1. Quantum States

- Preskill, John. "Quantum computing and the entanglement frontier." arXiv preprint arXiv:1203.5813 (2012).
- Ballentine, Leslie E. Quantum mechanics: a modern development. 1998.
- Preskill, John. "Lecture notes for Physics 219: Quantum computation." California Institute of Technology 3 (2004): 70. (Chapter 4 on entanglement)

1.1 Introduction

The field of quantum information science has its roots in the careful reasoning and deep thought that went into understanding the foundational concepts of quantum theory. Quantum mechanics is a theory which was invented initially to do a number of "useful" things – to describe the radiation emitted from thermal bodies, the structure of atoms, their spectra, scattering probabilities for subatomic particles, and the properties of matter and radiation. However, already in the 1920s, it raised significant philosphical questions about the nature of reality and required a reworking of basic conceptual frameworks that had been used for centuries to understand the world around us. Thought experiments such like Schrödinger's cat, the work of Einstein, Podolsky, and Rosen on entanglement in 1935, Bohr's complementarity principle, and Heisenberg's microscope thought experiment were developed initially to either create a philosophical scaffolding, or to form a basis for criticizing the nascent theory. These concepts now form the basis of fields in quantum computing, sensing, and communications. It was only in the 1960s and 1970s that some of the questions raised by these early pioneers began to be addressed in a rigorous manner, soon giving rise to the field of quantum communication, information, and computing. The field has come full circle. Today we try to use these philosophical advances in our understanding of nature, to again do useful things – to build quantum sensors to detect fields with unprecedented precision and develop enormously capable computers. Before diving into these new advances however, it is useful and important to remind ourselves of some of the the truly weird properties of quantum theory. We will also review a few practical things are important for describing real phenomena.

1.2 What is a quantum state?

Classical probability theory provides a framework to talk about and make calculations of the likelihood or relative frequency of events. It help us predict outcomes based on what we know about a system. At the heart of this framework is the probability distribution function, which gives us the probabilities of different outcomes of measurements on a given system. For instance, consider the probability distribution function:

$$p(x_1,\ldots,x_N) = \Pr[X_1 = x_1,\cdots,X_N = x_N], \tag{1.1}$$

This function encodes the likelihood that, upon measurement, the random variables $X_1,\cdots,X_N$ yield the results $x_1,\cdots,x_N$. We can use this to predict potential outcomes of measurements. However, it's essential to remember that while the probability distribution offers probabilities for various outcomes, it doesn't usually precisely predict a specific outcome — for instance, it doesn't indicate if X_1 will measure as 1.5 or 2.2 – it only offers the likelihood for either result.

■ **Example 1.1** Heads or tails (Bernoulli trials)

We define an experiment as such: we have an unbiased coin that we toss into the air, it lands, and we check whether it is heads (H) or tails (T). We do this experiment N times, insuring that the trials are independent, resulting in a vector $(x_1, x_2, \cdots, x_N)$ where each x_k is either H or T. The probability distribution

$$p(x_1, x_2, \cdots, x_N)$$

fully characterizes the result of any measurement. Convince yourself that $p(x_1, x_2, \cdots, x_N) = 2^{-N}$. ∎

In quantum mechanics, the theory's mathematical formulation enables us to calculate the probability distribution functions $p(x_1, \ldots, x_N)$ for any measurement.

The first element of this mathematical formulation is the Hilbert space ($\mathcal{H}$), a complex vector space. The main objects we use in the theory are either vectors in this Hilbert space or operators acting on the Hilbert space. The **observables** ($\hat{O}$) which are operators acting on the Hilbert space, can correspond to variables in classical physics (energy, momentum, position) or represent the question we are asking of the system (*e.g.*, "what is your spin?", "where are you?", "Is the voltage 1.5 V?", etc.). The **state** ($\hat{\rho}$) is another operator which *encodes all of our knowledge of the state of the system.* **Pure states**, a category among them, are also indicated by a **state vector** $|\psi\rangle$ with the corresponding state or **density matrix** being $\hat{\rho} = |\psi\rangle\langle\psi|$. All states, including pure states, can be represented via the density matrix. Remarkably, even given as complete knowledge as possible in quantum mechanics, the exact result of a specific measurement may still elude prediction.

This is already a *significant* departure from the classical, mechanistic understanding of the world. For example, while knowing a coin's 50% probability for heads doesn't predict a specific flip's result, theoretically, with sufficient understanding and modeling of the coin's physics within a classical theory, one could predict the outcome of a single coin toss. This is perhaps best captured in Laplace's famous claim "Give me the positions and velocities of all the particles in the universe, and I will predict the future". Randomness in classical theories only arises from our limited knowledge or incomplete information of initial conditions and interactions. In contrast, quantum mechanics introduces a different kind of randomness that persists despite complete knowledge of the state and dynamics.

1.2.1 Combining states and observables

We combine states and observables to make predictions in quantum theory. These predictions are in the form of probabilities or probability distribution functions. The way the two operators, observable and state, are combined is through the trace operation.

First we introduce the trace operation:

Definition 1.1 — Trace. The trace of a matrix A is the sum of its diagonal elements. Mathematically, for an $n \times n$ matrix A, the trace is given by:

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} A_{ii}$$

In the context of quantum mechanics, the trace of an operator $\hat{A}$ in a Hilbert space with an orthonormal basis $\{|\phi_i\rangle\}$ is given by:

$$\mathrm{Tr}[\hat{A}] = \sum_i \langle\phi_i|\hat{A}|\phi_i\rangle,$$

where the sum runs over all basis vectors in the Hilbert space.

For operators with continuous spectrum,

$$\mathrm{Tr}[\hat{A}] = \int \mathrm{d}x \langle x|\hat{A}|x\rangle.$$

Theorem 1.1 — Cyclic property of trace. For any two operators $\hat{A}$ and $\hat{B}$:

$$\mathrm{Tr}[\hat{A}\hat{B}] = \mathrm{Tr}[\hat{B}\hat{A}]$$

This property implies that the trace remains invariant under cyclic permutations of the matrices inside the trace.

Proof. Using the orthonormal basis $\{|\phi_i\rangle\}$, we can express the trace of the product of two operators as:

$$\mathrm{Tr}[\hat{A}\hat{B}] = \sum_i \langle \phi_i|\hat{A}\hat{B}|\phi_i\rangle$$

Expanding the identity operator the same basis orthonormal basis, we have:

$$\hat{1} = \sum_j |\phi_j\rangle\langle\phi_j|$$

Inserting this into our expression for the trace, we get:

$$\mathrm{Tr}[\hat{A}\hat{B}] = \sum_i \sum_j \langle \phi_i|\hat{A}|\phi_j\rangle\langle\phi_j|\hat{B}|\phi_i\rangle$$

Since $\langle \phi_i|\hat{A}|\phi_j\rangle$ and $\langle \phi_j|\hat{B}|\phi_i\rangle$ are just numbers, we rearrange them to prove the cyclic property of the trace:

$$\mathrm{Tr}[\hat{A}\hat{B}] = \sum_i \sum_j \langle \phi_j|\hat{B}|\phi_i\rangle\langle\phi_i|\hat{A}|\phi_j\rangle = \sum_j \langle \phi_j|\hat{B}\hat{A}|\phi_j\rangle = \mathrm{Tr}[\hat{B}\hat{A}].$$

■

Quantum mechanics provides a framework for calculating the probabilities of different measurement outcomes and expected values of observables. Let's consider an observable $\hat{O}$. All observables are represented by Hermitian operators. By the spectral theorem, we can express any Hermitian operator as a sum of projectors:

$$\hat{O} = \sum_k o_k |o_k\rangle\langle o_k|$$

where o_k are real eigenvalues (possible measurement outcomes) and $|o_k\rangle$ are the corresponding eigenstates forming an orthonormal basis.

