid or not. Further, when we look at the fully-quantized description of a circuit interaction with a two-level system, we will see that the terms that would be discarded if we make the RWA describe somewhat unintuitive physical processes. To partially put minds at ease, we include here the full solution of a semiclassical model that does not make the RWA for an approximately resonant π-pulse drive of a two-level system. The results are shown in Fig. 10.2. We see that if we zoom in very close to a particular trace (this is shown in the inset of the figure) that there are actually some small amplitude oscillations superimposed on the slowly-varying trace. These quicker oscillations have a frequency of $2\omega_{10}$, and constitute the contribution of the *counter-rotating wave* terms. Clearly, for this situation these effects are very small, and so are unnecessary to be considered to grasp the core physics of this system. However, if we need to be extremely precise in our calculations (which can certainly be necessary for emerging quantum technologies), then we may certainly need to consider the full model rather than only considering a system that has used the RWA.

10.3 Introduction to Fully-Quantized Circuit and Artificial Atom Interactions

Previously, we looked at how we could map an actual physical system of a charge qubit being driven by a classical voltage source onto a semiclassical model of a two-level system. We will now build on this example by considering a fully-quantized model of a charge qubit interacting with a quantum LC oscillator. The circuit schematic of the system we will consider is shown in Fig. 10.3. As with previous examples, we will begin with a classical description of this system and then quantize it through a canonical quantization procedure.

Considering this, our first step is to write down the classical Lagrangian for the circuit shown in Fig. 10.3. This can be done easily building on our prior semiclassical analysis, and gives us

$$\mathcal{L} = \frac{1}{2}C_1\dot{\phi}_1^2 + E_J\cos\left(\frac{2q}{\hbar}\phi_1\right) + \frac{1}{2}C_2\dot{\phi}_2^2 - \frac{1}{2L_2}\phi_2^2 + \frac{1}{2}C_c\left(\dot{\phi}_1 - \dot{\phi}_2\right)^2. \tag{10.46}$$

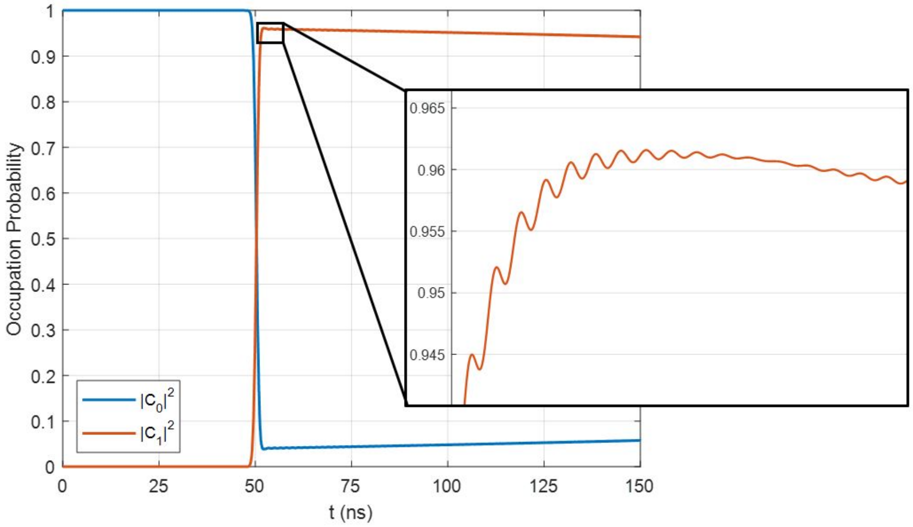


Figure 10.2: Results of an approximately resonant π-pulse for a semiclassical model of a two-level system that does not make the RWA. The impact of the counter-rotating terms are shown in the inset, where it is seen that they make only a very small contribution to the overall result.

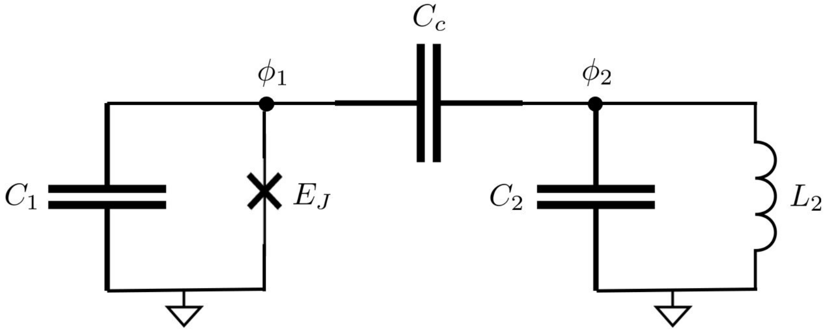


Figure 10.3: Circuit schematic for a LC tank circuit capacitively-coupled to a charge qubit.

The capacitance matrix for this system is identical to the cases we have considered previously, so we sill have that the inverse capacitance matrix that helps us convert between generalized "velocities" $\dot{\phi}_j$ to conjugate momenta Q_j will be

$$[C]^{-1} = \begin{bmatrix} 1/C_{\Sigma_1} & \beta/C_{\Sigma_2} \\ \beta/C_{\Sigma_2} & 1/C_{\Sigma_2} \end{bmatrix}. \tag{10.47}$$

Recall that our C_{Σ_j}'s were the total capacitance to ground at nodes 1 and 2, which are

$$C_{\Sigma_1} = C_1 + \frac{C_c C_2}{C_c + C_2}, \tag{10.48}$$

$$C_{\Sigma_2} = C_2 + \frac{C_c C_1}{C_c + C_1}, \tag{10.49}$$

respectively, and that β was a voltage divider between the coupling capacitance C_c and the artificial atom capacitance C_1, given as

$$\beta = \frac{C_c}{C_c + C_1}. \tag{10.50}$$

Using these relationships, we can find our total Hamiltonian for this system. In our current case, this is

$$H = \frac{1}{2C_{\Sigma_1}} Q_1^2 - E_J \cos\left(\frac{2q}{\hbar}\phi_1\right) + \frac{1}{2C_{\Sigma_2}} Q_2^2 + \frac{1}{2L_2}\phi_2^2 + \frac{\beta}{C_{\Sigma_2}} Q_1 Q_2. \tag{10.51}$$

We can introduce the usual dimensionless parameters for the quantities related to the charge qubit to rewrite this as

$$H = 4E_C n^2 - E_J \cos\varphi + \frac{1}{2C_{\Sigma_2}} Q_2^2 + \frac{1}{2L_2}\phi_2^2 + \frac{2q\beta}{C_{\Sigma_2}} n Q_2. \tag{10.52}$$

We will now consider the full quantization of this Hamiltonian.

10.3.1 Quantum Treatment

We can now perform a canonical quantization on our system easily since we have written the Hamiltonian in appropriate conjugate functions. Hence, we promote n and φ to be quantum operators with commutation relation

$$[\hat{n}, \hat{\varphi}] = -i \tag{10.53}$$

and promote ϕ_2 and Q_2 to be quantum operators with commutation relation

$$[\hat{Q}_2, \hat{\phi}_2] = -i\hbar. \tag{10.54}$$

Our Hamiltonian operator then becomes

$$\hat{H} = 4E_C \hat{n}^2 - E_J \cos\hat{\varphi} + \frac{1}{2C_{\Sigma_2}} \hat{Q}_2^2 + \frac{1}{2L_2}\hat{\phi}_2^2 + \frac{2q\beta}{C_{\Sigma_2}} \hat{n} \hat{Q}_2. \tag{10.55}$$

In many cases, attempting to work directly with (10.55) is not particularly advantageous. To gain more insight into some of the fundamental physical processes that can occur in this quantized system, it is useful to recast our Hamiltonian in terms of different operators. To begin, we will follow the example of our semiclassical analysis and rewrite the charge qubit part of our Hamiltonian in terms of it's energy eigenstates so that we can utilize a generic two-level system representation for it. Making the same assumptions as in our previous analysis, we can modify (10.55) to become

$$\hat{H} = -\frac{1}{2}\hbar\omega_{10}\hat{\sigma}_z + \frac{1}{2C_{\Sigma_2}} \hat{Q}_2^2 + \frac{1}{2L_2}\hat{\phi}_2^2 + \frac{2q\beta}{C_{\Sigma_2}} \langle 0|\hat{n}|1\rangle \hat{Q}_2 \big(\hat{\sigma}_+ + \hat{\sigma}_-\big). \tag{10.56}$$

Recall that ω_{10} is the frequency associated with the transition between the ground and first excited state of the charge qubit and that we have assumed our charge operator $\hat{n}$ does not have diagonal matrix elements (i.e., $\langle m|\hat{n}|m\rangle = 0,\ \forall m$) and that we have selected phase shifts in our definitions of the eigenstates such that $\langle 0|\hat{n}|1\rangle = \langle 1|\hat{n}|0\rangle$.

Likewise, we can rewrite the operators of the LC tank circuit in terms of the bosonic ladder operators for this quantum LC oscillator. For this circuit, we will write these in the slightly different form as

$$\hat{a} = \frac{1}{\sqrt{2C_{\Sigma_2}\hbar\omega_r}}\hat{Q}_2 + i\frac{1}{\sqrt{2L\hbar\omega_r}}\hat{\phi}_2, \tag{10.57}$$

$$\hat{a}^\dagger = \frac{1}{\sqrt{2C_{\Sigma_2}\hbar\omega_r}}\hat{Q}_2 - i\frac{1}{\sqrt{2L\hbar\omega_r}}\hat{\phi}_2, \tag{10.58}$$

where $\omega_r = 1/\sqrt{LC_{\Sigma_2}}$. The reason for this shift in definition is mostly "cosmetic", it will causes the signs of later expressions to more closely match those of an extremely widely-used Hamiltonian. Regardless, these operators still obey the usual commutation relation $[\hat{a}, \hat{a}^\dagger] = 1$. In terms of these operators, we can also express the LC charge operator as

$$\hat{Q}_2 = Q_{\mathrm{ZPF}}(\hat{a} + \hat{a}^\dagger) \tag{10.59}$$

where

$$Q_{\mathrm{ZPF}} = \sqrt{\frac{C_{\Sigma_2}\hbar\omega_r}{2}}. \tag{10.60}$$

Using these various results, we can now write (10.56) as

$$\hat{H} = -\frac{1}{2}\hbar\omega_{10}\hat{\sigma}_z + \hbar\omega_r\hat{a}^\dagger\hat{a} + \frac{2q\beta}{C_{\Sigma_2}}Q_{\mathrm{ZPF}}\langle 0|\hat{n}|1\rangle\big(\hat{a} + \hat{a}^\dagger\big)\big(\hat{\sigma}_+ + \hat{\sigma}_-\big), \tag{10.61}$$

where we have dropped the zero point energy of the LC oscillator because it does not change the dynamics of the physical system that we are interested in. It is customary to consolidate the constant factors that define the strength of the overall interaction between the two-level system and the LC oscillator into a single parameter, typically denoted as g. In our current case, we would have that

$$g = \frac{2q\beta V_{\mathrm{ZPF}}}{\hbar}\langle 0|\hat{n}|1\rangle, \tag{10.62}$$

where we have also redefined the zero-point fluctuations in terms of the voltage in the LC oscillator as

$$V_{\mathrm{ZPF}} = \sqrt{\frac{\hbar\omega_r}{2C_{\Sigma_2}}}. \tag{10.63}$$

Hence, we can rewrite our Hamiltonian of (10.61) as

$$\hat{H} = -\frac{1}{2}\hbar\omega_{10}\hat{\sigma}_z + \hbar\omega_r\hat{a}^\dagger\hat{a} + \hbar g\big(\hat{a} + \hat{a}^\dagger\big)\big(\hat{\sigma}_+ + \hat{\sigma}_-\big). \tag{10.64}$$

This Hamiltonian is often used in what is referred to as the *quantum Rabi model*.

Before moving on, it is useful to recall a few details about working with a Hamiltonian like this one that is "composed" of Hamiltonians from different subsystems. We already discussed these points with respect to the quantum analysis of two capacitively-coupled LC oscillators, but it will be useful to revisit them in the context of the current system. First, we need to consider again how we will express the quantum state of the system. From the form of the operators we have written (10.64) in terms of, it seems clear that we will express the state of the charge qubit in terms of a ground and excited state of the two-level system and the state of the LC oscillator in terms of number or Fock states. Considering this, we write the complete state of the system in terms of the tensor product of the states of the individual systems. In abstract notation, we would write this as

$$\mathcal{H} = \mathcal{H}_{\text{TLS}} \otimes \mathcal{H}_{\text{LC}}, \tag{10.65}$$

where $\mathcal{H}_{\text{TLS}}$ is the Hilbert space of the two-level system and $\mathcal{H}_{\text{LC}}$ is the Hilbert space of the LC oscillator. To help with differentiating the notation of the different subsystems, we will choose to write the states of the two-level system as

$$\mathcal{H}_{\text{TLS}} = \{|\text{g}\rangle, |e\rangle\}, \tag{10.66}$$

where $|\text{g}\rangle$ is the ground and $|e\rangle$ is the excited state of the two-level system. We can then use our traditional number or Fock state notation to denote the states of the LC oscillator without any ambiguity. As an example, the state $|\text{g}, 3\rangle$ of the composite system would denote that the two-level system is in its ground state and that there are three photons in the LC oscillator. Finally, it is useful to recall that the operators from the different Hilbert spaces $\mathcal{H}_{\text{TLS}}$ and $\mathcal{H}_{\text{LC}}$ still only operate on the portion of the quantum state that is defined within that respective Hilbert space.

Now, we can return to the quantum Rabi model of (10.64) and expand the various terms in the interaction part of the Hamiltonian to consider them further. In particular, we get that

$$\hat{H} = -\frac{1}{2}\hbar\omega_{10}\hat{\sigma}_z + \hbar\omega_r \hat{a}^\dagger \hat{a} + \hbar g\left[\hat{a}\hat{\sigma}_+ + \hat{a}^\dagger\hat{\sigma}_- + \hat{a}^\dagger\hat{\sigma}_+ + \hat{a}\hat{\sigma}_-\right]. \tag{10.67}$$

To gain more insight into the relative importance of the various interaction terms, it is instructive to temporarily "factor out" the underlying time variation of the different operators. Recall that for the two-level system operators, we had that the free oscillations of $\hat{\sigma}_\pm$ will follow $\exp[\pm i\omega_{10}t]$. We have also previously done this kind of analysis for $\hat{a}$ and $\hat{a}^\dagger$, with the results being their respective free oscillations follow $\exp[-i\omega_r t]$ and $\exp[i\omega_r t]$. With this in mind, we write these dependencies into our Hamiltonian to get

$$\begin{aligned}\hat{H} = -\frac{1}{2}\hbar\omega_{10}\hat{\sigma}_z + \hbar\omega_r \hat{a}^\dagger \hat{a} + \hbar g\Big[&\hat{a}\hat{\sigma}_+ e^{i(\omega_{10}-\omega_r)t} + \hat{a}^\dagger\hat{\sigma}_- e^{-i(\omega_{10}-\omega_r)t} \\ &+ \hat{a}^\dagger\hat{\sigma}_+ e^{i(\omega_{10}+\omega_r)t} + \hat{a}\hat{\sigma}_- e^{-i(\omega_{10}+\omega_r)t}\Big].\end{aligned} \tag{10.68}$$

Just like with the semiclassical analysis, we see that if we have a near resonance situation (i.e., $\omega_r \approx \omega_{10}$) then terms with $\exp[\pm i(\omega_{10} - \omega_r)t]$ dependence will oscillate very slowly in comparison to terms with $\exp[\pm i(\omega_{10}+\omega_r)t]$ dependence. These quickly oscillating terms are usually highly suppressed in this situation, allowing us to neglect them in our analysis. We still refer to this as the *rotating wave approximation (RWA)* in this context. Making the RWA and embedding the time dependence back into the operators finally gives us

$$\hat{H} = -\frac{1}{2}\hbar\omega_{10}\hat{\sigma}_z + \hbar\omega_r\hat{a}^\dagger\hat{a} + \hbar g\big(\hat{a}\hat{\sigma}_+ + \hat{a}^\dagger\hat{\sigma}_-\big). \tag{10.69}$$

This Hamiltonian is one of the most widely-studied Hamiltonians in the design of quantum technology, and is referred to as the *Jaynes-Cummings model.*

10.3.2 Vacuum Rabi Oscillations

The Jaynes-Cummings model is so widely studied because it is very useful in describing many practical effects in the field of *cavity quantum electrodynamics (QED).* In cavity QED, one is typically interested in analyzing the quantum interactions between an (artificial) atom and an electromagnetic cavity. An electromagnetic cavity can take many different forms, but fundamentally it is a kind of resonator, just like our LC resonator. These systems have been investigated so thoroughly because they provide us with a well-controlled setting to investigate and harness fundamental properties of these quantum interactions. These properties form the foundation of many emerging quantum technologies because they can be used to generate and process quantum information. We will return to considering what one can do with quantum information later in the course. For now, we will look at one particularly important and simple effects described by the Jaynes-Cummings model – *vacuum Rabi oscillations.*

This is an analogue to the driven Rabi oscillations we considered earlier for a semiclassical system. There we saw that we could apply a classical driving field to the two-level system to cause the state of the two-level system to oscillate between being in the ground and excited state. Vacuum Rabi oscillations are similar, but now can occur with *no classical drive.* In particular, if we have the two-level system start in its excited state we will find that that the excitation in the system coherently oscillates back and forth between being in the two-level system and being in the LC oscillator. To describe this effect mathematically, we will follow a somewhat different approach to what we did when discussing driven Rabi oscillations.