For each eigenvalue o_k, the corresponding projection operator

$$\hat{E}_k \equiv |o_k\rangle\langle o_k|$$

allows us to calculate the probability of a particular measurement outcome o_k when the system is in state $\hat{\rho}$:

$$p(\hat{O} = o_k) = \mathrm{Tr}[\hat{\rho}\hat{E}_k] = \langle o_k|\hat{\rho}|o_k\rangle.$$

The expected value of the observable $\hat{O}$ is then given by the weighted average of all possible outcomes:

$$\begin{aligned}\langle\hat{O}\rangle &= \sum_k o_k p(\hat{O} = o_k) \\ &= \sum_k o_k \langle o_k|\hat{\rho}|o_k\rangle \\ &= \mathrm{Tr}[\hat{\rho}\sum_k o_k|o_k\rangle\langle o_k|] \\ &= \mathrm{Tr}[\hat{O}\hat{\rho}]\end{aligned}$$

This last expression provides a general formula for calculating the expected value of any observable $\hat{O}$ when the system is in state $\hat{\rho}$.

■ **Example 1.2** The position operator

Given a particle in one dimension, we consider the operator $\hat{X}$ representing its position. Since the outcome of measuring the position of the particle is a real number $x \in \mathbb{R}[-\infty,\infty]$, we associate with each of these possible positions a distinct and orthogonal eigenvector $|x\rangle$, so

$$\hat{X} = \int_{-\infty}^{\infty} x|x\rangle\langle x|\,\mathrm{d}x.$$

■

Exercise 1.1 — Projection Operators as Observables. Consider a projection operator $\hat{P} = |\psi\rangle\langle\psi|$, where $|\psi\rangle$ is a normalized state vector.

1. Show that $\hat{P}$ is Hermitian.
2. Find the eigenvalues and eigenvectors of $\hat{P}$.
3. If we consider $\hat{P}$ as an observable, what are the possible measurement outcomes and their corresponding probabilities when measuring a system in state $|\phi\rangle$?
4. Given an arbitrary state $|\phi\rangle = \alpha|\psi\rangle + \beta|\psi_\perp\rangle$, where $|\psi_\perp\rangle$ is orthogonal to $|\psi\rangle$ and $|\alpha|^2 + |\beta|^2 = 1$, calculate the expectation value $\langle\hat{P}\rangle$.
5. Express the variance of this observable, $\mathrm{Var}(\hat{P}) = \langle\hat{P}^2\rangle - \langle\hat{P}\rangle^2$, in terms of $|\alpha|^2$.

■

We can summarize these properties of the density matrix and their physical interpretation:

1. Unit trace: $\mathrm{Tr}[\hat{\rho}] = 1$
   - This is a statement about probabilities. If we have an orthonormal basis $\{|\alpha_k\rangle\}$, then $p_k = \langle\alpha_k|\hat{\rho}|\alpha_k\rangle$ is the probability of being in state k. The sum $\sum_k p_k = 1$, is exactly the trace. *(probabilities add up to 1)*
2. The density matrix is Hermitian and positive.
   - $\hat{\rho}^\dagger = \hat{\rho}$. *(probabilities are real)*
   - $\langle\psi|\hat{\rho}|\psi\rangle \geq 0$ for all vectors $|\psi\rangle$. *(probabilities are positive)*

Some (but not all) density matrices can be expressed as $\hat{\rho} = |\psi\rangle\langle\psi|$. The vector $|\psi\rangle$ is called a state vector and $\hat{\rho}$ is a *pure state*. A state that is not pure is called a *mixed state*. For example, we may not know whether a system is in a pure state $|\psi_1\rangle$ or $|\psi_2\rangle$, but can assign probabilities p_1 and p_2 for these two different possibilities. The state of system would then be the linear combination $\hat{\rho} = p_1|\psi_1\rangle\langle\psi_1| + p_2|\psi_2\rangle\langle\psi_2|$.

1. Conservation of trace over unitary evolution: $\hat{\rho}(t) = \hat{U}(t)\hat{\rho}_0\hat{U}^\dagger(t)$, then using cyclic property of trace we have that $\mathrm{Tr}[\hat{\rho}(t)] = \mathrm{Tr}[\hat{U}(t)\hat{\rho}_0\hat{U}^\dagger(t)] = \mathrm{Tr}[\hat{U}^\dagger(t)\hat{U}(t)\hat{\rho}_0] = \mathrm{Tr}[\hat{\rho}_0]$. *(probabilities add up to 1 for all time)*
2. **Purity**: We define the purity $P = \mathrm{Tr}[\hat{\rho}^2]$. $P = 1$ if and only if $\hat{\rho}$ is a pure state ($\hat{\rho} = |\psi\rangle\langle\psi|$). proof:
   - $\hat{\rho} = |\psi\rangle\langle\psi|$ implies $\mathrm{Tr}[\hat{\rho}^2] = 1$ trivially.
   - Assume $\mathrm{Tr}[\hat{\rho}^2] = 1$. Since $\mathrm{Tr}[\hat{\rho}] = 1$ and all eigenvalues p_k of $\hat{\rho}$ are greater than 0 and less than 1 (property 1 and 2), for both $\sum_k p_k = 1$ and $\sum_k p_k^2 = 1$, we require that $p_{k'} = 1$ for some k' and zero otherwise, so $\hat{\rho}$ is a pure state.
3. Convex combinations of density matrices are also density matrices.
   - If $\{\hat{\rho}_s\}$ are density matrices, then $\hat{\rho} = \sum_s p_s\hat{\rho}_s$ is also a valid density matrix if $p_s \geq 0$ and $\sum_s p_s = 1$.

Exercise 1.2 — States prepared using different procedures.

1. You're handed two boxes (1) and (2), (1) emits photons with polarization $|H\rangle$ or $|V\rangle$randomly and with equal probability. (2) emits photons with polarization $|+\rangle$ or $|-\rangle$ (where$|\pm\rangle = (|H\rangle \pm |V\rangle)/\sqrt{2}$) randomly and with equal probability.
   a. What are the density matrices $\hat{\rho}_1$ and $\hat{\rho}_2$ representing the state of the photon in the basis with vector representation $|H\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|V\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.
   b. Can you think of an experiment to distinguish between the two boxes?
2. **(Tricky)** Someone has made a box that emits photons with polarization $|H\rangle$ with $p_H = 0.99$ and $|V\rangle$ with $p_V = 0.01$. Design a box that emits photons in state $|\psi\rangle_1$ or $|\psi\rangle_2$ randomly and with equal probability, and that is indistinguishable from the original box. What are $|x\rangle$ and $|y\rangle$, expressed in the original basis?

■

1.2.2 Measurements

We will study measurements in quantum mechanics in significantly greater detail in later chapters. For now we will just consider projective measurements and how this formalism allows us to use the state and an observable to find the probability distribution for measurement result.

Definition 1.2 — Projective Measurements. Projective Measurements

A measuring apparatus gives us a measurement result s, and changes the state of the system. In a projective measurement, we characterize the measurement with set of projection operators $\{\hat{E}_s\}$, with the following properties: (1) $\hat{E}_s = \hat{E}_s^\dagger$, (2) $\sum_s \hat{E}_s = 1$, (3) $\hat{E}_s\hat{E}_{s'} = \hat{E}_s\delta_{ss'}$.