In particular, we will begin by simplifying our analysis to assume that there is only a single quanta in our system and that our two subsystems are resonant such that $\omega_{10} = \omega_r = \omega_0$. Since there is only one quanta in our system, the only two states that we need to consider are $|\mathrm{g},1\rangle$ and $|e,0\rangle$. We can now find the matrix representation of the Jaynes-Cummings Hamiltonian given in (10.69) in terms of these two states, which is

$$[H] = \begin{bmatrix} \frac{1}{2}\hbar\omega_0 & \hbar g \\ \hbar g & \frac{1}{2}\hbar\omega_0 \end{bmatrix}, \tag{10.70}$$

where we have considered our first state to be $|\mathrm{g},1\rangle$ and our second state to be $|e,0\rangle$. We can now look for the eigenvectors and eigenvalues of this matrix, which we can then consider to be the energy eigenstates of our composite system. These can be found quite easily, and are

$$|\psi_\pm\rangle = \frac{1}{\sqrt{2}}(|e,0\rangle \pm |\mathrm{g},1\rangle), \tag{10.71}$$

$$E_{\pm} = \hbar\bigg(\frac{1}{2}\omega_0 \pm g\bigg). \tag{10.72}$$

We see that in our original basis of $|\mathrm{g}, 1\rangle$ and $|e, 0\rangle$ we had two *degenerate states* as evidenced by the identical diagonal elements in (10.70), but that due to the presence of the interaction (off-diagonal elements in the Hamiltonian) our new energy eigenstates are no long degenerate. This is usually referred to as having the interaction *lift* the degeneracy in the original basis, with a *level splitting* given by $2g$.

Now that we have the energy eigenstates of our composite system, we can use them to see the vacuum Rabi oscillation effect. From our earlier discussion, we wish to have our system begin in the excited state of the two-level system. We can describe this starting state as a superposition of our two energy eigenstates $|\psi_{\pm}\rangle$ as

$$|\psi(0)\rangle = \frac{1}{\sqrt{2}}\big(|\psi_+\rangle + |\psi_-\rangle\big) = |e, 0\rangle. \tag{10.73}$$

Recalling our knowledge of how superpositions of energy eigenstates evolve in time through the time-dependent Schrödinger equation, we find that our time dependent state will be

$$|\psi(t)\rangle = \frac{1}{\sqrt{2}}\Bigg(|\psi_+\rangle e^{-i(E_+/\hbar)t} + |\psi_-\rangle e^{-i(E_-/\hbar)t}\Bigg). \tag{10.74}$$

To see what the time dependence looks like in terms of our original states $|\mathrm{g}, 1\rangle$ and $|e, 0\rangle$ of the composite system, we can expand the energy eigenstates in (10.74) as

$$|\psi(t)\rangle = \frac{1}{2}\Bigg(|e, 0\rangle\Bigg\{e^{-i(E_+/\hbar)t} + e^{-i(E_-/\hbar)t}\Bigg\} + |\mathrm{g}, 1\rangle\Bigg\{e^{-i(E_+/\hbar)t} - e^{-i(E_-/\hbar)t}\Bigg\}\Bigg). \tag{10.75}$$

We can simplify this expression further by substituting in explicitly for the energy eigenvalues $E_{\pm}$. This allows us to get

$$\begin{aligned} |\psi(t)\rangle &= \frac{1}{2}e^{-i(\omega_0/2)t}\Bigg(|e, 0\rangle\Bigg\{e^{-igt} + e^{igt}\Bigg\} + |\mathrm{g}, 1\rangle\Bigg\{e^{-igt} - e^{igt}\Bigg\}\Bigg) \\ &= e^{-i(\omega_0/2)t}\Bigg(\cos(gt)|e, 0\rangle - i\sin(gt)|\mathrm{g}, 1\rangle\Bigg). \end{aligned} \tag{10.76}$$

We can finally look at the probabilities of finding the quanta of the system in $|\mathrm{g}, 1\rangle$ or $|e, 0\rangle$ as

$$P_e(t) = |\langle e, 0|\psi(t)\rangle|^2 = \cos^2(gt), \tag{10.77}$$

$$P_{\mathrm{g}}(t) = |\langle \mathrm{g}, 1|\psi(t)\rangle|^2 = \sin^2(gt). \tag{10.78}$$

As promised, the system oscillates back and forth between the quanta being in the two-level system or in the LC oscillator with a frequency given by the coupling strength g. Hence, stronger coupling between the subsystems allows for the swapping of excitations to occur more rapidly.

Although we restricted our analysis to only a single quanta in the composite system, a similar effect will be seen for larger numbers of quanta. In particular, because the two-level system can only absorb one quanta at a time the Jaynes-Cummings model can only describe transitions between "nearest-neighbor" excitation numbers in the LC oscillator. As a result, it is possible to repeat our prior analysis for various *subspaces* of possible quantum states that are able to interact with each other through the Jaynes-Cummings Hamiltonian. What one sees is that the degenerate states continue to be split, but that due to larger photon occupation numbers the splitting grows as the number of photons grows. This is illustrated in Fig. 10.4, which is often referred to as the *Jaynes-Cummings ladder* in analogy to the "ladder" of energy states described by a quantum harmonic oscillator. Regardless, the energy eigenstates still consist of symmetric and antisymmetric combinations of the original states of the composite system, like those described in (10.71). These states are often referred to as *dressed states*, with the original states of the uncoupled system referred to as *bare states*. We see that the dressed states, which are the true "excitations" of the coupled system, are no longer a photon or a two-level system, but instead are compositions of these notions. Hence, if our system is strongly coupled to the point where these effects are occurring, it can become difficult to try and "hang onto" our notions of the states of the individual subsystems in thinking about the physics of the coupled system.

10.4 Spontaneous and Stimulated Emission

We will now discuss two very important effects that arise due to the interactions between a circuit and artificial atom. Namely, spontaneous and stimulated emission. Although these effects can certainly happen in the context of a circuit and artificial atom, they are often more readily thought of in terms of an atom interacting with an electromagnetic field. We will continue to use our circuit picture for consistency, but some of the circuits we need to consider may be on the more "abstract" side.

10.4.1 Spontaneous Emission

We will discuss spontaneous emission in the context of time-dependent perturbation theory that we considered previously. We will now focus on adapting our time-dependent perturbation theory to the current situation of a quantum LC oscillator coupled to an artificial atom. For simplicity, we will still treat the artificial atom as a two-level system, although these considerations can be extended to handle the more complicated level structure of an actual artificial atom.

Now, following the time-dependent perturbation theory approach, we need to identify a part of the Hamiltonian as the "unperturbed" part and the remainder as the "perturbation". For our current system, it becomes useful to consider the "unperturbed" part of the Jaynes-Cummings Hamiltonian to be

$$\hat{H}_o = -\frac{1}{2}\hbar\omega_{10}\hat{\sigma}_z + \hbar\omega_r\hat{a}^\dagger\hat{a} \tag{10.79}$$

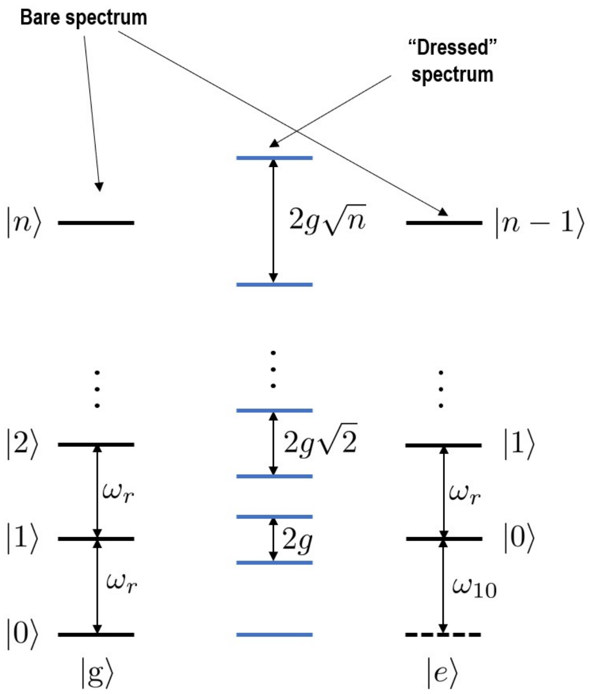


Figure 10.4: The generation of the Jaynes-Cummings ladder. On the left (right), is the energy levels for the composite system when the two-level system is in its ground (excited) state. We see that there are always two degenerate states at every "rung" of these two bare energy spectra. The interaction between these states through the Jaynes-Cummings model leads to a lifting of the degeneracy, with the amount of splitting dependent on the number of quanta in the system. The resulting Jaynes-Cummings ladder, which describes the spectra of the actual coupled system, is shown in the middle of the figure.

and the "perturbation" to be the interaction terms given by

$$\hat{H}_p = \hbar g\big(\hat{a}\hat{\sigma}_+ + \hat{a}^\dagger\hat{\sigma}_-\big). \tag{10.80}$$

The eigenstates of $\hat{H}_o$ that we will describe the perturbation processes through are then $|\text{g}, m\rangle$ and $|e, m'\rangle$, where m and m' will be the number of photons in the LC oscillator.

We can now follow the typical time-dependent perturbation theory process and find that the first-order change in the unperturbed states of the system will be given by

$$\dot{c}_q^{(1)} \approx \frac{1}{i\hbar}\sum_m c_m^{(0)} e^{-i(E_m-E_q)t/\hbar}\langle\psi_q|\hat{H}_p|\psi_m\rangle, \tag{10.81}$$

where E_m and E_q are the total energies of the complete state of the system (two-level system plus LC oscillator) described by $|\psi_m\rangle$ and $|\psi_q\rangle$, respectively. Now, because we are not able to "switch" off the interaction in this case (because it is not due to some external drive that we control), it becomes sensible to look at a transition rate in the manner of Fermi's golden rule. We can proceed with the same basic steps and find that for our particular case the transition rate from state $|\psi_m\rangle$ to state $|\psi_q\rangle$ is

$$w_{qm} = \frac{2\pi}{\hbar} |\langle \psi_q | \hat{H}_p | \psi_m \rangle|^2 \delta(E_m - E_q) \tag{10.82}$$

To describe the spontaneous emission process, we need to look at the case where the two-level system is in the excited state and there are no photons in the oscillator. Hence, we have that $|\psi_m\rangle = |e, 0\rangle$. We can now see how $\hat{H}_p$ operates on this state to gain insight into which output states will be relevant to consider. We find that

$$\hat{H}_p|\psi_m\rangle = \hbar g \hat{a}\hat{\sigma}_+|e,0\rangle + \hbar g \hat{a}^\dagger \hat{\sigma}_-|e,0\rangle = \hbar g|\mathrm{g},1\rangle. \tag{10.83}$$

Considering this, the state that we need to consider for $|\psi_q\rangle$ that this process will be able to create an appreciable transition probability to will be $|g, 1\rangle$. However, this will only be able to occur if the LC oscillator and two-level system are resonant. This ensures that $E_q = \hbar\omega_{10}$ will equal $E_m = \hbar\omega_r$ so that the delta function in (10.82) will be "non-zero". We then see that the transition rate for spontaneous emission is proportional to g^2.

Before continuing on, it is important to emphasize an important nuance about this effect. In particular, spontaneous emission is a purely quantum process. What we mean by this is that if we had not quantized our LC oscillator then there would be no $\hat{H}_p$ for the two-level system to decay through. Instead, it would be in an eigenstate of the two-level system and would then stay there indefinitely according to the time-dependent Schrödinger equation. Along these lines, we can interpret spontaneous emission occurring as being due to vacuum fluctuations in the LC oscillator. These vacuum fluctuations are constantly "interacting" with the two-level system, providing an avenue for emission to occur. In this perspective, we can actually think of spontaneous emission as being a kind of "stimulated emission" that is being stimulated through the vacuum fluctuations of the LC oscillator. Finally, it is worth noting that most of the light that we interact with on a day-to-day basis is actually due to spontaneous emission.

10.4.2 Stimulated Emission

We have already discussed stimulated emission at a high-level when we introduced time-dependent perturbation theory. We will now consider it for our current situation of a quantum LC oscillator coupled to a two-level system.

To describe the stimulated emission process, we need to look at the case where the two-level system is in the excited state and there is at least one photon in the oscillator. For simplicity, we will consider this particular situation, hence, $|\psi_m\rangle = |e, 1\rangle$. We can now see how $\hat{H}_p$ operates on this state to gain insight into which output states will be relevant to consider. We find that

$$\hat{H}_p|\psi_m\rangle = \hbar g \hat{a}\hat{\sigma}_+|e,1\rangle + \hbar g \hat{a}^\dagger \hat{\sigma}_-|e,1\rangle = \hbar\sqrt{2}g|\mathrm{g},2\rangle. \tag{10.84}$$

Considering this, the state that we need to consider for $|\psi_q\rangle$ that this process will be able to create an appreciable transition probability to will be $|g, 2\rangle$. We note again that for the delta function in (10.82) to be "non-zero" we need the two-level system and LC oscillator to be resonant.

It is interesting to note that the transition rate for stimulated emission is proportional to $2g^2$ when there was only a single photon in the LC oscillator. We can readily see from the properties of the ladder operators that as we increase the number of photons in the LC oscillator, this will continue to increase the transition rate to mg^2 for $m-1$ photons beginning in the LC oscillator. We can compare this to the spontaneous emission rate of g^2 to see that the presence of a photon has actually *enhanced* the probability of the two-level system emitting a photon. This is the origin of the name *stimulated emission*, since the presence of the photon is stimulating the two-level system to be more likely to emit another photon.

We are now in a position to have a basic understanding of how a laser works. As mentioned previously, the term laser is actually an acronym that stands for "light amplification by stimulated emission of radiation". The idea is that if we put a single photon into an oscillator that is connected to a large number of (artificial) atoms (shown in Fig. 10.5) in their excited states then stimulated emission will cause a cascade of emission processes to occur. For instance, we can consider the first photon to interact with a particular atom that stimulates the emission of a second photon. These two photons can then interact with two additional atoms, causing 2 more photons to be emitted, bringing us to a total of 4 photons. This process can continue until we have a very large number of photons.

Although this is the basic idea, there are many other important considerations that go into achieving successful laser operation (often referred to as *lasing*). For instance, once an atom has emitted a photon it is now in the ground state. As a result, it is possible for it to *absorb* a photon out of our oscillator and reduce our total number of photons. Hence, there is a kind of "damping" to the laser amplification due to the natural interaction with the atoms in the system. Considering this, it is essential that we are able to maintain our system in an overall state where there are more atoms in the excited state than the ground state. This is referred to as *population inversion*.

How do we maintain population inversion if we are constantly stimulating the atoms to emit photons? The answer lies in *pumping* the atoms with some other energy source. As is the case with any amplification process, there is "no free lunch". We have to expend energy in some other form to ensure that the atoms are in their excited state, and we must continue doing this during the laser operation to maintain population inversion. The exact way that this pumping is done depends on the specific design of the laser, and many different approaches have been developed. Overall, there is usually a delicate balance in achieving lasing that requires careful engineering of a system to ensure successful operation.

Although many different pumping approaches have been developed, it is often the case that the most useful "atoms" for use in a laser will actually require more than two levels. The idea is that we can pump the systems into a third level that is higher than the excited state that we will interact with for stimulated emission. The most useful third level is one that has some other natural relaxation process that will occur quickly to cause the system to drop into the desired excited state. Having this third state not be resonant with our stimulated emission process is beneficial so that whatever we use for our pump will not be in the same spectral location as the signal emitted from the laser. This results in a purer laser signal, which is often a priority for many applications that utilize lasers. In reality, actual physical systems may not have such a convenient "third level" to be used in the laser, and so more complicated schemes may be utilized (e.g., ones with four levels, etc.).

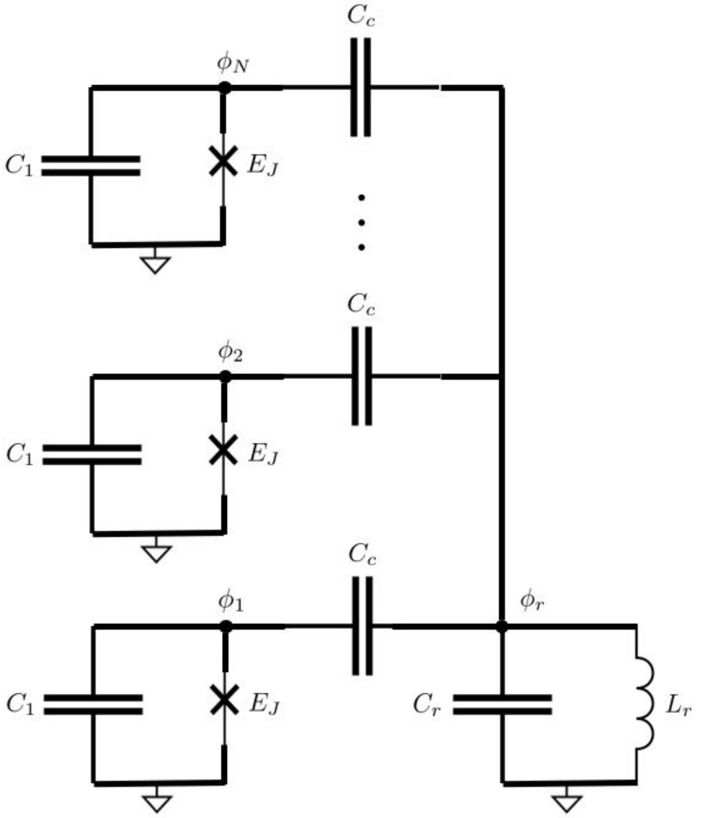


Figure 10.5: Schematic of a "circuit laser" where an LC oscillator is coupled to a large number of artificial atoms.

10.4.3 Purcell Effect

Our treatment of spontaneous emission made it clear that this rate depends on the interaction Hamiltonian between the artificial atom and the LC oscillator. However, this link was not always so obvious, and so for many years it was originally thought that the spontaneous emission rate of an atom was an *intrisic* property of the system. That this notion persisted for quite some time was mostly due to the fact that spontaneous emission was mainly able to be investigated experimentally in the form of a collection of atoms in "free space" (i.e., more or less in empty air rather than in a heavily-engineered system). However, later studies and theoretical analysis made it clear that the spontaneous emission rate is actually an *extrinsic* property that depends on details of the system that the atom is coupled to. This is often referred to as the *Purcell effect*, and is an important effect because it makes it clear that we can engineer our systems in particular ways to suppress or enhance the spontaneous emission rate.

Now, we will not go into detail on how to derive the specific form of the Purcell effect for a particular system. Instead, we will simply comment on the overall result in the context of circuits and look at some simple examples. Overall, one of the very important aspects of

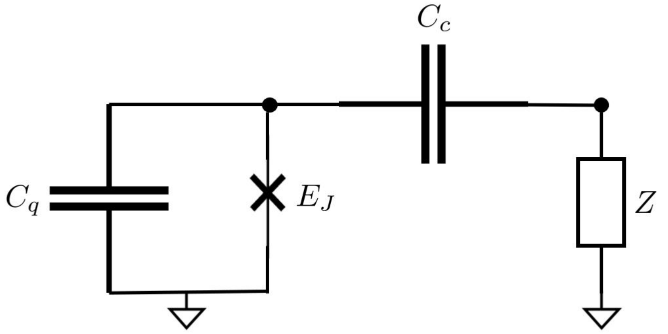


Figure 10.6: Generic representation of a charge qubit capacitively coupled to an environment represented by Z.

the Purcell effect is that it links the spontaneous emission rate (a purely quantum effect) to *classical* properties of the system the atom is coupled to. Typical terminology would refer to the system the atom is coupled to as the *environment.*

In the case of a charge qubit with qubit capacitance C_q, we can consider the environment to be represented as a general impedance Z that the qubit is capacitively coupled to, as illustrated in Fig. 10.6. One then finds that the relaxation time T_1 of the qubit (which is the inverse of the spontaneous emission rate) is given by the classical decay time of the qubit capacitance, i.e., $R_{\text{eq}}C_q$, where $R_{\text{eq}} = 1/\text{Re}\{Y_{\text{eq}}\}$ and Y_{eq} is the equivalent admittance of the environment and coupling capacitance across the terminals of C_q [31].

As an example, we can consider that the environment is simply a resistor R_0. This is a good representation of an "open" environment from the perspective of the qubit because the dissipation of the resistor means that the excitation of the qubit will effectively be "lost" to the environment once it has been emitted. If the environment were instead an LC oscillator, we would have a more complicated situation where the quanta emitted into the LC oscillator could potentially be absorbed back by the charge qubit, like in the vacuum Rabi oscillations we considered previously. It should be noted, however, that the case of vacuum Rabi oscillations occur when the qubit and LC oscillator are *strongly* coupled to each other. In that case, the assumptions built into our time-dependent perturbation theory treatment arriving at the spontaneous emission rate through Fermi's golden rule would have broken down so that our current analysis would no longer be applicable.

For now, we just need to determine the equivalent admittance seen by the qubit capacitance. For the circuit in Fig. 10.6 with $Z = R_0$, we would have Y_{eq} being given by the series combination of the coupling capacitance C_c and the terminating resistance R_0. The equivalent impedance in this case can be easily found to be

$$Z_{\text{eq}} = R_0 + \frac{1}{i\omega C_c}, \tag{10.85}$$

so that the equivalent admittance is

$$Y_{\text{eq}} = 1/Z_{\text{eq}} = \frac{i\omega C_c}{1 + i\omega R_0 C_c}. \tag{10.86}$$

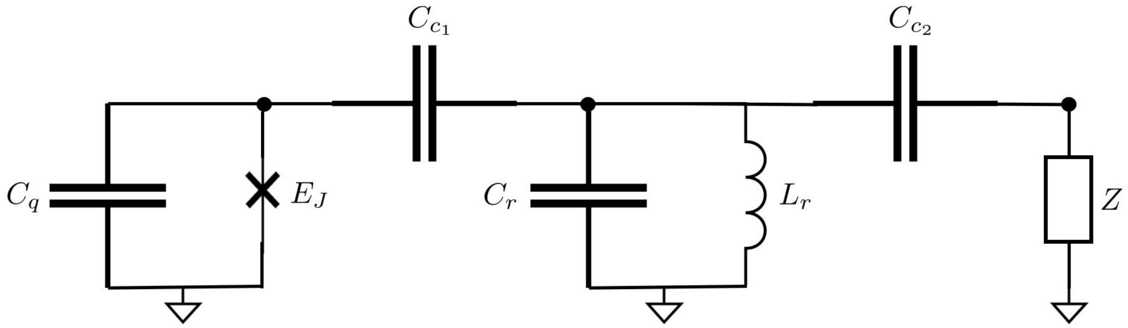


Figure 10.7: Schematic of a charge qubit capacitively coupled to an LC oscillator that is then coupled to the environment.

With a little more algebra, we can find that the real part of the admittance is then

$$\mathrm{Re}\{Y_{\mathrm{eq}}\} = \frac{\omega^2 R_0 C_c^2}{1 + (\omega R_0 C_c)^2}. \tag{10.87}$$

To arrive at a bit simpler result, it is useful to consider some relevant values for actual physical systems. In these, the qubit capacitance will typically be in the tens of femtofarad range, with the coupling capacitance some fraction of that. The environmental resistance is typically 50Ω, and the frequency is usually in the ones of gigahertz range. In this case, $(\omega R_0 C_c)^2 \ll 1$, so we can usually approximate $\mathrm{Re}\{Y_{\mathrm{eq}}\}$ as

$$\mathrm{Re}\{Y_{\mathrm{eq}}\} \approx \omega^2 R_0 C_c^2. \tag{10.88}$$

The resulting relaxation time is then

$$T_1 \approx \frac{C_q}{\omega^2 R_0 C_c^2}. \tag{10.89}$$

Let's now consider what happens to the relaxation time when we couple the charge qubit to an LC oscillator that is then coupled to the environment, as illustrated in Fig. 10.7. It is relatively straightforward to analyze this circuit, but the formulas become rather tedious to write down. Considering this, we will instead present numerical results of T_1 as a function of frequency for an LC oscillator with resonant frequency of 5 GHz. These results are shown in Fig. 10.8, where they are compared to the results of the qubit capacitively coupled directly to the environment.

It is seen that if the qubit frequency is away from the resonance of the LC oscillator than the relaxation time becomes much higher than the case of the qubit capacitively coupled to the environment. This is equivalent to *suppressing* the spontaneous emission rate, which leads to the qubit staying in its excited state for a longer period of time. This is often very advantageous for many quantum technologies, especially quantum computers, because we do not often want our qubit to spontaneously relax to its lower state in the middle of using it to try and process some kind of quantum information. However, for other technologies

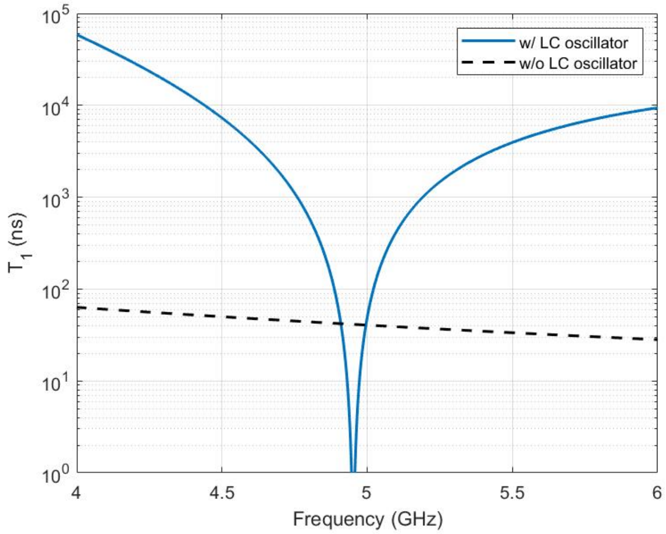


Figure 10.8: Relaxation time as a function of qubit frequency when it is coupled directly to the environment or to the environment through an LC oscillator. Parameters for the circuits are $R_0 = 50\Omega$, $C_q = 50\,\text{fF}$, $C_{c_1} = 5\,\text{fF}$, $C_{c_2} = 50\,\text{fF}$, $C_r = 2.8\,\text{pF}$, and $L_r = 0.362\,\text{nH}$.

it can be advantageous to increase the spontaneous emission rate (e.g., if this emission will act as a source for a later process). We see that if we bring the qubit closer to resonance with the LC oscillator than we can *enhance* the spontaneous emission rate. However, we must be careful with our operating parameters to ensure that the system does not become so strongly coupled to the LC oscillator that other effects, like vacuum Rabi oscillations, become prevalent.

10.5 Practice Problems

1. Consider a circuit like that shown in Fig. 10.9 that uses a phenomenological model to describe a classical voltage source (due to the large capacitance C_2) that drives a charge qubit. Following a semiclassical treatment of this problem, we found that the Hamiltonian could eventually be written as

$$\hat{H}_{\text{TLS,RWA}} = -\frac{1}{2}\hbar\delta\hat{\sigma}_z + \frac{1}{2}\hbar g V_x(t)\hat{\sigma}_x + \frac{1}{2}\hbar g V_y(t)\hat{\sigma}_y \tag{10.90}$$

where δ is the detuning between the qubit's transition frequency and the oscillating drive of the classical voltage source. If we only consider drives along the V_y channel with detuning $\delta = 0$, use the time-dependent Schrödinger equation to formulate appropriate ordinary differential equations to solve for the time evolution of the qubit's expansion coefficients $c_0(t)$ and $c_1(t)$. For a constant envelope function of $V_y(t) = V_{y,0}$, solve for the driven Rabi oscillations that will occur for this qubit if it starts in its ground state. Use these solutions to demonstrate that this drive will result in counterclockwise rotations of the qubit state vector on the Bloch sphere around the $+y$-axis.

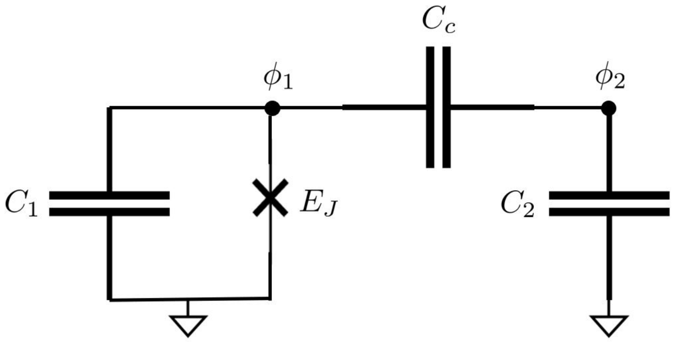


Figure 10.9: Circuit schematic for a large capacitance (C_2) capacitively-coupled to a charge qubit.

2. This problem will consider the Jaynes-Cummings model Hamiltonian given by

$$\hat{H} = -\frac{1}{2}\hbar\omega_{10}\hat{\sigma}_z + \hbar\omega_r\hat{a}^\dagger\hat{a} + \hbar g\big(\hat{a}\hat{\sigma}_+ + \hat{a}^\dagger\hat{\sigma}_-\big). \tag{10.91}$$

   (a) The operator

$$\hat{N} = \hat{\sigma}_+\hat{\sigma}_- + \hat{a}^\dagger\hat{a} \tag{10.92}$$

   is the excitation number operator for the Jaynes-Cummings model. Show that the expectation value of this operator for any valid state of the Jaynes-Cummings model returns the total number of excitations in the system (e.g., the TLS in its excited state and the oscillator with 4 quanta would have an excitation number of 5; if the TLS was in its ground state the total excitation would then be 4).

   (b) Show that $[\hat{N}, \hat{H}] = 0$. From the perspective of the Heisenberg picture, what does this mean about the Jaynes-Cummings model and the excitation number as a function of time?

3. In practical quantum computing applications, it is common to couple a transmon qubit to a microwave transmission line resonator to facilitate measuring the state of the qubit. The stronger the coupling is between the two systems, the quicker the state measurement can be made (which is key in building toward error-correctable quantum computers). Unfortunately, stronger coupling between the two systems also leads to stronger decay of the qubit state through spontaneous emission due to the Purcell effect, limiting the length of time an algorithm can be reliably executed for. This tradeoff can be circumvented by introducing an additional circuit between the microwave transmission line resonator and the "external environment". This additional circuit acts as a filter to protect the qubit from spontaneous emission. In this problem, you will design such a filtering circuit.

   (a) To start, write a computer program to compute the relaxation time of the qubit for the circuit shown in Fig. 10.10 from 4 to 6 GHz.

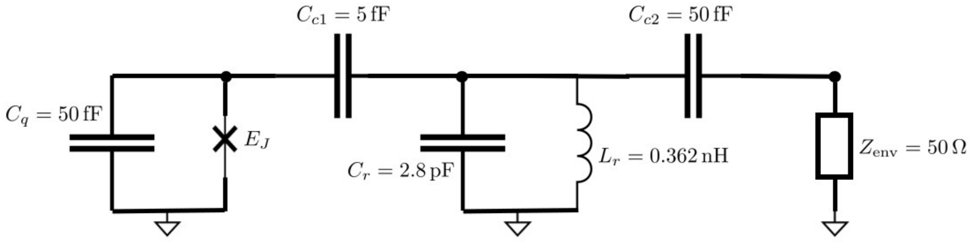


Figure 10.10: Equivalent circuit model of a transmon qubit capacitively coupled to a microwave transmission line resonator. The transmission line resonator is modeled as a single LC tank circuit (this approximation is not always reasonable) that is capacitively coupled to an external environment represented by Z_{env}.

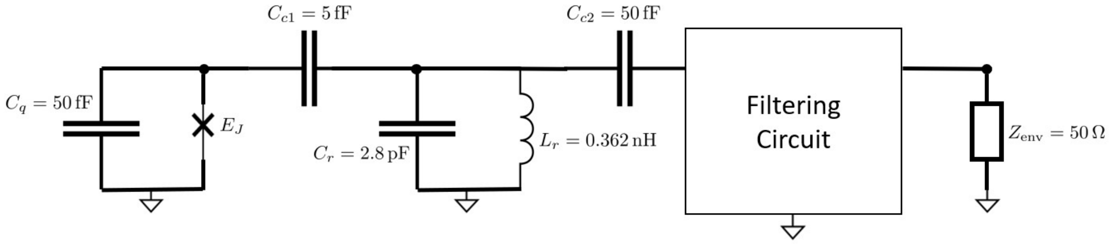


Figure 10.11: Modified circuit model of a transmon qubit capacitively coupled to a microwave transmission line resonator to include an additional filtering circuit.

(b) Now, determine an appropriate circuit that can be placed in the "filtering circuit" black box of Fig. 10.11 to lower the qubit's spontaneous emission rate. To achieve this, your filtering circuit should lower the effective impedance seen by the qubit at the frequency of the qubit, but should otherwise have a minimal effect on a signal originating from the qubit being able to "reach" the environmental impedance at the frequency of the intermediary resonator (this is necessary to maintain the ability to measure the qubit state). For this problem, assume that the qubit has a nominal operating frequency of 4.3 GHz. To keep the results more reasonable, assume that any resonator circuits you use in your filter have some small amount of parasitic resistance to not have "zero" or "infinite" response functions at the qubit operating frequency (a value of $1\,\mathrm{m}\Omega$ is usually reasonable for a series RLC circuit and a value of $1\,\mathrm{M}\Omega$ for a parallel RLC). Compute and plot the relaxation time of the qubit for your solution from 4 to 6 GHz and compare it to the answer from (a).

Note: There are many valid solutions to this problem. Try and focus on the *simplest* solutions.

4. For this problem, consider the Jaynes-Cummings model described by

$$\hat{H} = -\frac{1}{2}\hbar\omega_0\hat{\sigma}_z + \hbar\omega_0\hat{a}^\dagger\hat{a} + \hbar g\big(\hat{a}\hat{\sigma}_+ + \hat{a}^\dagger\hat{\sigma}_-\big). \tag{10.93}$$

You have been told that the total number of quanta in the system is equal to 3 and that at $t = 0$ the two-level system can be considered to be in state $|g\rangle$ (that is the ground state, located at the north pole of the Bloch sphere). Describe the trajectory that the state of the two-level system will take on the Bloch sphere for this system. Further, what will be the smallest value of t that the two-level system will be found to be at the south pole of the Bloch sphere if $g/(2\pi) = 100\,\text{MHz}$?

Chapter 11

Quantum Information and Its Applications

11.1 Density Matrix Formalism

Up to this point, we have been considering quantum systems in a manner where the only "randomness" involved was due to the inherent statistical aspects of quantum physics. However, there are many situations in classical physics and engineering where we must rely on statistical methods to describe systems because there is a "classical" kind of randomness in the system. These methods often are used when we are in a position where it is not possible for us to have a definite description of all aspects of our system; e.g., fluctuations of electrons in the atmosphere causing distortion to electromagnetic signals or a large ensemble of atoms in a gas interacting with one another. Just as these statistical methods were needed for complicated classical systems, as we add more and more complexity to the quantum systems we consider we find that we also need to make use of these statistical tools. The main approach used to connect the tools of statistical mechanics with quantum mechanics is the *density operator*, or more commonly just referred to as the *density matrix*. This formalism for quantum systems is a powerful technique that allows us to naturally blend the "classical" statistical methods with the underlying statistical interpretation of quantum mechanics.

We will now introduce the basics of the density matrix approach to describing quantum systems. As you learn more about quantum physics, you will increasingly encounter this approach to analyzing systems. We will use it in this course mainly to make more concrete certain concepts of "coherence" and "decoherence" that are key in beginning to understand quantum information, and play an important role in applications that seek to generate and process this new kind of information.