Measurement is the processes by which a state vector $|\psi\rangle$ is transformed:

$$|\psi\rangle \xrightarrow[\text{measurement}]{} \left\{ |\psi_s\rangle = \frac{\hat{E}_s|\psi\rangle}{\sqrt{\langle\psi|\hat{E}_s|\psi\rangle}} \right.$$

with the probability of obtaining measurement result s leading to a final state $|\psi_s\rangle$ given by $p_s = \langle\psi|\hat{E}_s|\psi\rangle$. Extending this definition to density matrices, we have:

$$\hat{\rho} \underbrace{\longrightarrow}_{measurement} \left\{\hat{\rho}_s = \frac{\hat{E}_s\hat{\rho}\hat{E}_s^\dagger}{p_s}\right.$$

with probability $p_s = \mathrm{Tr}[\hat{E}_s\hat{\rho}]$.

■ **Example 1.3** Measuring position

Consider the harmonic oscillator with Hamiltonian

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2.$$

The position operator has a real continuous spectrum of eigenvalues $x \in \mathbb{R}[-\infty,\infty]$. We assume we have a detector that tells us whether the particle is inside a bin $(x_s, x_s + \Delta x] \subset \mathbb{R}[-\infty,\infty]$ for $x_s = s\Delta x$, for every integer s. The associated measurement operator is $\hat{E}_s = \int_{x_s}^{x_s+\Delta x} dx'|x'\rangle\langle x'|$. Verify that the set of operators $\{\hat{E}_s\}$ form a valid set of operators defining a projective measurement. Experimentally, such a system would be realized were we have a detector that takes a state, and gives us a number s, corresponding to the position of the operator. The probability of measuring s signifying that the particle has a position of $x_s < x \leq x_s + \Delta x$ is given by

$$p_s = \mathrm{Tr}[\hat{E}_s\hat{\rho}].$$

■

1.3 Composite Systems and Entanglement

Let's say we have two systems, system A and system B. We have a way of describing these two systems, including their Hilbert spaces, operators, and other relevant quantum mechanical properties. We should also have a way of describing both systems at once as a single larger *composite* or *bipartite* system.

For example, if the first system is in state $|\alpha_i\rangle$ and the second one is in state $|\beta_j\rangle$, then the joint state of the composite system is $|\alpha_i\rangle \otimes |\beta_j\rangle$, which we usually write more compactly as just $|\alpha_i\rangle|\beta_j\rangle$. This new vector is obtained by performing the tensor product operation. It's a vector that lives in a larger tensor product Hilbert space, which combines the Hilbert spaces of the individual systems. If the states $|\alpha_i\rangle$ and $|\beta_j\rangle$ form orthonormal bases for the N_A- and N_B-dimensional Hilbert spaces of system A and B, then $|\alpha_i\rangle|\beta_j\rangle$ forms an orthonormal basis for the the new tensor product space $\mathcal{H}_{\text{tot}} = \mathcal{H}_A \otimes \mathcal{H}_B$, which consequently is of dimension $N_A \times N_B$.

Since in quantum mechanics, any linear combination of these joint states is also a valid state, pure states of the bipartite will in general take the form:

$$|\Psi\rangle = \sum_{i,j} c_{ij}|\alpha_i\rangle|\beta_j\rangle \in \mathcal{H}_{\text{tot}}$$

where c_{ij} are complex coefficients.

1.3.1 Density Matrix as the Partial Trace of a Pure State

Consider a composite system, composed of two subsystems A and B, with a state

$$|\Psi\rangle = \sum_{ij} c_{ij}|\alpha_i\rangle|\beta_j\rangle \tag{1.2}$$

Here, $\{|\alpha_i\rangle\}$ and $\{|\beta_i\rangle\}$ represent orthonormal bases for systems A and B respectively. Generally, observables or operators in this joint system influence both subsystems. However, suppose we only have experimental access to observables that act on system A. This situation might arise when system B is physically distant from A, or in scenarios where one system's properties are manipulated using another. For example we may be using an optical field described by system A to make measurements on or modify the mechanical position or spin described by B. In such cases, we deal with operators of the form $\hat{O}_A \otimes \hat{1}_B$ – operators that solely influence system A and whose expectation values are independent of system B's state. The expected value of these observables can be written as:

$$\begin{aligned}\langle\Psi|\hat{O}_A \otimes \hat{1}_B|\Psi\rangle &= \sum_{ij}\sum_{i'j'} c_{ij}^* c_{i'j'} \langle\alpha_i|\hat{O}_A|\alpha_{i'}\rangle\langle\beta_j|\beta_{j'}\rangle \\ &= \sum_{ij}\sum_{i'} c_{ij}^* c_{i'j} \langle\alpha_i|\hat{O}_A|\alpha_{i'}\rangle \\ &= \sum_{ii'}\left(\sum_j c_{ij}^* c_{i'j}\right)\langle\alpha_i|\hat{O}_A|\alpha_{i'}\rangle \end{aligned} \tag{1.3}$$

Assuming no knowledge about system B, let's say we aim to determine the expected value of a local operator $\hat{O}_A$, given the state of system A. We represent the system's state through a **density matrix**, $\hat{\rho}_A$. The observable's expectation can be calculated as:

$$\mathrm{Tr}[\hat{\rho}_A\hat{O}_A] = \sum_{i'}\langle\alpha_{i'}|\hat{\rho}_A\hat{O}_A|\alpha_{i'}\rangle = \sum_{ii'}\langle\alpha_{i'}|\hat{\rho}_A|\alpha_i\rangle\langle\alpha_i|\hat{O}_A|\alpha_{i'}\rangle \tag{1.4}$$

On comparing equations (1.3) and (1.4), we observe that as long as we restrict ourselves to observables acting only on system A, the measurement outcome is adequately represented if we assume system A is in state $\hat{\rho}_A = \sum_{ii'}\sum_k c_{ik}^* c_{i'k}|\alpha_{i'}\rangle\langle\alpha_i|$.

The composite system's state is given by

$$\begin{aligned}\hat{\rho} &= |\Psi\rangle\langle\Psi| \\ &= \sum_{ij}\sum_{i'j'} c_{ij}^* c_{i'j'}|\alpha_{i'}\rangle|\beta_{j'}\rangle\langle\alpha_i|\langle\beta_j|\end{aligned}$$

We find that the operation $\sum_k\langle\beta_k|\hat{\rho}|\beta_k\rangle = \sum_{ii'}\sum_k c_{ik}^* c_{i'k}|\alpha_{i'}\rangle\langle\alpha_i|$ provides us with $\hat{\rho}_A$. This operation, $\sum_k\langle\beta_k|\hat{\rho}|\beta_k\rangle$, is known as a **partial trace** and is denoted as $\mathrm{Tr}_B[\hat{\rho}]$. Thus, the state of the subsystem can be represented as

$$\hat{\rho}_A = \sum_j\langle\beta_j|\hat{\rho}|\beta_j\rangle \equiv \mathrm{Tr}_B[\hat{\rho}]. \tag{1.5}$$

Let's make things more concrete with a specific quantum system: a two-qubit system.

Exercise 1.3 — Two-Qubit System and Joint Measurement. Consider a two-qubit system, Q_1 and Q_2, which is prepared in the Bell state given by

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \tag{1.6}$$

Here, $|00\rangle$ and $|11\rangle$ are basis states representing both qubits being in state 0 and both

qubits being in state 1, respectively.