11.1.1 Pure and Mixed States

As alluded to previously, the only randomness we have been considering up to this point has been related to the outcomes of performing a quantum mechanical measurement on our system of interest. For instance, if we consider a two-level system, we have expressed the general state of such a system as

$$|\psi\rangle = c_0|0\rangle + c_1|1\rangle. \tag{11.1}$$

The randomness occurs when we take a measurement of this system, with the expectation that when we make our measurement we will find the system in state $|0\rangle$ with probability $|c_0|^2$ and in state $|1\rangle$ with probability $|c_1|^2$. If we consider this system from the perspective of the time-dependent Schrödinger equation, we know that each of the states involved in superposition $|\psi\rangle$ will evolve with the characteristic frequency of $\exp[-i(E_m/\hbar)t]$, where E_m is the energy of the mth state. We can then look at the expectation value for different measurable quantities and expect to see a time dependence due to interaction of these different time dependencies. This kind of situation is referred to as a *pure state*. In essence, the system is completely "definite" in the sense that we have a complete description of it and the only randomness occurs due to the randomness of quantum mechanical measurements.

Although pure states are certainly of interest, they are not the only kind of situation we may encounter. For instance, it is also reasonable for us to want to describe the statistical results of the situation where we have a large ensemble of two-level systems that we randomly prepare in the $|0\rangle$ or $|1\rangle$ state according to some probability (e.g., a 50-50 split). In this case, our two-level system is in an energy eigenstate so that if we evaluate the expectation value of any particular measurable quantity for any of our systems we will find that they do not vary in time. If we then take the average of these measurements *for all* of the randomly-prepared two-level systems, we will also see no oscillation in time (this kind of operation is referred to as an *ensemble average*). This situation is then distinct from the idea of a pure state, since our statistical preparation approach of the two-level systems cannot be described as a linear superposition of states like in (11.1). We refer to this situation as a *mixed state*.

The purpose of the density matrix is to provide a unified approach that can represent both mixed and pure states, and seamlessly give the correct statistical results without requiring any special treatment for one situation versus the other. As a result, it represents a very powerful tool for describing complex quantum systems. However, it does require us to introduce some new techniques to accomplish these goals.

11.1.2 Density Operator and Density Matrix

As alluded to by the name, we will need to introduce a new kind of "operator" to describe the state of our system. It is important to emphasize that the density operator is not like the operators we have worked with up to this point, which were often related to a measurable quantity and regardless were always defined in terms of how they *operated* on the state of the system. Since the density operator will be *describing* the state of our system, it is obviously not able to *operate* on the state of the system in the usual sense of our operators. So why do we bother with calling the density operator an operator if it is so different? The reason is because the density operator takes the mathematical form of

$$\rho = \sum_n P_n|\psi_n\rangle\langle\psi_n|, \tag{11.2}$$

where P_n represents the "classical" probability describing the preparation of the mixed state. Hence, we see that (11.2) looks like a special kind of bilinear expansion of an operator, and is why it is then referred to as an operator. As is commonly done, we leave off the hat for ρ to emphasize its unique role and not confuse it with our prior notion of an operator. Before moving on, it is important to also note that the pure states $|\psi_n\rangle$ used in describing our mixed state need not be orthogonal (i.e., $\langle\psi_m|\psi_n\rangle$ may or may not equal 0, depending on the context).

To see how to use the density matrix to compute ensemble averages of a physical observable, let's consider a general case where the states $|\psi_n\rangle$ making up the density operator are *not* orthonormal. Our first step then is to find the matrix representation of the density operator within some orthonormal basis. If we denote these basis vectors as $|\phi_k\rangle$, then we can expand each of the pure states $|\psi_n\rangle$ as

$$|\psi_n\rangle = \sum_k c_k^{(n)}|\phi_k\rangle. \tag{11.3}$$

This representation can then be substituted into (11.2) to express the density operator as

$$\rho = \sum_n P_n\bigg(\sum_k c_k^{(n)}|\phi_k\rangle\bigg)\bigg(\sum_\ell (c_\ell^{(n)})^*\,\langle\phi_\ell|\bigg). \tag{11.4}$$

We can rearrange terms to lump all the scalars together to get

$$\rho = \sum_{k,\ell}\bigg(\sum_n P_n c_k^{(n)}(c_\ell^{(n)})^*\bigg)|\phi_k\rangle\langle\phi_\ell|. \tag{11.5}$$

From this expression, we can clearly determine the matrix representation of ρ using the $|\phi_k\rangle$ basis. In particular, we see that a given matrix element is

$$\rho_{k\ell} = \langle\phi_k|\rho|\phi_\ell\rangle = \sum_n P_n c_k^{(n)}(c_\ell^{(n)})^*. \tag{11.6}$$

We can recognize that this matrix element is the ensemble average of the product of coefficients involved. A useful shorthand notation for this is

$$\overline{c_k c_\ell^*} = \sum_n P_n c_k^{(n)}(c_\ell^{(n)})^*. \tag{11.7}$$

Now, since the density operator is essentially always represented in this way, it is often referred to as the density matrix instead. Although there is technically a distinction, it is so often unimportant that the terms density operator and density matrix are used interchangeably.

We can now look at a few properties of the density matrix. First, we can readily identify from the expression for the matrix elements given in (11.6) that the density matrix (and, so also the density operator) is Hermitian, i.e., $\rho_{k\ell} = \rho_{\ell k}^*$. We can also look at the different matrix elements and ask the question of what these quantities represent. Looking at the diagonal elements first, we see that this gives us the ensemble average (i.e., the overall probability) of finding our system in a particular state. This is the case because we can recognize for ρ_{kk} that the $c_k^{(n)}(c_k^{(n)})^*$ term in (11.6) corresponds to the probability of finding the nth pure state $|\psi_n\rangle$ in the basis state $|\phi_k\rangle$. Summing this probability multiplied by the P_n's then gives the ensemble average of the overall system being in the basis state $|\phi_k\rangle$, as mentioned earlier.

We can also note that the sum of the diagonal elements of the density matrix is then equal to 1. Summing the diagonal elements of a matrix is an important tool in mathematical analysis in its own right, and is known as the *trace* of the matrix. For the density matrix, we denote this as

$$\mathrm{tr}\{\rho\} = \sum_k \rho_{kk} = \sum_k \sum_n P_n |c_k^{(n)}|^2. \tag{11.8}$$

Now, because the pure states must be normalized, we have that $\sum_k |c_k^{(n)}|^2 = 1$ for each n. We then have that

$$\mathrm{tr}\{\rho\} = \sum_n P_n = 1, \tag{11.9}$$

since the mixed state probabilities must also sum to 1 to be well-defined probabilities.

Returning to our discussion about the meaning of the different matrix elements, we turn our attention to the off-diagonal elements. Interpreting the off-diagonal elements is more subtle, but these terms provide very important insight into the physics of a system. To start understanding these terms better, let's look at a simple (but important) use of the density matrix. In particular, we will consider the density matrix for the pure state $|\psi\rangle = c_0|0\rangle + c_1|1\rangle$. If we use the states $|0\rangle$ and $|1\rangle$ as our orthonormal basis for expanding the density matrix, then we find that the density matrix for this pure state will be

$$\rho = \begin{bmatrix} |c_0|^2 & c_0 c_1^* \\ c_0^* c_1 & |c_1|^2 \end{bmatrix}. \tag{11.10}$$

At this stage, the important point to note is that the diagonal elements of the density matrix are real and purely positive, while the off-diagonal elements are complex-valued. Since these off-diagonal elements are complex-valued, they carry important information about the relative phase between the states. As a result, they are sometimes informally referred to as "coherences". Additionally, the complex-valued nature allows the summation of terms like $c_0 c_1^*$ for *different* pure states to *interfere* with one another. If we have a mixed state that is significantly "random", we would expect the summations of all the different $c_0 c_1^*$ terms to approximately average to 0. Hence, these off-diagonal elements also provide us with information about whether our system is still *coherently* evolving as a linear superposition of states or if it has devolved into a completely random (often called *incoherent*) mixed state. This is an extremely important measure when considering the processing of quantum information in a real system, and we will return to discussing it later.

However, this conversation also raises the interesting question of whether we can tell if a density matrix corresponds to a pure or mixed state if we are only given the density matrix and no other information. Fortunately, it is possible to make this distinction, which is accomplished by looking at the trace of ρ^2. If we have a density matrix that corresponds to a specific pure state $|\psi\rangle$, then we can quickly find that

$$\rho^2 = |\psi\rangle\langle\psi|\psi\rangle\langle\psi| = |\psi\rangle\langle\psi| = \rho. \tag{11.11}$$

As a result, we see that

$$\mathrm{tr}\{\rho^2\} = \mathrm{tr}\{\rho\} = 1. \tag{11.12}$$

For a mixed state, we will find that $\mathrm{tr}\{\rho^2\} < 1$. Hence, one can evaluate $\mathrm{tr}\{\rho^2\}$ to determine if a density matrix represents a pure or mixed state.

Next, we need to finish determining how we actually *use* the density matrix to compute something like the ensemble average of a particular measurable quantity associated with an operator $\hat{A}$. What we will find is that this can be accomplished by evaluating $\mathrm{tr}\{\rho\hat{A}\}$, but before we can do this we will first need to understand what the diagonal elements of the product $\rho\hat{A}$ "look" like. Considering this, we begin by noting that

$$\rho\hat{A} = \sum_{k,\ell}\left(\sum_n P_n c_k^{(n)}(c_\ell^{(n)})^*\right)|\phi_k\rangle\langle\phi_\ell|\hat{A}. \tag{11.13}$$

A specific diagonal element of $\rho\hat{A}$ is then

$$\langle\phi_q|\rho\hat{A}|\phi_q\rangle = \sum_{k,\ell}\left(\sum_n P_n c_k^{(n)}(c_\ell^{(n)})^*\right)\langle\phi_q|\phi_k\rangle\langle\phi_\ell|\hat{A}|\phi_q\rangle. \tag{11.14}$$

Using the orthonormality of the $|\phi_k\rangle$ basis, we can replace $\langle\phi_q|\phi_k\rangle$ as δ_{qk} to simplify (11.14) to be

$$\langle\phi_q|\rho\hat{A}|\phi_q\rangle = \sum_{\ell}\left(\sum_n P_n c_q^{(n)}(c_\ell^{(n)})^*\right)\langle\phi_\ell|\hat{A}|\phi_q\rangle. \tag{11.15}$$

We can now evaluate the trace of $\rho\hat{A}$ as

$$\mathrm{tr}\{\rho\hat{A}\} = \sum_q\langle\phi_q|\rho\hat{A}|\phi_q\rangle = \sum_q\sum_\ell\left(\sum_n P_n c_q^{(n)}(c_\ell^{(n)})^*\right)\langle\phi_\ell|\hat{A}|\phi_q\rangle. \tag{11.16}$$

We can regroup terms in a more suggestive way to see that

$$\mathrm{tr}\{\rho\hat{A}\} = \sum_n P_n\left(\sum_\ell(c_\ell^{(n)})^*\langle\phi_\ell|\right)\hat{A}\left(\sum_q c_q^{(n)}|\phi_q\rangle\right). \tag{11.17}$$

We can recognize the terms in the large parentheses as being the basis expansion of the nth pure state via (11.3), so that we can now write (11.17) as

$$\mathrm{tr}\{\rho\hat{A}\} = \sum_n P_n\langle\psi_n|\hat{A}|\psi_n\rangle. \tag{11.18}$$

We see that the right-hand side is the ensemble average of the expected value of $\hat{A}$ for each pure state $|\psi_n\rangle$, and so $\mathrm{tr}\{\rho\hat{A}\}$ will indeed calculate this important statistic for us. We also see that if the density matrix happened to be for a pure state $|\psi\rangle$ so that there would be only a single P_n that would then equal 1 the result of (11.18) then matches what we already know, i.e., that $\langle A\rangle = \langle\psi|\hat{A}|\psi\rangle$. Hence, as promised, the density matrix formalism will still work for everything that we have done up to this point in terms of pure states.

Now that we can evaluate the desired statistics, the next piece of machinery needed for the density matrix formalism is to be able to compute the time evolution of the density matrix. For a pure state, we used the time-dependent Schrödinger equation

$$\hat{H}|\psi_n\rangle = i\hbar\frac{\partial}{\partial t}|\psi_n\rangle. \tag{11.19}$$

We can now generalize this result to a mixed state represented by the density matrix. We will find that the equation of motion looks very similar to the Heisenberg equations of motion for a Heisenberg picture operator, but that there is a small difference.

To see this, we start by expressing the pure state in (11.19) using the orthonormal basis set $|\phi_\ell\rangle$ to get

$$i\hbar\sum_\ell \dot{c}_\ell^{(n)}|\phi_\ell\rangle = \sum_\ell c_\ell^{(n)}\hat{H}|\phi_\ell\rangle. \tag{11.20}$$

We can now take the inner product of this equation with $|\phi_k\rangle$ to get

$$i\hbar\dot{c}_k^{(n)} = \sum_\ell c_\ell^{(n)}H_{k\ell}, \tag{11.21}$$

where $H_{k\ell} = \langle\phi_k|\hat{H}|\phi_\ell\rangle$ is a matrix element of the Hamiltonian operator. We can also take the complex conjugate of (11.21) to get

$$-i\hbar(\dot{c}_k^{(n)})^* = \sum_\ell (c_\ell^{(n)})^* H_{\ell k}, \tag{11.22}$$

where we noted that due to the hermiticity of the Hamiltonian we have $H_{k\ell}^* = H_{\ell k}$. Next, we look at the time derivative of a density matrix element, which we see from (11.6) is

$$\dot{\rho}_{k\ell} = \sum_n P_n\big[c_k^{(n)}(\dot{c}_\ell^{(n)})^* + (c_\ell^{(n)})^*\,\dot{c}_k^{(n)}\big]. \tag{11.23}$$

We can substitute in from (11.21) and (11.22) after changing the dummy summation indices as needed to get

$$\dot{\rho}_{k\ell} = \frac{i}{\hbar}\sum_n P_n\left[c_k^{(n)}\sum_r (c_r^{(n)})^* H_{r\ell} - (c_\ell^{(n)})^*\sum_s c_s^{(n)}H_{ks}\right]. \tag{11.24}$$

Regrouping terms then gives

$$\dot{\rho}_{k\ell} = \frac{i}{\hbar}\left[\sum_r\left(\sum_n P_n c_k^{(n)}(c_r^{(n)})^*\right)H_{r\ell} - \sum_s H_{ks}\left(\sum_n P_n c_s^{(n)}(c_\ell^{(n)})^*\right)\right]. \tag{11.25}$$

We can recognize the terms in the large parentheses as being the definition of density matrix elements from (11.6). Using this shorthand, we then see that (11.25) can be written as

$$\dot{\rho}_{k\ell} = \frac{i}{\hbar}\left[\sum_r \rho_{kr}H_{r\ell} - \sum_s H_{ks}\rho_{s\ell}\right]. \tag{11.26}$$

We can next note that terms like

$$\sum_r \rho_{kr} H_{r\ell} \tag{11.27}$$

correspond to evaluating a term in the matrix-matrix multiplication of $\rho\hat{H}$. In the case of (11.27), this corresponds to the $k\ell$ element of the matrix $\rho\hat{H}$, which we denote as $(\rho\hat{H})_{k\ell}$. Using this notation, we see that (11.26) is

$$\dot{\rho}_{k\ell} = \frac{i}{\hbar}\left[(\rho\hat{H})_{k\ell} - (\hat{H}\rho)_{k\ell}\right] = \frac{i}{\hbar}[\rho, \hat{H}]_{k\ell}. \tag{11.28}$$

Since this holds for a single generic matrix element $k\ell$, it holds for them all, so that we can write the general result as

$$\dot{\rho} = \frac{i}{\hbar}[\rho, \hat{H}]. \tag{11.29}$$

This is sometimes referred to as the *von Neumann equation*, or the *Liouville-von Neumann equation* due to its similarity to the Liouville equation in classical statistical mechanics.

We can use (11.29) to evaluate the time evolution of the density matrix just as we would use the time-dependent Schrödinger equation for a particular pure state. Recalling the Heisenberg equation of motion for an operator is

$$\frac{d}{dt}\hat{X}_H(t) = -\frac{i}{\hbar}[\hat{X}_H(t), \hat{H}], \tag{11.30}$$

we see that this is very similar to that of (11.29). However, there is a sign difference, which must be included to get the correct result in the end. It is also important to emphasize that when we work with (11.29) we are treating the density matrix in the Schrödinger picture. It is in fact possible to use the density matrix formalism within the Heisenberg picture, but we will not concern ourselves with this in this particular course.

11.2 Introduction to Quantum Information

One of the central areas of electrical engineering is *information theory*. This discipline concerns itself with the mathematical study of how we can represent "information" (in an abstract sense), transmit it, and process it. In the classical case, we think of information as being encoded in a binary form as a sequence of bits that take a value of 0 or 1, where we do not concern ourselves with the underlying physical system that we have used to represent this bit. One of the fundamental reasons information theory is so important is because it focuses on developing engineering solutions to process information in non-ideal systems. Many results of information theory come in the form of performance bounds for various practical situations, like the famous noiseless- and noisy-channel coding theorems of Claude Shannon. These results have widespread implications on the design of modern communication systems and error correcting codes, as well as touch on topics like why we can compress data files (e.g., .zip files) and then recover the full information at a later time.

Quantum information theory is concerned with many of the same kinds of questions of classical information theory, but instead looks at the "information" that can be generated and processed in a quantum mechanical system. Instead of having a typical bit that equals 0 or 1, we concern ourselves with quantum bits (qubits) that have two distinct states $|0\rangle$ and $|1\rangle$ that the qubit can collapse into after measurement. However, prior to measurement, the qubit follows quantum mechanical rules, and so can exist in a superposition of $|0\rangle$ and $|1\rangle$. There are also additional resources available to quantum systems that do not have a classical counterpart, such as *quantum entanglement*, which lead to the revolutionary ability of quantum information processing systems to outperform classical systems for specific (important) tasks. Also in similarity to classical information theory, it is important to note that quantum information theory studies the information processing ability of quantum systems in non-ideal situations – e.g., in the presence of noise. Landmark mathematical theorems in the area of quantum information theory underlie why it is expected that we can actually build large-scale quantum computers some day. Prior to these theoretical accomplishments, large-scale quantum computers were thought of as potentially being a "pipe dream" that would never be achievable due to the difficulty in maintaining *quantum coherence* for all of the qubits in a system under the presence of noise.