1. Compute the density matrix $\hat{\rho}_{Q_1}$ for the first qubit Q_1.
2. Suppose we want to measure the z-component of the spin (also known as the Pauli-Z operator) of the first qubit, represented by the observable $\hat{Z}$. The Pauli-Z operator can be represented in the computational basis as:

$$\hat{Z} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{1.7}$$

Calculate the expected value of this measurement.
3. Now, consider a joint measurement on both qubits using the Pauli-ZZ operator, defined as the tensor product of the Pauli-Z operators acting on the two subsystems. Calculate the expected value of this joint measurement. ■

1.3.2 Entanglement

Entanglement is one of the most radical notions in quantum mechanics.

Definition 1.3 — Entangled States. An **entangled pure state** is a state of a system that cannot be written as a product state:

$$|\Psi\rangle \neq |\alpha\rangle|\beta\rangle.$$

Similarly, an **entangled mixed state** is a state that cannot be expressed as a statistical (convex) mixture of product states

$$\hat{\rho} \neq \sum_k p_k \hat{\rho}_{A,k} \otimes \hat{\rho}_{B,k}.$$

Entangled states have counter-intuitive properties. Why is entanglement such a radical notion? For an entangled state, even if the state $|\Psi\rangle$ of the joint system is known perfectly, the reduced density matrix $\hat{\rho}_A$ describing a subsystem will be mixed. Loosely speaking, even if we know perfectly the state of the two systems taken together, we may still know very little about the state of each subsystem taken separately.

1.3.3 Bell's Theorem

1.3.3.1 Background

Quantum mechanics is a statistical theory and has an irreducible randomness. Repeated measurements on identical quantum states can give different outcomes. Quantum theory only gives us the relevant probability distributions. This raises the question: does quantum mechanics emerge from some deeper theory with additional "hidden variables" that determine the outcomes of experiments? Is there some more fundamental theory from which quantum mechanics emerges, in analogy to how statistical physics emerges from an underlying deterministic set of classical theories? Theories postulating an underlying theory (with "hidden variables") have been proposed to provide a more fundamental explanation.

Exercise 1.4 — EPR State. In this exercise we go over some states and their properties which are useful for understanding EPR's argument and Bell's inequality.

1. Consider two entangled particles 1 and 2 in the EPR state where their relative position and total momentum are precisely correlated:

$\hat{X}_1 + \hat{X}_2 = 0$ and $\hat{P}_1 - \hat{P}_2 = 0$
   a. Show that $[\hat{X}_1 + \hat{X}_2, \hat{P}_1 - \hat{P}_2] = 0$, confirming that this is a valid quantum state.
   b. If particle 1's position is measured to be x, what can you conclude about particle 2's position? What about their momenta if particle 1's momentum is measured instead?
2. Write the (unnormalized) wavefunction for the EPR state. (*Hint: A state with the particle 1 at position x_1 and particle 2 at position x_2 is given by $|\Psi\rangle = |\hat{X}_1 = x_1\rangle|\hat{X}_2 = x_2\rangle$.*)

■

An influential argument challenging the completeness of quantum mechanics was presented by Einstein, Podolsky, and Rosen (EPR) in 1935. EPR argued that quantum mechanics faces a dilemma: either it is "incomplete" or it violates locality (the principle that distant objects cannot influence each other instantaneously). Here, "incomplete" means that quantum mechanics fails to account for all "elements of physical reality". The argument considers two particles in an entangled state where their relative position and total momentum are precisely correlated, such that $\hat{X}_1 + \hat{X}_2 = 0$ and $\hat{P}_1 - \hat{P}_2 = 0$. This is a state that's allowed by quantum physics since the operators $\hat{X}_1 + \hat{X}_2$ and $\hat{P}_1 - \hat{P}_2$ commute and therefore their simultaneous eigenstate with eigenvalues 0 for both is a valid state. It is now known as an EPR state. After the particles separate, the observer at one end has the choice of measuring *either* $\hat{X}_1$, *or* $\hat{P}_1$ of particle 1. Let's assume that they decide to measure $\hat{X}_1$. Whatever the result of the measurement of $\hat{X}_1$, we immediate know precisely what value of a $\hat{X}_2$ measurement would be if the second observer were to measure it. Assuming there are nonlocal effects allowed between the two particles, *i.e.*, no so-called "Spooky Action at a Distance", this means that position of the second particle, $\hat{X}_2$, is fully determined and is thus an "element of physical reality."[1] On the other hand, the observer at position 1 could just as well have decided to measure momentum instead, and so we can apply the same argument to $\hat{P}_2$: it to has a certain value and is an element of reality. But this calls into question one of the fundamental precepts of quantum mechanics which states that noncommuting observable can not be assigned values simultaneously with certainty. It also contradicts the view that the quantum state provides a complete description of a physical system, as it cannot simultaneously represent definite values for both position and momentum.

In the wake of the EPR argument, physicists continued to grapple with the completeness of quantum mechanics. Some physcists, inspired by Einstein's critique, tried to develop *hidden variable* theories that could provide a "complete," deterministic description of quantum phenomena. These theories proposed that underlying the quantum world were hidden variables that, if known, would allow precise predictions of particle behavior. In 1932, John von Neumann claimed to prove that hidden variable theories were generally mathematically impossible, ostensibly closing the door on this line of inquiry. But then in 1952, David Bohm proposed a hidden variable theory that appeared to work, reproducing all the predictions of quantum mechanics, clearly showing that von Neumann's proof was incorrect.[2] Bohm's theory, while deterministic, was nonlocal, and allowed for instantaneous influences between distant particles. The natural question then became: is there "complete" and "local" hidden variable theory?

[1] EPR state that if we can predict with certainty (i.e., with probability equal to unity) the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity.

[2] John von Neumann's proof was "foolish" according to Bell's later analysis.

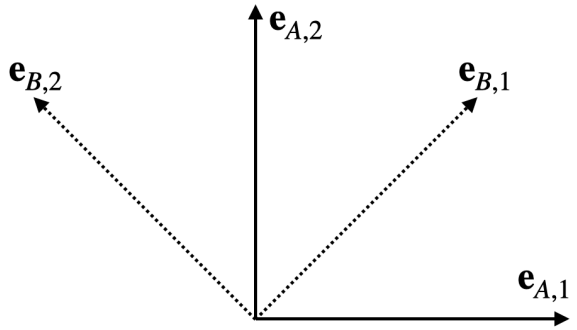

Figure 1.1: Measurement directions for the CHSH inequality. The vectors $\mathbf{e}_{A,1}$ and $\mathbf{e}_{A,2}$ represent the measurement directions for system A. Vectors $\mathbf{e}_{B,1}$ and $\mathbf{e}_{B,2}$ represent those for system B. The vectors are arranged in the x-y plane with 45-degree separations.

This is where John Bell's work came in – he proved that if quantum mechanics is correct, then there can be no hidden variable theory that is both deterministic and local. Importantly, Bell's work showed that the issue was not merely philosophical but experimentally testable. His theorem demonstrated that any hidden variable theory satisfying certain reasonable locality conditions must satisfy inequalities, now known as Bell inequalities, in its predictions for the outcomes of certain experiments. Quantum mechanics, on the other hand predicts violations of these inequalities. This insight transformed the landscape of quantum foundations. It shifted the debate from purely theoretical and philosophical grounds to the realm of experiment. Over the subsequent decades, a series of increasingly sophisticated experiments were conducted to test these predictions. While these experiments consistently confirmed the quantum mechanical predictions, by violating the Bell inequalities, they were subject to certain loopholes that could allow for alternative explanations. The two most significant were the "locality" loophole (the possibility that the detector settings could be communicated between the locations of the two particles) and the "detection" loophole (given inefficient measurements, there is always the possibility that the detected values were not a fair sample of all the values). It wasn't until 2015 that experiments were finally conducted that closed both of these loopholes simultaneously, providing the most conclusive evidence to date against local hidden variable theories. Nearly sixty years after the original publication of Bell's paper, the 2022 Nobel Prize in Physics was awarded to Alain Aspect, John Clauser, and Anton Zeilinger for their pioneering experiments in this field.

1.3.3.2 Proving Bell's Theorem: CHSH Inequality

We consider two separated systems with local observers, Alice (A) and Bob (B). At location A, we have local observables $\hat{a}_1$ and $\hat{a}_2$, while at location B, we have $\hat{b}_1$ and $\hat{b}_2$. Each of these observables yield a value of ± 1.

Let's assume these observables have pre-assigned values a_1, a_2, b_1, and b_2. We collect the results and calculate $C = (a_1 + a_2)b_1 + (a_2 - a_1)b_2$. Note that $a_1 + a_2$ or $a_2 - a_1$ will be 0, and the other will be ± 2. Consequently, $C = \pm 2$. If we repeat this measurement multiple times and average the result, we get $-2 \leq \langle C \rangle \leq 2$, or $|\langle C \rangle| \leq 2$. It's important to recognize that there's an assumption of hidden variable theory present in this reasoning, since we assumed that that all of the variables have some definite values.

Now, let's calculate the expected value of C for a specific quantum mechanical setting. We

assume that the observables are the spin along certain axes $\mathbf{e}_{j,k}$ where $j = A, B$, and $k = 1, 2$. Thus, $\hat{a}_k = \hat{\sigma}_A \cdot \mathbf{e}_{A,k}$ and $\hat{b}_k = \hat{\sigma}_B \cdot \mathbf{e}_{B,k}$. We also assume that the two spins at locations A and B are in an entangled state $|\Psi^-\rangle = 2^{-1/2}(|01\rangle - |10\rangle)$.

Exercise 1.5 Show that $\langle\Psi^-|\hat{\sigma}_i\hat{\sigma}_j|\Psi^-\rangle = -\delta_{ij}$. Use this to demonstrate that $\langle\Psi^-|(\hat{\sigma}_A \cdot \mathbf{e}_A)(\hat{\sigma}_B \cdot \mathbf{e}_B)|\Psi^-\rangle = -\mathbf{e}_A \cdot \mathbf{e}_B$. ■

We choose the vectors for the observables $\hat{a}_1$, $\hat{a}_2$, $\hat{b}_1$, $\hat{b}_2$ as follows (see Figure 1.1): $\mathbf{e}_{A,1}$ is along the x-axis, $\mathbf{e}_{B,1}$ is rotated by 45 degrees, $\mathbf{e}_{A,2}$ by 90 degrees, and $\mathbf{e}_{B,2}$ by 135 degrees about the z-axis. Evaluating C, we find:

$$\begin{aligned} C &= \langle\hat{a}_1\hat{b}_1\rangle + \langle\hat{a}_2\hat{b}_1\rangle + \langle\hat{a}_2\hat{b}_2\rangle - \langle\hat{a}_1\hat{b}_2\rangle \\ &= -\mathbf{e}_{A,1}\cdot\mathbf{e}_{B,1} - \mathbf{e}_{A,2}\cdot\mathbf{e}_{B,1} - \mathbf{e}_{A,2}\cdot\mathbf{e}_{B,2} + \mathbf{e}_{A,1}\cdot\mathbf{e}_{B,2} \\ &= -\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}} = 2\sqrt{2}. \end{aligned}$$

This result violates the inequality $|\langle C\rangle| \leq 2$ derived under the hidden variable assumption, demonstrating Bell's theorem.

1.3.4 Schmidt Decomposition

An entangled state $|\Psi\rangle$ can't be expressed as $|\alpha\rangle|\beta\rangle$. The closest we can come to this form is called the Schmidt decomposition. We can express

$$|\Psi\rangle = \sum_{k=1}^{\min(N_A,N_B)} \sqrt{\lambda_k}|\alpha_k\rangle|\beta_k\rangle,$$

with $\{|\alpha_k\rangle\}$, and $\{|\beta_k\rangle\}$ with $\{0 \leq \lambda_k \leq 1\}$ forming orthonormal sets of vectors in their respective Hilbert spaces. This decomposition is pretty remarkable. Remember that in general, the vector $|\Psi\rangle$ is expressed as $|\Psi\rangle = \sum_{ij} c_{ij}|\alpha_i\rangle|\beta_j\rangle$ and so $N_A \times N_B$ coefficients are required in an arbitrary basis. The Schmidt decomposition tells us that there is a basis within which only need $\min(N_A, N_B)$ coefficients. In fact the actual number of coefficients we need depends on how much entanglement there is. For separable states, the best basis is the one where $|\Psi\rangle$ is obviously a product state $|\alpha_1\rangle|\beta_1\rangle$ and so only single coefficient, $\lambda_1 = 1$, is needed. The Schmidt decomposition effectively cuts through the chase and gives us a representation of the system that is commensurate to how much entanglement there really is in a state. This aspect of the Schmidt decomposition makes it the key component of numerical methods (like DMRG) and representation techniques (such as tensor networks) that have been developed to understand quantum correlated many-body states.

Theorem 1.2 — Schmidt Decomposition. Schmidt Decomposition

Any state $|\Psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ can be expressed as

$$|\Psi\rangle = \sum_{k=1}^{\min(N_A,N_B)} \sqrt{\lambda_k}|\alpha_k\rangle|\beta_k\rangle,$$

with $\{|\alpha_k\rangle\}$, and $\{|\beta_k\rangle\}$ with $\{0 \leq \lambda_k \leq 1\}$ forming orthonormal sets of vectors in their respective Hilbert spaces.

Proof. We start with the density matrix for the full system $\hat{\rho} = |\Psi\rangle\langle\Psi|$. We take a partial trace over subsystem B, so $\hat{\rho}_A = \mathrm{Tr}_B[\hat{\rho}]$. Since this is a valid density matrix for

the subsystem A , it is also Hermitian, and so can be diagonalized. We call this basis where the density matrix is diagonal $\{|\alpha_k\rangle\}$, and so

$$\hat{\rho}_A = \sum_k p_k |\alpha_k\rangle\langle\alpha_k|.$$

In this basis, we can express the original vector as $|\Psi\rangle = \sum_{jk} c_{jk}|\alpha_k\rangle|b_j\rangle$ for some basis $|b_j\rangle$ for system B. Here we play with the order of summation a little:

$$\begin{aligned}|\Psi\rangle &= \sum_{jk} c_{jk}|\alpha_k\rangle|b_j\rangle \\ &= \sum_k |\alpha_k\rangle \left(\sum_j c_{jk}|b_j\rangle\right)\end{aligned}$$