Quantum information theory can be a challenging subject, and so we will not be able to explore it in depth in this course. We will comment on some of the basic ideas as we look at some specific examples of revolutionary quantum information processing technologies, such as quantum communication systems, quantum computers, and quantum sensors. For those interested in more details, the standard introductory text (at a graduate level) for quantum information theory is [32].

11.2.1 Quantum Decoherence and Noise

To begin discussing quantum information, it is helpful to return to considering the meaning of various elements of the density matrix and the idea of quantum coherence. Recall that if we had a two-level system in the pure state

$$|\psi\rangle = c_0|0\rangle + c_1|1\rangle, \tag{11.31}$$

then the density matrix represented in the $|0\rangle$ and $|1\rangle$ basis would be

$$\rho = \begin{bmatrix} |c_0|^2 & c_0 c_1^* \\ c_0^* c_1 & |c_1|^2 \end{bmatrix}. \tag{11.32}$$

We see that the off-diagonal elements contain information about the relative phase between the $|0\rangle$ and $|1\rangle$ components of the pure state. When this relative phase is "perfectly" specified, we can think of the two states as being *coherent.* As they evolve in time, we can see coherent interference effects between the states that represent different physical ideas depending on the underlying physical system. For instance, it could be describing the oscillation of the wavefunction of an electron trapped in a potential well, or could be producing the wave interference fringes in a double-slit diffraction experiment. When we lose this coherence (referred to as *decoherence*), we stop being able to see these interference effects that are characteristic of many quantum phenomena.

This idea of decoherence can be very closely linked to the idea of a mixed state. As a simple example, if we assume that we randomly prepare identical two-level systems with a 50-50 probability of starting in either the $|0\rangle$ or $|1\rangle$ state, we can find that our density matrix will become

$$\rho = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{11.33}$$

We see that the off-diagonal elements have vanished because there is no coherence in the mixed state, we are simply in the $|0\rangle$ or $|1\rangle$ state, there is no longer any *superposition*. However, this is not the only way that we can arrive at this density matrix for a mixed state. Let's now assume that we randomly prepare our two-level system in a superposition state so that our c_0 and c_1 coefficients in (11.31) have the same magnitude but random phases. We then always have $|c_0|^2 = |c_1|^2 = 1/2$, but that the off-diagonal terms will look like

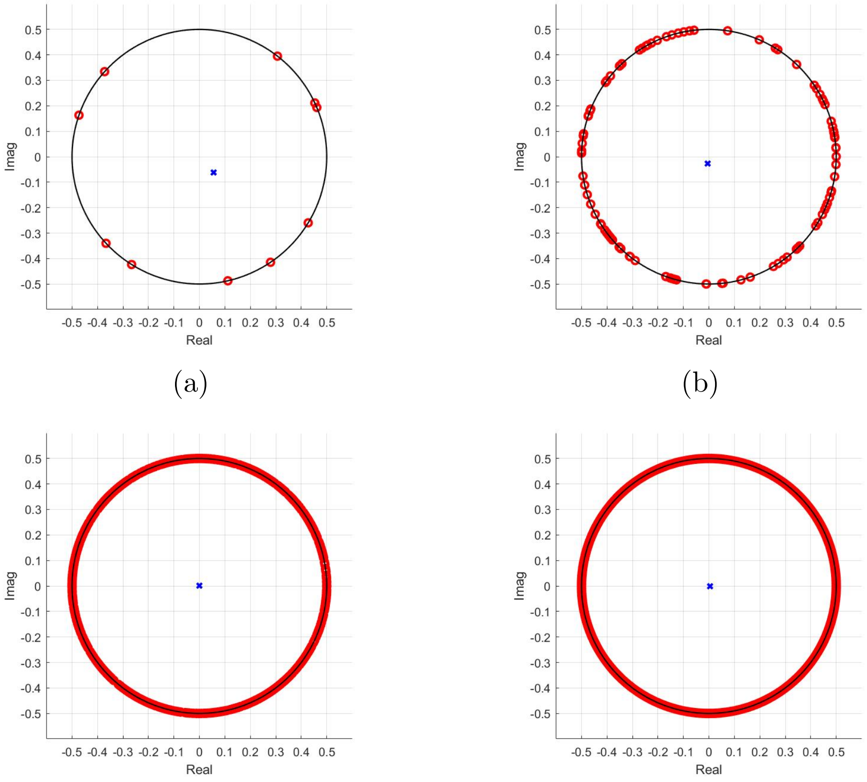
$$c_0 c_1^* = \frac{1}{2} e^{i\delta}, \tag{11.34}$$

where δ is the (random) relative phase between c_0 and c_1. If we now take the mean of all the random realizations of (11.34) for a large number of δ's, we find that this value very rapidly goes to 0. This is illustrated visually in Fig. 11.1. The resulting density matrix for this mixed state is then also given by (11.33).

This example illustrates an important point about the density matrix. We can have completely different "scenarios", but still end up with an identical density matrix because from the perspective of the statistical predictions of quantum theory our different "scenarios" are actually physically indistinguishable. Hence, we see that preparing our qubit in a random equal-weight linear superposition state is statistically the same as our earlier example of randomly placing the qubit in the $|0\rangle$ or $|1\rangle$ states with equal probability. We already commented on this situation as having no "coherence", and so we see this classification also applies to our random equal-weight linear superposition system as well.

Let's now extend some of these notions to a very practical situation of considering an *open quantum system*. The basic idea of an open quantum system is that we have a particular subsystem that we are focused on analyzing, e.g., a qubit. However, in a real physical system our qubit is *never* perfectly isolated from other physical systems. In general, these other systems our qubit is coupled to are "vast" in the sense that they likely need an infinite-dimensional Hilbert space to describe their state. As a result, attempting to fully describe the system we care about (i.e., the qubit) and the *environment* it is coupled to becomes something of an impossible task in most situations. The theory of open quantum systems is focused on determining how the system we care about (i.e., the qubit) evolves in time given that it is coupled to this vast environment, without attempting to describe in detail the exact state of the environment. This naturally leads us to the idea of characterizing the effect of the environment as some form of *noise* that our subsystem of interest is subjected to.

For instance, we can have our two-level system subjected to noise due to thermal fluctuations in an adjacent system. In the case of our quantized circuits, this could come from the finite resistance of circuit components, where the loss in the resistor is then converted to heat. However, we can also have noise in the form of zero-point fluctuations of nearby



(a) (b)

(c) (d)

Figure 11.1: Evaluating the off-diagonal element of the density matrix for (a) 10, (b) 100, (c) 1,000, and (d) 10,000 random equal-weight linear superposition states. The red circles are the realizations of $c_0 c_1^*$ and the blue "x" is the mean value that would be the actual matrix entry ρ_{01}.

systems, e.g., an LC oscillator that our qubit is coupled to. There can also be noise in the applied control signals, such as classical voltages, that we use to attempt to control the qubit. Hence, to consider quantum information processing in a practical scenario we must invariably consider the presence of noise.

However, not all noise is created equal. In general, we need to understand how the noise couples to our two-level system to properly characterize its effect. To see this more explicitly, let's recall our discussion about a semiclassical treatment of a classical voltage source driving a two-level system. To begin, we considered a bias voltage of

$$V_b(t) = \begin{cases} V_x(t)\cos(\omega_d t) + V_y(t)\sin(\omega_d t), & 0 < t < t_g \\ 0, & \text{otherwise}, \end{cases} \tag{11.35}$$

and saw that upon substituting this into our Hamiltonian and making the rotating-wave approximation we could write our resulting system Hamiltonian as

$$\hat{H}_{\text{TLS,RWA}} = -\frac{1}{2}\hbar\delta\hat{\sigma}_z + \frac{1}{2}\hbar g V_x(t)\hat{\sigma}_x + \frac{1}{2}\hbar g V_y(t)\hat{\sigma}_y. \tag{11.36}$$

We then discussed the idea of a driven Rabi oscillation, which involved us designing $V_x(t)$ or $V_y(t)$ to rotate the state vector of our two-level system around the x- or y-axis of the Bloch sphere, respectively. If we then were to have noise in this applied voltage, we would see that this would drive rotations around the respective axes of the Bloch sphere, leading to the exact location of the state vector on the Bloch sphere becoming less "certain". We can express this uncertainty by considering our two-level system to now be in a mixed state, with the probabilities determining the mixed state linked to the statistics we know about the noise coupling to the two-level system. More generally, we see that any noise source that couples to the $\hat{\sigma}_x$ and $\hat{\sigma}_y$ operators of the two-level system will lead to this same kind of effect.

However, we also had another notion of a spontaneous transition occurring in the state of our two-level system. For instance, when we considered the fully-quantized treatment of a two-level system coupled to an LC oscillator, we saw that vacuum fluctuations in the LC oscillator coupled to the $\hat{\sigma}_x$ and $\hat{\sigma}_y$ operators lead to a spontaneous emission rate that characterized how likely it was for the two-level system to drop from its excited to ground state. This kind of effect is sometimes referred to as a *longitudinal relaxation* process because it modifies the qubit state along the axis of the qubit basis states in the Bloch sphere representation. It is typical to lump all longitudinal relaxation effects together into an overall *longitudinal relaxation rate* given by $\Gamma_1 \equiv 1/T_1$. The T_1 time notation has its origins in the study of nuclear magnetic resonance (the effect used in magnetic resonance imaging), and is a common terminology that is now applied to many quantum systems. This relaxation time is not a perfect characterization of the different noise processes that impact a two-level system, but it is used as a common figure of merit because it does an adequate job in many situations. In practice, we operate our two-level systems so that their equilibrium state would be $|0\rangle$ (i.e., if we let the system sit for a very long time we would expect to find the system in its equilibrium state) and that there is a negligible chance of "spontaneous absorption" raising the two-level system to its $|1\rangle$ state. The T_1 relaxation time then describes the exponential decay process as a function of time in terms of our density matrix for initial pure state $|\psi\rangle = c_0|0\rangle + c_1|1\rangle$ as

$$\rho \text{ ``}=\text{''} \begin{bmatrix} 1 + (|c_0|^2 - 1)e^{-\Gamma_1 t} & c_0 c_1^* \\ c_0^* c_1 & |c_1|^2 e^{-\Gamma_1 t} \end{bmatrix}. \tag{11.37}$$

Note that the " = " is used here because this expression does not yet fully show how T_1 relaxation effects impact all elements of the density matrix. We turn to discussing this broader picture now.

Although the T_1 time does encapsulate many important effects, like spontaneous emission, it does not completely describe all possible types of noise. For instance, when we discussed the transmon qubit we mentioned that we could make the operating frequency of the qubit tunable using a SQUID instead of a single Josephson junction. If we have magnetic flux noise intersecting the SQUID loop, then we will find that the operating frequency of the qubit becomes "noisy". Looking back at our Hamiltonian in (11.36), we see that this kind of noise will couple to our system in terms of the $\hat{\sigma}_z$ operator. As with the other Pauli matrices, this operator drives rotations around its axis, in this case the z-axis. This kind of noise impacts the coherence of a linear superposition state. For instance, one can consider the case of a qubit in an equal-weight superposition and recognize that the resulting state vectors lie on the equator of the Bloch sphere. Noise applied to the $\hat{\sigma}_z$ operator will then cause the relative phase in the superposition to become noisy. We have already seen from our earlier discussions that this will impact the off-diagonal terms of the density matrix.

More generally, we sometimes refer to this kind of effect as a *transverse relaxation* process because it modifies the qubit state in the plane transverse to the qubit axis on the Bloch sphere. Broadly speaking, any "phase-breaking" process that upsets a superposition state will be incorporated into the *transverse relaxation rate* Γ_2. Since longitudinal relaxation does constitute a "phase-breaking" process, we find that the Γ_1 will actually also be included in our Γ_2 relaxation rate. More specifically, we have that

$$\Gamma_2 \equiv \frac{1}{T_2} = \frac{\Gamma_1}{2} + \Gamma_\varphi, \tag{11.38}$$

where T_2 is the transverse relaxation time and Γ_φ is referred to as the *pure dephasing rate*. This pure dephasing rate specifically characterizes the impact of the noise coupled to the $\hat{\sigma}_z$ operator. As alluded to earlier, the transverse relaxation impacts the off-diagonal terms of the density matrix. The typical format of this then causes (11.37) to become

$$\rho = \begin{bmatrix} 1 + (|c_0|^2 - 1)e^{-\Gamma_1 t} & c_0 c_1^* e^{-\Gamma_2 t} \\ c_0^* c_1 e^{-\Gamma_2 t} & |c_1|^2 e^{-\Gamma_1 t} \end{bmatrix}. \tag{11.39}$$

This model for qubit decoherence is sometimes referred to as the *Bloch-Redfield model*. It should be noted that this model is an idealization that assumes the noise impacting the system is rather "simple" (e.g., like white noise), and though real systems can often interact with noise that deviates from this simple model it still serves as a useful starting point in many cases for understanding the coherence of a quantum system. Rigorously considering open quantum systems with more realistic noise models is a complicated endeavor, and still represents an active area of research.

These two relaxation processes help us start to understand why processing quantum information can be so difficult. Our systems will always be subjected to noise, which work together to destroy the coherence of our linear superpositions of qubit states. This leads to the quantum information stored in the qubit state to decay. A significant component of designing quantum information processing systems today involves attempting to increase the T_1 and T_2 times so that we can accomplish more processing tasks within the finite time that we can expect a linear superposition to maintain coherence. Once the coherence has vanished, the system has lost all of its meaningful quantum information, and any further operations we perform on the quantum state will not be performing the intended effect on the superposition state that we used to have.

11.2.2 Quantum Entanglement

Quantum entanglement is a rather mysterious, but fundamental property, that is central to the performance advantage possible with many quantum information processing technologies. We will first introduce it through a rather mathematical definition, and will then see how it "pops up" time and time again as we discuss the basics of various revolutionary quantum technologies.

Entanglement is a property that can exist between two or more "quantum systems". For now, let's make our conversation explicit by considering as our "quantum systems" two distinct qubits that can interact with one another. We have already discussed how to describe such a composite quantum system earlier in the course when we considered the fully-quantized treatment of two LC oscillators capacitively coupled to one another or a single LC oscillator capacitively coupled to an artificial atom (e.g., a transmon or charge qubit). In particular, we found that the total Hilbert space that the state of the composite system should be described in was given by the tensor product of the Hilbert spaces of the individual systems. For two qubits, we could then write a generic state as

$$|\psi\rangle = c_{00}|0\rangle_1|0\rangle_2 + c_{01}|0\rangle_1|1\rangle_2 + c_{10}|1\rangle_1|0\rangle_2 + c_{11}|1\rangle_1|1\rangle_2, \tag{11.40}$$

where $|m\rangle_n$ is the mth state of the nth qubit. Another more common notation would be to consolidate the different kets as

$$|\psi\rangle = c_{00}|00\rangle + c_{01}|01\rangle + c_{10}|10\rangle + c_{11}|11\rangle. \tag{11.41}$$

We see that our two 2-dimensional Hilbert spaces have combined into a new Hilbert space with a dimension of 4. This trend continues, so that if we were to have a system of N qubits coupled to each other the dimension of the resulting Hilbert space would be 2^N, which is a number that "explodes" very quickly.

One important property of the generic state given in (11.41) is that it is *factorizable*. In particular, we can think of it as a tensor product between the two linear superposition states $|\psi_1\rangle = a_0^{(1)}|0\rangle_1 + a_1^{(1)}|1\rangle_1$ and $|\psi_2\rangle = a_0^{(2)}|0\rangle_2 + a_1^{(2)}|1\rangle_2$. Explicitly, we have

$$\begin{aligned} |\psi_1\rangle \otimes |\psi_2\rangle &= \Big(a_0^{(1)}|0\rangle_1 + a_1^{(1)}|1\rangle_1\Big) \otimes \Big(a_0^{(2)}|0\rangle_2 + a_1^{(2)}|1\rangle_2\Big) \\ &= a_0^{(1)}a_0^{(2)}|00\rangle + a_0^{(1)}a_1^{(2)}|01\rangle + a_1^{(1)}a_0^{(2)}|10\rangle + a_1^{(1)}a_1^{(2)}|11\rangle. \end{aligned} \tag{11.42}$$

We can then identify by matching to (11.41) that $c_{00} = a_0^{(1)}a_0^{(2)}$, $c_{01} = a_0^{(1)}a_1^{(2)}$, and so on. In this case, we see that the states of the two subsystems are independent of one another. If we make a measurement on one of our two subsystems, nothing happens to the other subsystem.

However, this is not the only kind of state that is allowable. For example, we could have the state

$$|\Phi^{\pm}\rangle = \frac{1}{\sqrt{2}}\big(|00\rangle \pm |11\rangle\big), \tag{11.43}$$

which cannot be factorized into a product of the states of the individual subsystems like in the first line of (11.42). Such a state that cannot be factorized in this way is called *entangled*. In this case, we are no longer able to think of the states of the two subsystems individually anymore. In particular, if we were to make a measurement on the first qubit, we would find that this would not only collapse the wavefunction for qubit 1, but *also* for qubit 2. For instance, if we measure qubit 1 and find that it is in the state $|0\rangle_1$, then we immediately know for the state given in (11.43) that qubit 2 has also collapsed into the state $|0\rangle_2$. This effect can happen even if the two qubits are spatially separated, and apparently occurs instantaneously. This is a very peculiar effect, that has bemused physicists for decades due to certain paradoxes that seem to be possible with this "spooky action at a distance" (to quote Einstein's views on the subject). As bizarre as these entangled states seem, they have many important consequences in developing technology and are a central resource in quantum information theory. We will discuss some of these ideas in more depth in the final sections of this course.