We label these vectors, $\sqrt{\lambda_k}|\beta_k\rangle$, choosing λ_k so that the $|\beta_k\rangle$ are normalized:

$$\sqrt{\lambda_k}|\beta_k\rangle = \sum_j c_{jk}|b_j\rangle.$$

We'll show now that these vectors, $|\beta_k\rangle$ are also orthonormal, so that choosing the express the original wavefunction in this basis of the eigenvectors $\{|\alpha_k\rangle\}$ of $\hat{\rho}_A$ has also in sense "diagonalized" it in the basis for system B. Notice that we can express the partial trace operation also as:

$$\begin{aligned}\hat{\rho}_A &= \mathrm{Tr}_B\left[\sum_{kk'} \sqrt{\lambda_k\lambda_{k'}}|\alpha_k\rangle|\beta_k\rangle\langle\alpha_{k'}|\langle\beta_{k'}|\right] \\ &= \sum_{kk'} \mathrm{Tr}_B\left[\sqrt{\lambda_k\lambda_{k'}}|\beta_k\rangle\langle\beta_{k'}|\right]|\alpha_k\rangle\langle\alpha_{k'}| \\ &= \sum_{kk'} \sqrt{\lambda_k\lambda_{k'}}\langle\beta_{k'}|\beta_k\rangle|\alpha_k\rangle\langle\alpha_{k'}|.\end{aligned}$$

Comparing this expression to the initial $\hat{\rho}_A = \sum_k p_k|\alpha_k\rangle\langle\alpha_k|$, and taking inner products with different $|\alpha_k\rangle$ on both sides, it is clear that $\langle\beta_{k'}|\beta_k\rangle = \delta_{kk'}$, and that $\lambda_k = p_k$. Since $0 \leq p_k \leq 1$, we've proven the proposition. ■

Exercise 1.6 — Entanglement of two qubits.

1. Is the 2-qubit state $|\Psi\rangle = (|00\rangle + |01\rangle + |10\rangle + |11\rangle)/2$ separable? If yes express as a product.
2. Is the 2-qubit state $|\Psi\rangle = (|00\rangle + |01\rangle + |10\rangle - |11\rangle)/2$ separable? If yes, express as a product.
3. Is the N-qubit state

   $$|\Psi\rangle = 2^{-n/2}(|00\ldots00\rangle + |00\ldots01\rangle + |00\ldots10\rangle + |00\ldots11\rangle + \cdots + |11\ldots11\rangle)$$

   separable?
4. We start with a pure state that is *separable*, i.e. can be written as $|\Psi\rangle = |\alpha\rangle|\beta\rangle$. Show that:
   a. The density matrix $\hat{\rho}_A$ is pure.
   b. The system now evolves according to a Hamiltonian that doesn't have any interaction between A and B, i.e., $\hat{H} = \hat{H}_A + \hat{H}_B$, where $\hat{H}_A$ and $\hat{H}_B$ generate

the evolution of system A and B separately. Show that under this evolution, the state $|\Psi\rangle$ remains separable.

5. Starting with a 2-qubit state (qubit A and B) $|\Psi\rangle = (|00\rangle + |01\rangle + |10\rangle + |11\rangle)/2$, we perform a *cPHASE* gate, which results in a state $|\Psi\rangle = (|00\rangle + |01\rangle + |10\rangle + e^{i\phi}|11\rangle)/2$.
   a. Calculate the reduced density matrix $\hat{\rho}_A$ in terms of ϕ.
   b. Calculate the purity of the reduced density matrix, $\mathrm{Tr}[\hat{\rho}_A^2]$.
   c. We can quantify entanglement using the purity of the partial trace. Let $\mathcal{E} = 1 - \mathrm{Tr}[\hat{\rho}_A^2]$. What is the entanglement in as a function of ϕ?

■

2. Transformations and Pictures

2.1 Continuous Unitary Transformations

Unitary transformations are central in describing the evolution of quantum systems. They are also a key concept for describing symmetries in physical systems, an important step to understanding the emergence of various dynamical variables. These transformations can be intuitively understood as rotations within the Hilbert space. The defining characteristic of a unitary transformation, represented by $\hat{U}$, is that it preserves inner products. This means that for any two states $|\psi\rangle$ and $|\phi\rangle$ in the Hilbert space, the inner product before and after the application of the unitary transformation remains unchanged:

$$\langle\psi|\hat{U}^\dagger\hat{U}|\phi\rangle = \langle\psi|\hat{1}|\phi\rangle$$

This implies that $\hat{U}^\dagger\hat{U} = \hat{1}$.

Consider unitary transformations that can be parametrized continuously. The parameter may be time, but can also be something else, for example the rotation angle, a displacement, or the amount of charge present on an island of a superconductor. Taking s for the parameter, such a continuously parameterized unitary transformation is given by $\hat{U}(s)$ for each s, such that $\hat{U}(s)$ is unitary. Moreover, we require that the transformation is the multiplicative over the parameter,

$$\hat{U}(s_1 + s_2) = \hat{U}(s_1)\hat{U}(s_2),$$

and that $\hat{U}(0) = \hat{1}$. These properties make the family of transformations into a **group**.

Let's consider the case of an infinitesmial transformation. For a small change in the parameter s by δs, the transformation can be expressed as:

$$\hat{U}(s + \delta s) = \hat{U}(s)\hat{U}(\delta s)$$

For very small δs, the transformation $\hat{U}(\delta s)$ can be approximated as:

$$\hat{U}(\delta s) = \hat{1} + \frac{\mathrm{d}}{\mathrm{d}s}\hat{U}(0)\delta s + O(\delta s^2)$$

Applying the condition $\hat{U}^\dagger(\delta s)\hat{U}(\delta s) = \hat{1}$, we find that:

$$\frac{\mathrm{d}}{\mathrm{d}s}\hat{U}(0) + \frac{\mathrm{d}}{\mathrm{d}s}\hat{U}^\dagger(0) = 0$$

This leads to the conclusion that the derivative $\frac{\mathrm{d}}{\mathrm{d}s}\hat{U}(0)$ is anti-Hermitian, or in other words equivalent to $i\hat{K}$, where $\hat{K}$ is a Hermitian operator. The operator $\hat{K}$ is called the **generator** of the group of transformations $\hat{U}(s)$. Since $\hat{K}$ is Hermitian, it corresponds to a physical observable. Notice that we obtain a differential equation

$$\frac{\mathrm{d}}{\mathrm{d}s}\hat{U}(s) = \hat{U}(s)i\hat{K}, \tag{2.1}$$

resulting in the expression

$$\hat{U}(s) = \exp(i\hat{K}s).$$

2.1.1 Momentum as the Generator of Displacement

We now apply some of the definitions above to the the single particle in one dimension, to see how momentum as a dynamical variable arises from considering transformations in position space.

Definition 2.1 — Translation operator. Consider the particle in one dimension. The eigenstates of the position operator $\hat{X}$ were defined to be $|x\rangle$ for real numbers x. A translation in space has the following effect:

$$\hat{U}(a)|x\rangle = |x+a\rangle$$

We denote the generator of this transformation as $-\hat{P}/\hbar$, so

$$\hat{U}(a) = \exp\left(-\frac{i\hat{P}}{\hbar}a\right)$$

We considered above the action of the transformation $\exp\left(-\frac{i\hat{P}a}{\hbar}\right)$ on states $|x\rangle$. But how does it act on operators? Consider the position operator $\hat{X}$,

$$\hat{X} = \int_{-\infty}^{\infty} x|x\rangle\langle x|\mathrm{d}x$$

so

$$\hat{U}(a)\hat{X}\hat{U}^{\dagger}(a) = \int_{-\infty}^{\infty} x\hat{U}(a)|x\rangle\langle x|\hat{U}^{\dagger}(a)\mathrm{d}x = \int_{-\infty}^{\infty} x|x+a\rangle\langle x+a|\mathrm{d}x = \int_{-\infty}^{\infty}(x-a)|x\rangle\langle x|\mathrm{d}x,$$

from which we obtain:

$$\hat{U}(a)\hat{X}\hat{U}^{\dagger}(a) = \exp\left(-\frac{i\hat{P}}{\hbar}a\right)\hat{X}\exp\left(\frac{i\hat{P}}{\hbar}a\right) = \hat{X} - a\hat{1}$$

Exercise 2.1 — Commutation relation between momentum and position. Show that the above relation implies $[\hat{X},\hat{P}] = i\hbar$. ∎

2.1.2 The Hamiltonian as the Generator of Time Evolution

Here we follow a similar logic as above, and consider how the state of a system evolves in time, changing from $|\psi(t_0)\rangle$ to $|\psi(t_1)\rangle = |\psi(t_0+t)\rangle$. We modify the notation slightly so

$$|\psi(t)\rangle = \hat{U}(t,t_0)|\psi(t_0)\rangle.$$

Then from equation 2.1, we find that

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{U}(t,t_0) = -\frac{i\hat{H}}{\hbar}\hat{U}(t,t_0), \tag{2.2}$$

where $\hat{H}$ is Hermitian. We find that equation 2.2 leads to

$$\hat{U}(t,t_0) = \exp\left(-i(t-t_0)\hat{H}/\hbar\right).$$

Note that the above equations hold only for when the Hamiltonian is not changing in time. We will often be faced with system where the Hamiltonian itself is time dependent, *i.e.*, $\hat{H}(t)$. Then as we will study later, the differential equation for $\hat{U}$ would still hold, but the solution would need to be modified.

2.2 Time evolution and pictures

2.2.1 Equations of Motion

The quantum states of a system are transformed by time evolution – this is referred to as dynamics. In cases where this evolution is deterministic and **closed**, *i.e.*, we know the dynamics completely and the components of the system are not interacting with external degrees of freedom[1], then this evolution is described by the unitary operator $\hat{U}(t,t_0)$, such that

$$|\psi(t)\rangle = \hat{U}(t,t_0)|\psi(t_0)\rangle$$

We can take unitary time evolution to be a fundamental postulate of quantum mechanics. We call *generator* of the this unitary operator the Hamiltonian. We obtain Schrödinger's equation by taking the time derivative of the $|\psi(t)\rangle$ with respect to time. To see this, we remind ourselves that saying $\hat{U}(t,t_0)$ is generated by a (time-dependent) Hamiltonian is simply another way of stating:

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{U}(t,t_0) = -\frac{i}{\hbar}\hat{H}(t)\hat{U}(t,t_0).$$

Note the slight generalization over the discussion in the last section[2] where we have taken $\hat{H}(t)$ to be itself time-dependent – so the Hamiltonian at a certain time t generates the time evolution at that instance at time t. This generalization means that our original $\exp(\cdots)$ solution no longer holds and we will study the more general solutions to the above differential equation in chapter **??**. For now we see that the above equation implies **Schrödinger's equation**:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = \hat{H}(t)|\psi(t)\rangle$$

Time evolution of a density matrix is given by

$$\hat{\rho}(t) = \hat{U}(t,t_0)\hat{\rho}(t_0)\hat{U}^\dagger(t,t_0),$$

which leads to **Schrödinger's equation for density matrices**:

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\rho}(t) = -\frac{i}{\hbar}[\hat{H}(t),\hat{\rho}(t)] \tag{2.3}$$

When connecting the density matrix to experiment, we saw in chapter 1 that the only connection to experiment comes from expressions of the form

$$O(t) = \mathrm{Tr}[\hat{\rho}(t)\hat{O}]$$

where $\hat{O}$ is an observable, and $\hat{\rho}$ is the state. This means that a completely equivalent theory can be obtained by assuming that $\hat{\rho}$ is constant in time, and the operator $\hat{O}(t)$ has the time dependence needed to produce the same $O(t)$. We can easily obtain this equivalent representation by using the cyclic property of trace:

$$O(t) = \mathrm{Tr}[\rho(t)\hat{O}] = \mathrm{Tr}[\hat{U}(t,t_0)\hat{\rho}(t_0)\hat{U}^\dagger(t,t_0)\hat{O}] = \mathrm{Tr}[\hat{\rho}(t_0)\hat{U}^\dagger(t,t_0)\hat{O}\hat{U}(t,t_0)].$$

[1] Later we will consider 1) time evolution that is classically stochastic, for example due to fluctuations in parameters, and 2) open quantum systems where we study the time evolution of a smaller subsystem interacting with or within a larger quantum system.

[2] For dynamics that have time-dependent Hamiltonians, the postulate $\hat{U}(s_1+s_2) = \hat{U}(s_1)\hat{U}(s_2)$ is replaced by $\hat{U}(t_1,t_0) = \hat{U}(t_1,t_m)\hat{U}(t_m,t_0)$.

In this way, we find an equivalent representation of the theory where the time evolution is completely in the observables, so $\hat{O}_{\mathrm{H}}(t) = \hat{U}^{\dagger}(t,t_0)\hat{O}\hat{U}(t,t_0)$.

Taking the time derivative of $\hat{O}_{\mathrm{H}}(t)$, we obtain **Heisenberg's equation**[3]:

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{O}_{\mathrm{H}}(t) = \frac{i}{\hbar}[\hat{H}(t),\hat{O}_{\mathrm{H}}(t)] \tag{2.4}$$

This representation is called the **Heisenberg picture**, in contrast to the **Schrödinger picture**, which we considered by default. These pictures form only two of a family of many different possible pictures or frames, all of which are related to one other by unitary transformations as we will see below.

2.2.2 Generalizing pictures

We can further generalize the discussion in the previous section to go beyond the Schrödinger and Heisenberg pictures. We saw above that the expected value of a time-independent observable at time t is given by

$$O(t) = \mathrm{Tr}[\hat{U}(t,t_0)\hat{\rho}(t_0)\hat{U}^{\dagger}(t,t_0)\hat{O}].$$

Let us now write the time evolution operator as

$$\hat{U}(t,t_0) = \hat{U}_{\mathrm{T}}(t,t_0)\hat{U}_{\mathrm{I}}(t,t_0). \tag{2.5}$$

This allows us to write

$$O(t) = \mathrm{Tr}[\hat{U}_{\mathrm{I}}(t,t_0)\hat{\rho}(t_0)\hat{U}_{\mathrm{I}}^{\dagger}(t,t_0)\hat{U}_{\mathrm{T}}^{\dagger}(t,t_0)\hat{O}\hat{U}_{\mathrm{T}}(t,t_0)].$$

This expression shows that we can write our equations in a picture where operators evolve according a unitary operator $\hat{U}_{\mathrm{T}}$ (generated by $\hat{H}_{\mathrm{T}}(t)$), and our states evolve by an operator $\hat{U}_{\mathrm{I}}$ (generated by $\hat{H}_{\mathrm{I}}(t)$).

In other words, we can choose a picture by choosing whatever $\hat{H}_{\mathrm{T}}(t)$ we find convenient, and then calculating $\hat{H}_{\mathrm{I}}(t)$ so that eqn. 2.5 is satisfied. The Schrödinger and Heisenberg pictures are now just special cases, obtained for $\hat{H}_{\mathrm{T}} = \hat{1}$ and $\hat{H}_{\mathrm{T}} = \hat{H}$ respectively. Other pictures can be found by choosing other $\hat{H}_{\mathrm{T}}(t)$.