11.3 Bell's Inequalities and Interpretation of Quantum Mechanics

Previously, we briefly introduced the idea of quantum entanglement by looking at the example two-qubit states

$$|\Phi^{\pm}\rangle = \frac{1}{\sqrt{2}}\big(|00\rangle \pm |11\rangle\big). \tag{11.44}$$

Another set of two-qubit states with similar properties are

$$|\Psi^{\pm}\rangle = \frac{1}{\sqrt{2}}\big(|01\rangle \pm |10\rangle\big). \tag{11.45}$$

Collectively, these are referred to as *Bell states*, and they form a basis for all possible composite two-qubit states. We saw that there would be certain situations for these kind of states that if we apply the "rules" of quantum mechanics (in particular, quantum mechanical measurement) we have been using up to this point we could have some very peculiar results. The main reason for this was that these entangled states are arranged in such a manner that it becomes impossible for us to properly think about the state of an individual qubit irrespective of the other qubit. Mathematically, this was expressed by the fact that the states given in (11.44) and (11.45) cannot be factored into a tensor product of the two single qubit states.

The thought experiment that we did to consider one of the very peculiar properties of these entangled states was to imagine that our two qubits were spatially separated when we perform a measurement on one of the qubits. From our rules of quantum mechanics, this measurement causes the wavefunction of the system to collapse according to the measured result. Due to the inseparability of these different states, we see that the collapse due to the measurement on one qubit actually influences the other qubit as well, forcing it to collapse into the appropriate value according to the measurement on the first qubit. This effect apparently happens instantaneously, and to be frank about it, just seems completely *wrong* from our classical intuition about physical systems. This uncomfortable physical process was referred to as "spooky action at a distance" by Einstein, and was considered to be utterly ridiculous to him and his collaborators. He and his co-authors (Podolsky and Rosen) presented this paradoxical situation (referred to as the *EPR paradox* after the initials of the authors) and introduced some of their ideas on how to overcome this paradox.

Before discussing more about this, it is important to note that the Bell states given in (11.44) and (11.45) are not simply some imagined "crazy" quantum state that does not occur in practice. Rather, they can be generated quite readily due to various physical processes. We will not go into detail about these physical processes, but they often occur naturally in systems where a single "particle" decays into two others where the decay process must occur in a manner that conserves momentum. For example, a single high frequency photon propagating in a nonlinear crystal can cause a mixing process to occur that leads to the generation of two lower frequency photons (this is referred to as *spontaneous parametric down-conversion*). To conserve momentum, the two generated photons must propagate in opposite directions and have a fixed relationship between their polarization. The resulting polarization states can act as the states of a qubit, and because the two photons propagate in opposite directions we see that we can very quickly have two qubits spatially separated from one another quite quickly (the speed of light is 3×10^8 m/s, after all).

So, given that the Bell states do actually occur due to a number of possible physical processes, it was the notion of Einstein, Podolsky, and Rosen that quantum mechanics must be an incomplete theory in some manner. They proposed one possible attempt to overcome this kind of issue in 1935. At that time, quantum mechanics had been found to work very successfully, so they did not attempt to say that quantum mechanics was "wrong". Rather, they felt that the reason for the statistical nature had to come from the fact that we were having an incomplete description of the physical system in some manner. Their view was that there were some kind of *hidden variables* that we are unaware of that caused the two photons that were generated to have a specific set of complementary polarizations at the moment of creation. Within this physical picture, the measurement outcomes were decided at the time of creation, and so there was no wavefunction collapse or "spooky action at a distance" to contend with.

Work proceeded on developing quantum mechanics and possible *hidden variable theories* for some time, but the story changed rather dramatically in 1964 when John Bell determined a simple but clever *experimental* test to see whether quantum mechanics (as we have been learning it) or some hidden variable theory was correct. In particular, he developed a test that was applicable to *any* ***local*** hidden variable theory. The term *local* here is sometimes restated as the "principle of locality", which in vague terms means that a physical system can only be influenced by its "immediate surroundings". More specifically, these ideas of "local" and "influence" are usually framed in the context of the theory of relativity, which has a central tenet that no physical influence/information can propagate at a speed greater than that of light. What John Bell proved was that quantum mechanics was *incompatible* with any local hidden variable theory, and thus made it an experimental (rather than philosophical) question as to which viewpoint was correct.

The proof of this centers around the outcomes of measurements on Bell states, where it was shown that any local hidden variable theory would be bound by the same inequality, regardless of the possible hidden variables involved. Traditionally, any inequality of this form is referred to as a *Bell inequality.* The surprising result was that for this set of measurements the expected outcome based on a quantum mechanical calculation would violate the Bell inequality. As a result, it came down to engineering a physical system well enough to be able to perform the requisite experiments. This was completed (fairly) convincingly in 1982 by Alain Aspect, with the results coming out in overwhelming favor of quantum mechanics. These *Bell tests* have been completed by many other scientists with many different physical systems over the years with more and more refined setups to close various "loopholes" that may exist. In all cases, the results have agreed with quantum mechanics and the *Copenhagen interpretation* that we have learned in this course. As a result, we can conclusively say that nature is apparently *nonlocal*, either in the form of an instantaneous collapse of a wavefunction or due to some kind of nonlocal hidden variable theory.

There are various interpretations of quantum mechanics that attempt to provide a consistent way to explain these very peculiar effects. We will not discuss those here, but the interested reader can find more information in [8, 9] and the references therein. For now, we will conclude our discussion by stating that the loss of locality to quantum mechanics is a central component of what makes many quantum information processing technologies able to "work". If everything was determined by some local hidden variables upon creation, we would not be able to process the various superpositions of entangled states in a quantum mechanical manner that is essential for generating the *non-classical* results of many revolutionary quantum technologies. So although it is perhaps disconcerting to think about some of the odder possible implications of the nonlocality of quantum physics, there are certainly extremely exciting technological innovations to focus on instead to distract us from any philosophical questions!

11.4 Introduction to Quantum Communications

Quantum communication systems can broadly be thought of as a communication system that communicates quantum information, rather than just classical information. It turns out that there are a number of possible applications for such systems. One example is *superdense coding*, which is a quantum communication protocol that is able to communicate some number of classical bits of information using a smaller number of qubits. Another interesting use of quantum communications is for networking together quantum computers to lead to an overall larger quantum computing architecture (much like how we network together many classical computers to form supercomputers). However, possibly the most famous use of quantum communication systems is for quantum cryptography – i.e., a secure form of encryption for a communication that is impossible to "crack".

11.4.1 Quantum Key Distribution

To understand a basic form of quantum cryptography, known as *quantum key distribution*, it is necessary to first establish the *no-cloning theorem.* As the name suggests, the no-cloning theorem proves that it is impossible to perfectly copy one quantum state onto another. As a basic sketch of this proof, consider that we set our first system (the one we wish to clone) into an arbitrary state $|\psi_a\rangle_1$ and our second system (the one we wish to copy the first system's state onto) into a standard starting state $|\psi_s\rangle_2$. What we will find is that it is not possible to arrive at the composite state $|\psi_a\rangle_1|\psi_a\rangle_2$ for an arbitrary state $|\psi_a\rangle$.

To see why, we note that because we are dealing with quantum mechanical states we can imagine that our "copying" functionality must be accomplished by some linear time-evolution operator acting on our composite system. For instance, we have previously discussed that the time-evolution operator can be generically written as

$$\hat{T} = e^{-i(\hat{H}/\hbar)t}, \tag{11.46}$$

where our task then would be to design a Hamiltonian to perform the copying functionality. Assuming we have done this, we can consider how $\hat{T}$ will affect our composite system for two basis states we choose for qubit 1. To denote these generically, we will call the basis states $|\psi_1\rangle_1$ and $|\psi_2\rangle_1$. For the copy function to work correctly for these basis states we must then have,

$$\hat{T}|\psi_1\rangle_1|\psi_s\rangle_2 = |\psi_1\rangle_1|\psi_1\rangle_2 \tag{11.47}$$

and

$$\hat{T}|\psi_2\rangle_1|\psi_s\rangle_2 = |\psi_2\rangle_1|\psi_2\rangle_2. \tag{11.48}$$

If we now consider a superposition state $|\psi\rangle_1 = \alpha|\psi_1\rangle_1 + \beta|\psi_2\rangle_1$, we see that

$$\begin{aligned}\hat{T}\big(\alpha|\psi_1\rangle_1 + \beta|\psi_2\rangle_1\big)|\psi_s\rangle_2 &= \alpha\hat{T}|\psi_1\rangle_1|\psi_s\rangle_2 + \beta\hat{T}|\psi_2\rangle_1|\psi_s\rangle_2 \\ &= \alpha|\psi_1\rangle_1|\psi_1\rangle_2 + \beta|\psi_2\rangle_1|\psi_2\rangle_2,\end{aligned} \tag{11.49}$$

due to the linearity of quantum operators. We note that our result in (11.49) *does not* equal the correctly copied state, which should have been $|\psi\rangle = \big(\alpha|\psi_1\rangle_1 + \beta|\psi_2\rangle_1\big)\big(\alpha|\psi_1\rangle_2 + \beta|\psi_2\rangle_2\big)$. Hence, due to the linearity of quantum mechanics, it is impossible to perfectly clone an arbitrary superposition state.

To see how to use this for a simple quantum encryption scheme, we imagine that we have two people attempting to communicate securely with one another. These individuals are typically referred to as “Alice” and “Bob” by convention. In addition to Alice and Bob, there is a third person named “Eve” who is attempting to eavesdrop (this pun brought to you courtesy of quantum physicists...) on the communication. Our goal is to devise a scheme that allows Alice and Bob to determine whether Eve is present and intercepting the message or not.

To begin, let’s imagine that Alice sends a qubit to Bob in the $|0\rangle$ or $|1\rangle$ state, and that Bob is able to orient his measuring apparatus to measure these states perfectly. This approach is not secure, because if Eve is present she can intercept the message, take note of it, and then prepare a qubit of her own in the same state she just received and then transmit this to Bob. This is effectively how a classical communication system works, there is not in principle a sufficiently strong way to detect whether an eavesdropper is present or not.

If Alice changes her basis from the $|0\rangle$ and $|1\rangle$ basis to the $|+\rangle = 1/\sqrt{2}(|0\rangle + |1\rangle)$ and $|-\rangle = 1/\sqrt{2}(|0\rangle - |1\rangle)$ basis and Bob leaves his measurements in the $|0\rangle$ and $|1\rangle$ basis, we end up with a very different result. In this case, Bob always has a 50-50 chance of measuring a $|0\rangle$ or a $|1\rangle$, and so he does not receive any information from Alice.

Let’s now assume that Alice and Bob are both measuring in the $|\pm\rangle$ basis, but that Eve is present and is measuring in the $|0\rangle$ and $|1\rangle$ basis. Alice and Bob can now send a number of test qubit transmissions and then communicate with one another classically on their outcomes. They will quickly be able to deduce that something is wrong because their list of transmissions and measurement results will not be correctly correlated. In particular, each time Eve receives a qubit she will have a 50-50 chance of measuring a $|0\rangle$ or a $|1\rangle$, which she will then transmit to Bob. Each time Bob receives a $|0\rangle$ or a $|1\rangle$ and measures in his $|\pm\rangle$ basis, he will have a 50-50 probability of getting a $|+\rangle$ or a $|-\rangle$. By sending enough test qubits, the statistics of the situation can quickly make it abundantly clear whether something has gone wrong in their communication channel, possibly due to an eavesdropper being present.

This approach is again not actually secure, because it could happen that Eve has happened to measure in the correct basis. The key to overcoming this is for Alice and Bob to both randomly choose which basis they will transmit and receive in. They can then classically communicate with one another after performing a sufficient set of transmit and receive cycles to determine for which set of transmitted qubits they were aligned with one another. They can then compare records of transmitted and received qubits for some set of these to determine whether they are correctly correlated or if there is an eavesdropper present. If they are properly correlated, they can then be comfortable knowing that they have some set of bits that they have shared with one another that no one else has access to. For an idea of how many qubits need to be compared, the probability of finding a disagreement between Alice and Bob in the presence of Eve is

$$P = 1 - \left(\frac{3}{4}\right)^n, \tag{11.50}$$

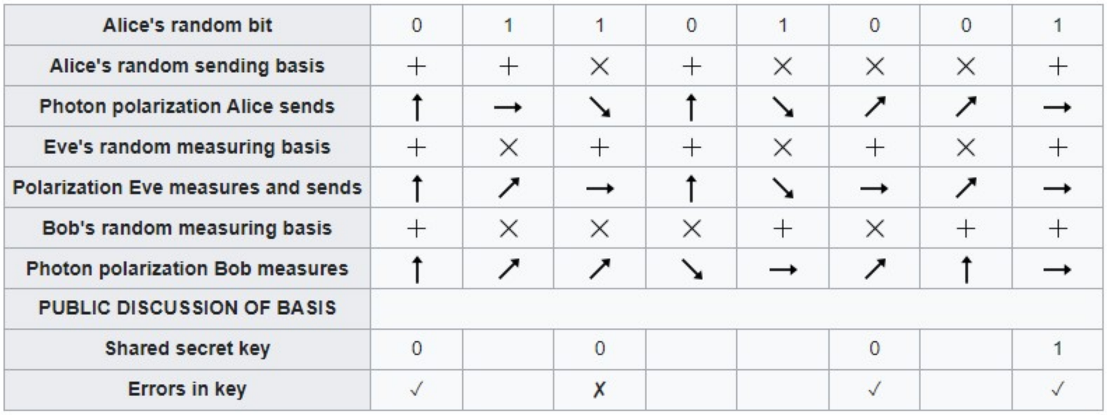
where n is the number of comparisons made. For a value of $n = 72$ comparisons, the probability of detecting an eavesdropper becomes $P = 0.999999999$. Overall, this communication strategy works because of the no-cloning theorem. This prevents Eve from being able to clone the qubit transmitted by Alice, and then send the clone on to Bob while keeping the original qubit for herself to try and gain some partial information about what is being communicated.

An example set of scenarios for this protocol with and without an eavesdropper present are tabulated in Figs. 11.2 and 11.3. In these examples, a basis composed of photon polarization states are used instead of our $\{|0\rangle, |1\rangle\}$ and $\{|+\rangle, |-\rangle\}$ sets. One does not need a detailed understanding of electromagnetic field polarization to interpret the examples in Figs. 11.2 and 11.3, simply know that sending within the + basis (vertical and horizontal polarizations) or the × basis (rotated basis by 45°) is equivalent to our scenario in terms of $\{|0\rangle, |1\rangle\}$ and $\{|+\rangle, |-\rangle\}$ sets.

In reality, this communication process is not particularly efficient for attempting to send an entire message. Instead, what is done is this process is utilized to distribute what is known in cryptography as a *key*. This key can be much smaller, and is a shared resource that is used to encrypt and then decrypt a classically communicated message. By securing the transmission of the key due to the rules of quantum mechanics, the resulting classical transmission is much more secure than existing cryptographic communication protocols. This is referred to as *quantum key distribution*, and is a quantum technology that has been successfully implemented in a growing number of more and more practical scenarios. However, this approach is hardly perfect, since if there is an eavesdropper present the outcome is that Alice and Bob must abort their attempt to communicate securely until Eve leaves (or the communication channel is repaired in case there was some other physical error source

Alice's random bit	0	1	1	0	1	0	0	1
Alice's random sending basis	+	+	×	+	×	×	×	+
Photon polarization Alice sends	↑	→	↘	↑	↘	↗	↗	→
Bob's random measuring basis	+	×	×	×	+	×	+	+
Photon polarization Bob measures	↑	↗	↘	↗	→	↗	→	→
PUBLIC DISCUSSION OF BASIS								
Shared secret key	0		1			0		1

Figure 11.2: Tabulation of a set of send and measurement scenarios without an eavesdropper present (image from Wikipedia [33]).



Alice's random bit	0	1	1	0	1	0	0	1
Alice's random sending basis	+	+	×	+	×	×	×	+
Photon polarization Alice sends	↑	→	↘	↑	↘	↗	↗	→
Eve's random measuring basis	+	×	+	+	×	+	×	+
Polarization Eve measures and sends	↑	↗	→	↑	↘	→	↗	→
Bob's random measuring basis	+	×	×	×	+	×	+	+
Photon polarization Bob measures	↑	↗	↗	↘	→	↗	↑	→
PUBLIC DISCUSSION OF BASIS								
Shared secret key	0		0			0		1
Errors in key	✓		✗			✓		✓

Figure 11.3: Tabulation of a set of send and measurement scenarios with an eavesdropper present (image from Wikipedia [33]).

present that was causing the issues). As a result, there is still significant work in the area of quantum communications and quantum cryptography to develop better protocols and analyze their level of security (and practicality).

11.4.2 Quantum Teleportation

Another important type of quantum communication comes in the form of *quantum teleportation*. Unfortunately, despite the exciting name, quantum teleportation is a far cry away from being able to teleport matter (sorry, Trekkies...). Rather, quantum teleportation is focused on transferring a *quantum state* between two places, not the underlying physical carrier of that state. Accomplishing this goal seems perhaps rather challenging, since we already know that we cannot clone a quantum state and then send that cloned state somewhere. The way that we end up accomplishing quantum teleportation is through sharing entanglement

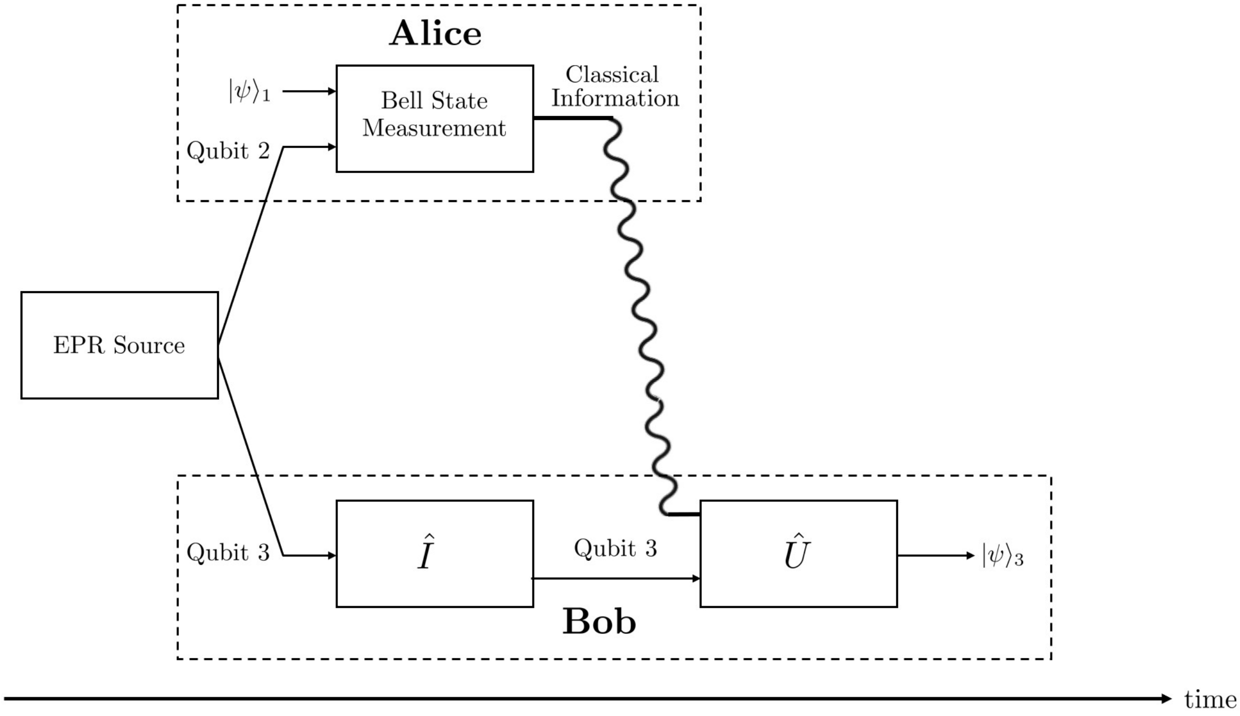


Figure 11.4: Basic schematic of how a quantum state can be teleported between two locations.

between two locations. A basic schematic of the quantum teleportation protocol is shown in Fig. 11.4.

The basic idea begins with an *EPR source*, which is named after the EPR paradox because this source produces a pair of qubits in a particular Bell state. The EPR source produces the Bell state

$$|\Psi^-\rangle_{23} = \frac{1}{\sqrt{2}}\big(|0\rangle_2|1\rangle_3 - |1\rangle_2|0\rangle_3\big) \tag{11.51}$$

in such a way that the qubit 2 travels to Alice and qubit 3 travels to Bob. Our goal is to now take an input qubit at Alice's location and "teleport" its state to Bob's location.

Let's now assume that the input qubit at Alice's location has the state

$$|\psi\rangle_1 = \alpha|0\rangle_1 + \beta|1\rangle_1. \tag{11.52}$$

The state of our entire composite system of the 3 qubits is then

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\big(\alpha|0\rangle_1 + \beta|1\rangle_1\big)\big(|0\rangle_2|1\rangle_3 - |1\rangle_2|0\rangle_3\big). \tag{11.53}$$

Recalling that the Bell states form a complete basis for two-qubit states, we can show after some algebra that it is possible to rewrite (11.53) in terms of Bell states for qubit's 1 and 2.

The result is

$$|\Psi\rangle = \frac{1}{2}\bigg(|\Phi^+\rangle_{12}\big(\alpha|1\rangle_3 - \beta|0\rangle_3\big) + |\Phi^-\rangle_{12}\big(\alpha|1\rangle_3 + \beta|0\rangle_3\big) \\ + |\Psi^+\rangle_{12}\big(-\alpha|0\rangle_3 + \beta|1\rangle_3\big) - |\Psi^-\rangle_{12}\big(\alpha|0\rangle_3 + \beta|1\rangle_3\big)\bigg). \quad (11.54)$$

Alice can now perform a Bell state measurement to determine which Bell state qubits 1 and 2 are in. If we assume she performs this measurement and the result corresponds to the $|\Phi^-\rangle_{12}$ state, then we know that the qubit at Bob's location has collapsed into the state

$$|\phi\rangle_3 = \alpha|1\rangle_3 + \beta|0\rangle_3. \quad (11.55)$$

This is *not* exactly the input state that Alice had for $|\psi\rangle_1$ given in (11.52). However, we can apply some known set of unitary operations to $|\phi\rangle_3$ to make the result eventually equal $|\psi\rangle_3$. The catch is that the particular unitary operations that Bob needs to perform depends on the Bell state measurement that Alice completes, as is evident from the different terms in (11.54). To fix this issue, Alice uses a classical communications channel to tell Bob what the result of her Bell state measurement was. Based on this, he can then perform whichever unitary operations are needed to recover $|\psi\rangle_3$, thus "teleporting" the state of the input qubit at Alice's location to Bob's location.

This teleportation is accomplished without Alice or Bob ever measuring the state to be teleported directly. By avoiding needing to determine the coefficients α and β, Alice and Bob are able to avoid the statistical uncertainties in quantum mechanics, as well as the issues surrounding cloning a quantum state. However, although we have instantaneously teleported the quantum state from Alice's location to Bob's location, it is important to note that no actual *information* has been transferred until Alice classically communicates with Bob as to what unitary operations he needs to perform to recover $|\psi\rangle_3$ at his location. This "loophole" (as one may wish to view it) allows for us to consider that quantum teleportation does not violate the principle of special relativity that no information can be transmitted at speeds faster than the speed of light.

11.5 Introduction to Quantum Sensors

Devices that use quantum mechanical principles to measure or "sense" a physical quantity have become extremely widespread technologies. Although these are very mature technologies with many products commercially available, most quantum sensors to date fall in the category of "first-generation" quantum technologies. That is, they require understanding quantum mechanical principles to have been designed and to be successfully operated, but they do not leverage many effects trademark of "second-generation" quantum technologies (e.g., quantum superposition or entanglement). However, these first-generation type of sensors continue to be developed and have had significant impacts on society in many ways (similar to how transistors have enabled classical information technology). For instance, atomic clocks have revolutionized how we measure and keep track of time, and is an enabling technology for satellite positioning and many other technologies. More broadly, there is an ongoing trend of using first-generation quantum sensors/effects in forming the standard that various SI units are based on due to the fundamental nature of these effects. In other areas, sensitive photodetectors that utilize the photoelectric effect to measure the presence of light enable many optoelectronic technologies that are a key component in the communications infrastructure that underlies the Internet. There are of course, many, many, other examples of first-generation quantum sensors, with more continually being designed as other quantum technologies are matured. We will not focus on these kinds of quantum sensors here.

Instead, we will discuss a few examples of quantum effects that are often of interest in designing second-generation quantum sensors. These effects require control over the quantum state of a system, and can provide better performance than is possible using classical sensing approaches. In particular, we will discuss the concept of a *squeezed state*, as well as how certain kinds of entangled states can be used for sensing purposes.

Although we will not discuss them, another broad class of second-generation quantum sensors involve using an artificial atom as a sensor. Artificial atoms can be very sensitive to different external physical stimuli, which can be leveraged to measure the external physical stimuli by monitoring the behavior of the artificial atom. Since artificial atoms are also typically extremely small in size (often in the nanometer range), these kinds of sensors can also provide extremely fine spatial resolutions compared to larger-scale traditional sensing methods. Some notable examples of these kinds of sensors are nitrogen-vacancy centers [34] and Rydberg atoms [35].

11.5.1 Squeezed States

One important technique that is often used with second-generation quantum sensors is to utilize *squeezed states* when making measurements. At a high level, a squeezed state makes clever use of the uncertainty principle to allow a measurement to be performed that is subject to less noise than would be possible using classical means. This allows for higher-precision measurements to be made, which can be exploited in certain applications.

Prior to discussing the specifics of squeezed states, it will be helpful for us to revisit some properties of coherent states once more. First, recall that coherent states are defined for quantum harmonic oscillators and that we typically characterize a coherent state by a single complex number α, with the property that

$$\hat{a}|\alpha\rangle = \alpha|\alpha\rangle, \tag{11.56}$$

where $\hat{a}$ is the annihilation operator of the quantum harmonic oscillator. We can evaluate the expectation value of the number operator to see that

$$\langle\alpha|\hat{a}^\dagger\hat{a}|\alpha\rangle = |\alpha|^2. \tag{11.57}$$

From this, we conclude that $|\alpha|$ is related to the mean number of quanta in the quantum harmonic oscillator.

Another important property of the coherent state is that it behaves close to classically, in the sense that the parameter α follows the trajectory expected of a classical harmonic oscillator. We found this property by computing the time evolution of $|\alpha\rangle$ by applying the time evolution operator of the quantum harmonic oscillator to it. For the free quantum harmonic oscillator (i.e., there is no interaction with another system), it can be shown that

$$e^{-i(\hat{H}/\hbar)t}|\alpha\rangle = e^{-i\omega_0(\hat{a}^\dagger\hat{a}+1/2)t}|\alpha\rangle = e^{-i\omega_0 t/2}|e^{-i\omega_0 t}\alpha\rangle. \tag{11.58}$$

That is, up to a global phase of $\exp[-i\omega_0 t/2]$, the coherent state simply evolves into a new coherent state oscillating at the frequency of the free quantum harmonic oscillator. It is common to express this trajectory for free evolution in terms of the two *quadratures* that we use to describe the harmonic oscillator. For the case of an LC oscillator, our quadratures were the nodal flux ϕ and charge Q. In terms of these, we can express the time evolution of the expected values of ϕ and Q as

$$\langle e^{-i\omega_0 t}\alpha|\hat{\phi}|e^{-i\omega_0 t}\alpha\rangle = 2\phi_{\text{ZPF}}\,\text{Re}\{\alpha e^{-i\omega_0 t}\},\;\; \langle e^{-i\omega_0 t}\alpha|\hat{Q}|e^{-i\omega_0 t}\alpha\rangle = 2Q_{\text{ZPF}}\,\text{Im}\{\alpha e^{-i\omega_0 t}\}, \tag{11.59}$$

where ϕ_{ZPF} and Q_{ZPF} characterize the magnitude of the zero-point fluctuations in their respective quantities. Due to the form of the expressions in (11.59), it is often convenient to picture the evolution of the coherent state using a phasor diagram. In our case, we would associate the real axis of the diagram with the ϕ quadrature and the imaginary axis with the Q quadrature. In addition to decomposing a complex number into its real and imaginary parts, it can equally well be characterized by its magnitude and phase. Here, we would have that

$$\alpha = |\alpha|e^{i\varphi}. \tag{11.60}$$

Before giving an example of such a phasor diagram, it is important to remember that our state is still quantum mechanical, and as a result still has a statistical uncertainty to it. One very important property of coherent states was that they are an example of a *minimum-uncertainty state* that *saturates* the uncertainty principle (i.e., the product of the two uncertainties equals the minimum value allowable according to quantum mechanics). Further, for a coherent state there is no "preference" between the two quadratures so the uncertainty is distributed evenly between the two. Considering this, we typically draw a coherent state on a phasor diagram as a circle with center defined by $\alpha(t)$ and the diameter of the circle given by the uncertainty. An example of a typical phasor diagram of a coherent state is shown in Fig. 11.5. Alternatively, due to the circular nature of the uncertainty region, we can also think of the uncertainty as being distributed between the number of photons N (which would change the length of the phasor) and the phase φ of the phasor. For some of the most famous applications of squeezed states, it is usually this second perspective that is more useful.

Now, a squeezed state is also an example of a minimum-uncertainty state that saturates the uncertainty principle. The difference with a coherent state is that a squeezed state *reduces* the uncertainty along one direction at the cost of *increasing* the uncertainty in the orthogonal direction. Examples of states *squeezed* in the phase and number directions are shown in Fig. 11.6. The purpose of this squeezing process is to reduce the noise in the measurements made in a particular "favored" quantity. For the examples in Fig. 11.6, squeezing along the N direction allows for more accurate amplitude measurements and squeezing along the

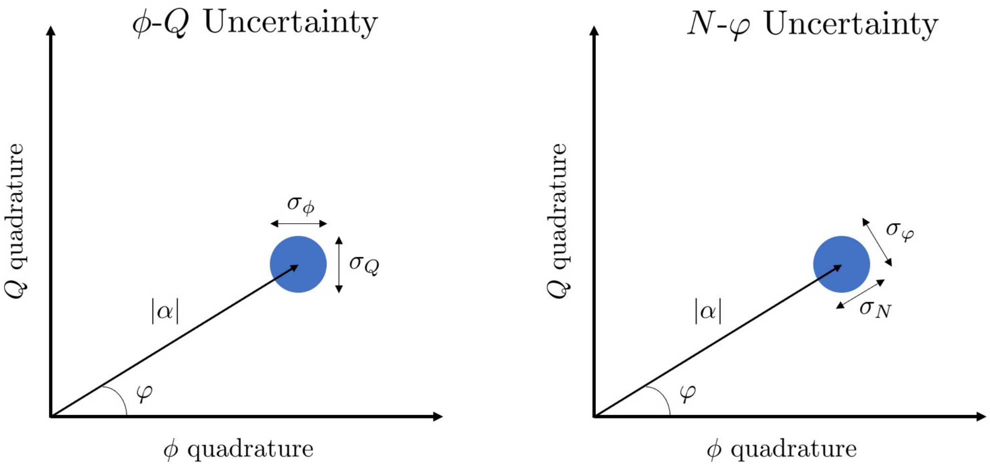


Figure 11.5: Typical phasor diagrams of a coherent state. The quantum uncertainty in the state is represented by the circular area centered at the $|\alpha|e^{i\varphi}$ point on the phasor diagram. (Left) Uncertainty is shown in the ϕ and Q axes, or (right) equivalently in terms of the number of photons N and the phase of the state φ.

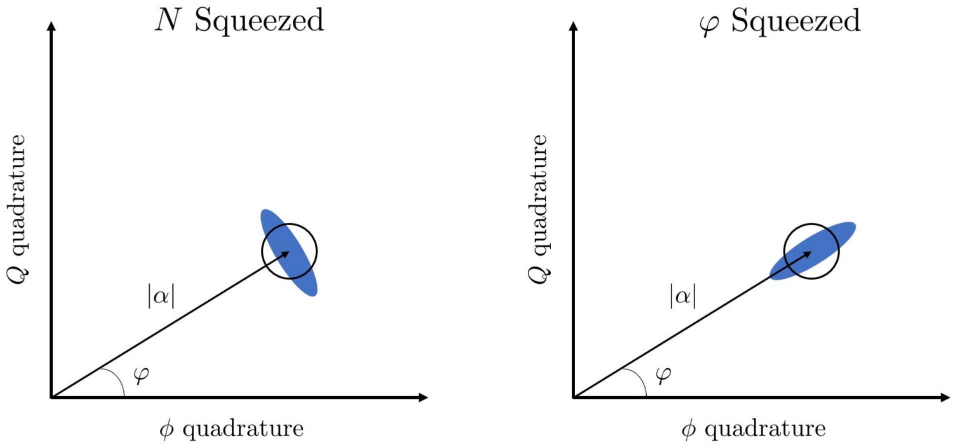


Figure 11.6: Phasor diagrams of squeezed states. The quantum uncertainty in the equivalent coherent state is represented by the circular outline centered at the $|\alpha|e^{i\varphi}$ point on the phasor diagram. The uncertainty region of the squeezed state is shown by the shaded blue region. (Left) State squeezed along the N direction and (right) the φ direction.

φ direction allows for more accurate phase measurements. Classical measurements are best described in terms of coherent states, so we see that using a squeezed state can provide a boost in measurement accuracy past what is possible with any classically-prepared state.

One famous application of this kind of quantum sensing approach is with interferometers. These have long histories in detecting properties of various kinds of waves by interfering two waves that have traveled down different "arms" of the interferometer. Small deviations in the phase accumulated by the two waves as they propagate down the different arms of the interferometer will modify the interference pattern produced when the waves are superposed back together later in the device. By squeezing the phase of the light waves propagating in an interferometer, it is possible to enhance the measurement accuracy of the device by reducing the noise in the system. This effect is exploited in the LIGO (laser interferometer gravitational-wave observatory) interferometer that has been used to detect gravitational waves [36]. These waves stretch and compress *space* in the different arms of the interferometer, which leads to changes in the interference observed with the interferometer. In the case of LIGO, using squeezed states of light was very beneficial as the effect of a gravitational wave can be *extremely* small – sometimes on the order of 10^{-19} m, which is around 1/10,000th of the width of a proton!

The details of generating squeezed states are outside of the scope of this course. At a high level, the generation of squeezed states is usually accomplished through some nonlinear effect. For optical systems, this can be accomplished using certain nonlinear materials that can allow for a mixing process to occur between a weak *signal* (sometimes just the vacuum fields) and a strong *pump* tone (see [11] for an introductory account). At microwave frequencies, a similar process can be achieved with specially-designed parametric amplifiers [37,38]. These devices can find potential use with the kinds of superconducting circuit systems we have learned about within this course.

11.5.2 Quantum Ghost Imaging Systems

In comparison to other areas of quantum information processing technologies (e.g., quantum computers or quantum communication systems), quantum sensing systems that utilize complex quantum states are not as mature. One reason for this is that quantum sensing systems by their nature *must* interact with an external system in a somewhat "uncontrolled" manner. This can make the issue of decoherence of the quantum states involved even more challenging to cope with than in other quantum technologies. As a result, experimental realizations of a substantial quantum advantage with a quantum sensor in comparison to classical techniques have not been widely observed yet. Despite this, the field of quantum sensing is still growing and there are theoretical proposals for technologies that are of significant interest for practical applications. Here, we will discuss one interesting quantum sensing approach; namely, a *quantum ghost imaging system*.

A quantum ghost imaging system uses correlations between entangled photons to image a target object. In its most simple configuration, a quantum ghost imaging system begins by generating a pair of entangled photons in a Bell state. This is typically achieved using *spontaneous parametric down-conversion* by sending a pump signal into a nonlinear crystal. The two photons leave the nonlinear crystal propagating in different directions (this occurs due to conservation of momentum), with both photons then collected into an optical setup to route the photons as needed.

One photon is sent to interact with the target object that is to be imaged. Behind the target is a "bucket detector", that is used to receive the photon. This bucket detector is large so that any photon that is able to pass through the target object will be received by it. A device is connected to this bucket detector that then keeps track of every time that a photon is received by the detector. By scanning the location of the photon propagation path with respect to the target object, different parts of the object can be interacted with.

The second of the entangled photons is sent through a separate optical path that does not interact with the target object. After passing through the optical path, this single photon is directed toward a multi-pixel detector that has a fine spatial resolution. This detector tracks when and where a photon is received, which can then be compared with the time history of photon detections made at the bucket detector. To enable this comparison to be made, the path length of the second photon's optical path can be roughly matched to that of the first photon to enable correlations to be made between the receive history of the two photons more readily.

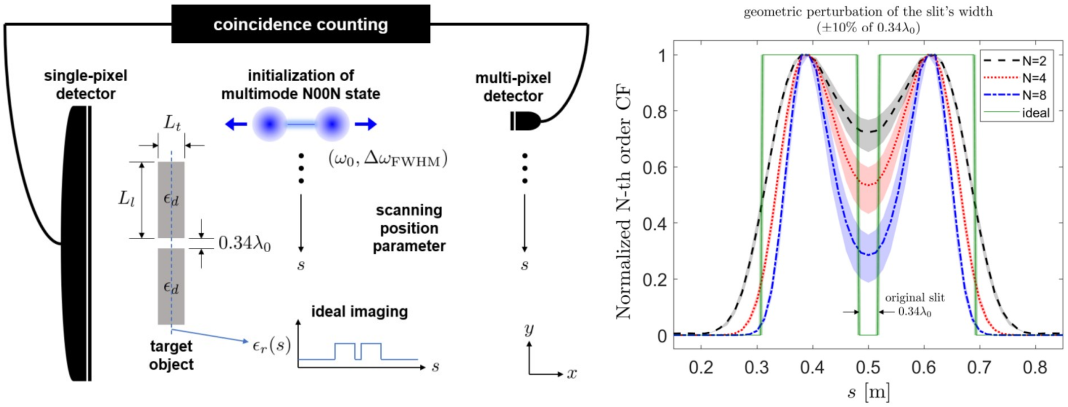
A spatially-resolved image of the target object can now be reconstructed by calculating the correlation of when photons are detected between the two detectors. Importantly, the spatial resolution of the image is set by the photon in the second optical path that does *not* interact with the target object and that is received by the multi-pixel detector. This property can be exploited in interesting ways for certain applications. For instance, when imaging biological media it is generally difficult to penetrate deep into a target and achieve fine spatial resolution. The reason for this is that to achieve fine spatial resolution one must use a high frequency photon, but these are more prone to loss in the biological media that prevents them from penetrating through a thick object. With a quantum ghost imaging setup, a lower frequency photon that *can* penetrate through the object can be directed to the target object and the bucket detector while a higher frequency photon can be sent to the multi-pixel detector. This provides greater penetration depth for a given spatial resolution than is possible with traditional imaging techniques.

It is also possible, in principle, to use higher entangled states to achieve further performance boosts in a quantum ghost imaging setup. One particular state that has been studied significantly is a *NOON state*, which has the mathematical form of

$$|\psi\rangle_{\mathrm{NOON}} = \frac{1}{\sqrt{2}}\Big(|N\rangle_1|0\rangle_2 + |0\rangle_1|N\rangle_2\Big). \tag{11.61}$$

By investigating higher-order correlation functions with this kind of state, it is possible to enhance the resolution of the quantum ghost imaging system beyond the classical *diffraction limit* (this is usually referred to as *super-resolution*). Although NOON states are of great interest, they are quite fragile and are very difficult to generate and measure experimentally. As a result, we show here simulation results of a NOON state quantum ghost imaging setup rather than experimental results.

The basic setup of the quantum ghost imaging system is shown in Fig. 11.7. In this numerical example, a 2D analysis is used that assumes all objects are infinitely-long in the z-dimension (into and out of the page). A dielectric object with a subwavelength slit is placed in front of the bucket detector. The simulation is then performed for various values of N and the correlation functions are computed to reconstruct the image. It is seen that by increasing N, the subwavelength slit (that is below the diffraction limit) becomes increasingly better resolved, which is an indication of super-resolution.



(a)

(b)

Figure 11.7: (a) Setup of a 2D quantum ghost imaging system using a NOON state and (b) simulation results of the correlation function. The higher values of N achieve a super-resolution result by detecting the slit that is smaller than the diffraction limit. A classical imaging system that does not have super-resolution would not be able to detect the slit at all. (Images courtesy of D.-Y. Na, Pohang University of Science and Technology.)

11.6 Practice Problems

1. An infinitely-deep potential well with width L has been prepared in the pure state

$$|\psi\rangle = \frac{1}{\sqrt{3}}|\psi_1\rangle - i\sqrt{\frac{2}{3}}|\psi_2\rangle, \tag{11.62}$$

where $|\psi_j\rangle$ denotes the jth stationary state of the system. For this scenario, answer the following questions.

   (a) Find the matrix representation of the position operator $\hat{x}$ using as basis the states $|\psi_1\rangle$ and $|\psi_2\rangle$.

   (b) Find the density matrix representation of $|\psi\rangle$ using as basis the states $|\psi_1\rangle$ and $|\psi_2\rangle$.

   (c) Compute the expectation value of position for state $|\psi\rangle$ using the density matrix.

   (d) Show that the density matrix found in (b) corresponds to a pure state by using properties of the density matrix.

2. Independently study a quantum topic of interest to you and prepare a presentation and a formal written report designed to teach someone with a similar background in quantum mechanics as yourself about your selected topic. Suggested project topics and initial resources for learning about the topics are included below.

   (a) Quantum mechanics of finite potential wells and electron tunneling through a potential barrier (Suggested Reference: D. A. B. Miller, Ch. 2.8 – 2.9).

   (b) Approximate solutions using time-independent perturbation theory and tight binding models (Suggested Reference: D. A. B. Miller, Ch. 6.3 – 6.5).

(c) Effective mass theory of an electron in a crystal (Suggested Reference: D. A. B. Miller, Ch. 8.1 – 8.3, 8.5, and 8.6).

(d) Quantum mechanics of angular momentum and relation to "s, p, d, f, etc." atomic orbitals (Suggested Reference: D. A. B. Miller, Ch. 9).

(e) Quantum mechanics of a hydrogen atom (Suggested Reference: D. A. B. Miller, Ch. 10 or D. Griffiths, Ch. 4.1 – 4.2. Note: This is an advanced topic, you will want to be quite comfortable with solving partial differential equations.).

(f) Use of the density matrix to describe the interaction of light with a two-level system (Suggested Reference: D. A. B. Miller, Ch. 14.5).

(g) Quantization of electromagnetic fields (Suggested Reference: D. A. B. Miller, Ch. 15.3 – 15.7. Note: Not everything in here has to be covered.).

(h) Example use of Bell's inequalities that supports the Copenhagen interpretation of quantum mechanics (Suggested Reference: D. A. B. Miller, Ch. 19.1).

(i) Particle in a 3-D box and relation to quantum dots (just a conceptual relation, not detailed mathematics) (Suggested Reference: Chemistry LibreTexts).

(j) Introduction to phase and/or flux superconducting artificial atoms (Suggested References: N. Langford's lecture notes [Section 4.3 and 4.4], Krantz's intro to superconducting qubits [Section II-B-2], S. Girvin's lecture notes [Section 4.2], and/or Kockum and Nori's superconducting qubit chapter [Ch. 17.3.2 and 17.3.3]).

(k) Design of two-qubit gates in superconducting processors (Suggested Reference: Krantz's intro to superconducting qubits [Section IV-E – IV-G, not all of these gates need to be covered])

(l) Measurement of the state of a transmon qubit via dispersive readout (Suggested References: Roth's intro to transmon qubits [Section about interfacing transmon qubits with other circuitry], Krantz's intro to superconducting qubits [Section V], and some more mathematical details in Gu's review on superconducting quantum circuits [Section 4.2.4 and B.3]).

(m) Use of Purcell filters for faster dispersive readout (Suggested Reference: Krantz's intro to superconducting qubits [Section V-D, and some earlier sections in V]).

(n) Introduction to lasers (Suggested Reference: Chemistry LibreTexts and related pages).

(o) Quantum communication protocol of superdense coding (Suggested Reference: Qiskit tutorial).

(p) Introduction to the quantum Fourier transform and its applications (Suggested Reference: Qiskit tutorial).

(q) Introduction to the Deutsch-Joza algorithm (Suggested Reference: Qiskit tutorial).

(r) Rydberg atom field sensors and radios (Suggested Reference: Holloway's introduction to the subject).

Bibliography

[1] Wikipedia contributors, “Black-body radiation — Wikipedia, the free encyclopedia,” https://en.wikipedia.org/w/index.php?title=Black-body_radiation&oldid=1203021712, 2024, [Online; accessed 5-February-2024].

[2] M. Stone and P. Goldbart, *Mathematics for Physics: A Guided Tour for Graduate Students*. Cambridge University Press, 2009.

[3] L. D. Landau and E. M. Lifshitz, *Mechanics: Course of Theoretical Physics, Volume 1*, 3rd ed. Elsevier Ltd., 1976.

[4] B. Yurke and J. S. Denker, “Quantum network theory,” *Physical Review A*, vol. 29, no. 3, p. 1419, 1984.

[5] S. M. Girvin, “Circuit QED: Superconducting qubits coupled to microwave photons,” *Quantum machines: Measurement and Control of Engineered Quantum Systems*, pp. 113–256, 2014.

[6] U. Vool and M. Devoret, “Introduction to quantum electromagnetic circuits,” *International Journal of Circuit Theory and Applications*, vol. 45, no. 7, pp. 897–934, 2017.

[7] Wikipedia contributors, “Davisson–germer experiment — Wikipedia, the free encyclopedia,” https://en.wikipedia.org/w/index.php?title=Davisson%E2%80%93Germer_experiment&oldid=1179862211, 2023, [Online; accessed 5-February-2024].

[8] D. A. B. Miller, *Quantum Mechanics for Scientists and Engineers*. Cambridge University Press, 2008.

[9] D. J. Griffiths and D. F. Schroeter, *Introduction to Quantum Mechanics*. Cambridge University Press, 2018.

[10] Wikipedia contributors, “Quantum harmonic oscillator — Wikipedia, the free encyclopedia,” https://en.wikipedia.org/w/index.php?title=Quantum_harmonic_oscillator&oldid=1197338185, 2024, [Online; accessed 5-February-2024].

[11] M. Fox, *Quantum Optics: An Introduction*. Oxford University Press, 2006.

[12] D. F. Walls and G. J. Milburn, *Quantum Optics*. Springer Science & Business Media, 2007.

[13] Wikipedia contributors, "Stern–gerlach experiment — Wikipedia, the free encyclopedia," https://en.wikipedia.org/w/index.php?title=Stern%E2%80%93Gerlach_experiment&oldid=1203205603, 2024, [Online; accessed 5-February-2024].

[14] D. P. DiVincenzo, "The physical implementation of quantum computation," *Fortschritte der Physik: Progress of Physics*, vol. 48, no. 9-11, pp. 771–783, 2000.

[15] P. Krantz, M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, and W. D. Oliver, "A quantum engineer's guide to superconducting qubits," *Applied Physics Reviews*, vol. 6, no. 2, p. 021318, 2019.

[16] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell *et al.*, "Quantum supremacy using a programmable superconducting processor," *Nature*, vol. 574, no. 7779, pp. 505–510, 2019.

[17] A. M. Dalzell, A. W. Harrow, D. E. Koh, and R. L. La Placa, "How many qubits are needed for quantum computational supremacy?" *Quantum*, vol. 4, p. 264, 2020.

[18] Z. Chen, K. J. Satzinger, J. Atalaya, A. N. Korotkov, A. Dunsworth, D. Sank, C. Quintana, M. McEwen, R. Barends, P. V. Klimov *et al.*, "Exponential suppression of bit or phase errors with cyclic error correction," *Nature*, vol. 595, no. 7867, pp. 383–387, 2021.

[19] R. Barends, J. Kelly, A. Megrant, A. Veitia, D. Sank, E. Jeffrey, T. C. White, J. Mutus, A. G. Fowler, B. Campbell *et al.*, "Superconducting quantum circuits at the surface code threshold for fault tolerance," *Nature*, vol. 508, no. 7497, pp. 500–503, 2014.

[20] F. Solgun, D. P. DiVincenzo, and J. M. Gambetta, "Simple impedance response formulas for the dispersive interaction rates in the effective Hamiltonians of low anharmonicity superconducting qubits," *IEEE transactions on microwave theory and techniques*, vol. 67, no. 3, pp. 928–948, 2019.

[21] Y. Chen, C. Neill, P. Roushan, N. Leung, M. Fang, R. Barends, J. Kelly, B. Campbell, Z. Chen, B. Chiaro *et al.*, "Qubit architecture with high coherence and fast tunable coupling," *Physical Review Letters*, vol. 113, no. 22, p. 220502, 2014.

[22] J. Aumentado, "Superconducting parametric amplifiers: The state of the art in Josephson parametric amplifiers," *IEEE Microwave Magazine*, vol. 21, no. 8, pp. 45–59, 2020.

[23] T. E. Roth, R. Ma, and W. C. Chew, "The transmon qubit for electromagnetics engineers: An introduction," *IEEE Antennas and Propagation Magazine*, vol. 65, no. 2, 2023.

[24] L. Castaños and A. Zuñiga-Segundo, "The forced harmonic oscillator: Coherent states and the RWA," *American Journal of Physics*, vol. 87, no. 10, pp. 815–823, 2019.

[25] J. J. Sakurai and J. Napolitano, *Modern Quantum Mechanics*, 3rd ed. Cambridge University Press, 2021.

[26] N. K. Langford, "Circuit QED – Lecture notes," *arXiv preprint arXiv:1310.1897*, 2013.

[27] A. F. Kockum and F. Nori, "Quantum bits with Josephson junctions," in *Fundamentals and Frontiers of the Josephson Effect.* Springer, 2019, pp. 703–741.

[28] M. H. Devoret and R. J. Schoelkopf, "Superconducting circuits for quantum information: An outlook," *Science*, vol. 339, no. 6124, pp. 1169–1174, 2013.

[29] Y. Wu, W.-S. Bao, S. Cao, F. Chen, M.-C. Chen, X. Chen, T.-H. Chung, H. Deng, Y. Du, D. Fan *et al.*, "Strong quantum computational advantage using a superconducting quantum processor," *Physical Review Letters*, vol. 127, no. 18, p. 180501, 2021.

[30] Wikipedia contributors, "Bloch sphere — Wikipedia, the free encyclopedia," https://en.wikipedia.org/w/index.php?title=Bloch_sphere&oldid=1187689877, 2023, [Online; accessed 5-February-2024].

[31] A. A. Houck, J. A. Schreier, B. R. Johnson, J. M. Chow, J. Koch, J. M. Gambetta, D. I. Schuster, L. Frunzio, M. H. Devoret, S. M. Girvin *et al.*, "Controlling the spontaneous emission of a superconducting transmon qubit," *Physical Review Letters*, vol. 101, no. 8, p. 080502, 2008.

[32] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition.* Cambridge University Press, 2010.

[33] Wikipedia contributors, "Quantum key distribution — Wikipedia, the free encyclopedia," https://en.wikipedia.org/w/index.php?title=Quantum_key_distribution&oldid=1199771503, 2024, [Online; accessed 5-February-2024].

[34] R. Schirhagl, K. Chang, M. Loretz, and C. L. Degen, "Nitrogen-vacancy centers in diamond: Nanoscale sensors for physics and biology," *Annual Review of Physical Chemistry*, vol. 65, no. 1, pp. 83–105, 2014.

[35] C. Holloway, M. Simons, A. H. Haddab, J. A. Gordon, D. A. Anderson, G. Raithel, and S. Voran, "A multiple-band Rydberg atom-based receiver: AM/FM stereo reception," *IEEE Antennas and Propagation Magazine*, vol. 63, no. 3, pp. 63–76, 2020.

[36] J. Aasi, J. Abadie, B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, C. Adams, T. Adams, P. Addesso, R. X. Adhikari *et al.*, "Enhanced sensitivity of the LIGO gravitational wave detector by using squeezed states of light," *Nature Photonics*, vol. 7, no. 8, pp. 613–619, 2013.

[37] B. Yurke, "Squeezed-state generation using a Josephson parametric amplifier," *JOSA B*, vol. 4, no. 10, pp. 1551–1557, 1987.

[38] B. Yurke, P. G. Kaminsky, R. E. Miller, E. A. Whittaker, A. D. Smith, A. H. Silver, and R. W. Simon, "Observation of 4.2-K equilibrium-noise squeezing via a Josephson-parametric amplifier," *Physical Review Letters*, vol. 60, no. 9, p. 764, 1988.