Theorem 2.1 Given the decomposition of the unitary time evolution operator $\hat{U}(t,t_0) = \hat{U}_{\mathrm{T}}(t,t_0)\hat{U}_{\mathrm{I}}(t,t_0)$, and choosing the generator for the transformation of the operator to be $\hat{H}_{\mathrm{T}}(t)$, we find that generator of the time-evolution for the state $\hat{H}_{\mathrm{I}}(t)$ is related the original and the transformation Hamiltonian as

$$\hat{H}(t) = \hat{H}_{\mathrm{T}}(t) + \hat{U}_{\mathrm{T}}(t,t_0)\hat{H}_{\mathrm{I}}(t)\hat{U}_{\mathrm{T}}^{\dagger}(t,t_0). \tag{2.6}$$

Exercise 2.2 Prove the above expression. ■

■ **Example 2.1 — Interaction picture.** we often split our Hamiltonian into two parts:

$$\hat{H}(t) = \hat{H}_0(t) + \hat{V}(t)$$

[3] Notice that here we have assumed that the operator in the Schrödinger picture has no time-dependence. If it were to have time dependence, an additional term $\hat{U}^{\dagger}(t,t_0)(d\hat{O}/dt)\hat{U}^{\dagger}(t,t_0)$ which is sometimes written as $\partial\hat{O}_{\mathrm{H}}/\partial t$ would need to be added on the right side.

The first part may describe and unperturbed part of the system where the dynamics are well understood, or have already been solved. The second part denotes interactions that have been now imposed on the system.

The idea behind the **Interaction Picture** is to move the well-understood part of the dynamics onto the operators, such that anytime evolution of the state is only generated by the interaction part.

As such we choose:

$$\hat{H}_{\mathrm{T}}(t) = \hat{H}_0(t).$$

We find then that

$$\hat{H}_{\mathrm{I}}(t) = \hat{U}_{\mathrm{T}}^{\dagger}(t,t_0)\hat{V}(t)\hat{U}_{\mathrm{T}}(t,t_0).$$

■

2.2.3 Why Move to a Different Frame?

We saw in the previous section that the unitary time evolution operator may be applied to either the states or the observables to obtain the Schrödinger and Heisenberg pictures respectively. These pictures or frames are equivalent in the sense that they predict the same physics. Moreoever, these and other equivalent representations can be used for simplifying the interpretation and solutions of the equations of motion in quantum mechanics. The question arises: *why choose one picture over another?* A particularly beneficial feature of moving to a different frame is that it may make it significantly easier to unmask dynamics that are of primary interest.

A quintessential example of this advantage emerges when we try to solve equations for a driven atom or a two-level system. Take, for instance, a system with a ground state $|g\rangle$ and an excited state $|e\rangle$. These states might have energy levels with a separation of 2 eV or roughly 500 THz ($\omega_{eg} \sim 10^{15}$ 1/s). By applying an electromagnetic field (*i.e.*light) at close to this transition frequency, we induce terms such as $\Omega|g\rangle\langle e| + \text{h.c.}$ in the Hamiltonian. The strength of this interaction, called the Rabi frequency Ω might be significantly smaller than the transition frequency. For example, a typical value may be on the order of 1 MHz ($\Omega \sim 10^7$ 1/s) – more than 8 orders of magnitude smaller.

This disparity causes complications when we attempt to solve the equations numerically in a naive way by just plugging the equations into numerical integrator:

1. The necessary time-step, δt, required to solve the equation would need to be considerably smaller than $1/\omega_o$, for example around $\delta t \sim 10^{-17}$ seconds. Then, roughly 10 billion steps would be needed to accurately capture the dynamics induced by the term proportional to Ω.
2. In the resulting equations of motion, we would be combining terms proportional to the various terms in the Hamiltonian, *i.e.*those proportional to ω_o and Ω. On a digital computer, adding and subtracting small and large numbers often leads to difficulties. If the computer's precision in representing these values is inadequate, the cumulative effects of this imprecision over multiple steps will lead to unreliable results.

Moving into a rotating frame on the other hand, would allow us to obtain accurate results with only $10 - 100$ steps of numerical integration – an 8 to 9 order of magnitude

improvement over the naive approach. It is clear that even if we are only interested in solving quantum mechanical equations numerically, finding the correct picture is an essential step that must be taken with care.

In the following sections, we will study techniques centered around moving between different frames and representations and obtain some practice. We will see that moving to a new frame often offers an elegant solution to the challenges outlined above. In addition to helping us find reliable computational solutions to the equations in question, moving to a new frame can also improve our understanding of the dynamics, remove time-dependencies that complicate analytic analysis, and allow us to make approximations in a controlled fashion.

2.2.4 A Derivation of the Interaction Picture for Time-independent $\hat{H}_0$ in terms of States and Coefficients

In the section 2.2.2, a very general description of pictures beyond Schrödinger and Heisenberg was provided. For clarity, we consider below the derivation of the **Interaction Picture** (given as example 2.1) that is somewhat more explicitly performed at the level of states and coefficients.

Let us begin by considering a system governed by a time-independent Hamiltonian, $\hat{H}_0$. This Hamiltonian, responsible for the system's dynamics, can be diagonalized to yield its eigenstates and eigenenergies $\{|\phi_k\rangle, \hbar\omega_k\}$. As these eigenstates form a complete basis, any quantum state at a given time t can be expressed as:

$$|\psi(t)\rangle_S = \sum_k d_k(t)|\phi_k\rangle,$$

where $d_k(t)$ are time-dependent coefficients and the subscript S indicates that this state is described in the Schrödinger picture.

We can derive the time evolution of $d_k(t)$ by utilizing the Schrödinger equation:

$$\frac{d}{dt}|\psi(t)\rangle_S = -\frac{i}{\hbar}\hat{H}_0|\psi(t)\rangle_S.$$

Taking the inner product of both sides with $\langle\phi_k|$, we find the set of differential equations:

$$\frac{d}{dt}d_k(t) = -i\omega_k d_k(t) \Rightarrow d_k(t) = c_k e^{-i\omega_k t}.$$

This time evolution of the coefficients is captured by an operator $\hat{U}_0(t)$ such that $|\psi(t)\rangle_S = \hat{U}_0(t)|\psi(0)\rangle_S$, where

$$\hat{U}_0(t) \equiv \sum_k |\phi_k\rangle\langle\phi_k|e^{-i\omega_k t} = \exp\left(-i\hat{H}_0 t/\hbar\right).$$

However, in many cases, our Hamiltonian $\hat{H}$ is a combination of a primary part, $\hat{H}_0$, and an additional, possibly time-dependent, interaction part $\hat{V}(t)$. Here, $\hat{H}_0$ is diagonal in its own eigenbasis, making it easier to work with, while $\hat{V}(t)$ may not be.

One common situation is when the dynamics of our quantum system are influenced by some external interaction, represented by $\hat{V}(t)$. In such cases, we're interested in understanding the additional dynamics induced by this interaction term.

To handle this, it's convenient to describe our system in the eigenbasis of $\hat{H}_0$, transforming the Schrödinger picture to what is known as the Dirac or interaction picture. We relate the

state vectors in the interaction and Schrödinger pictures, $|\psi(t)\rangle_{\mathrm{I}}$ and $|\psi(t)\rangle_{\mathrm{S}}$, to each other by the rotation $\hat{U}_0(t)$ in Hilbert space:

$$|\psi(t)\rangle_{\mathrm{S}} = \sum_k c_k(t) e^{-i\omega_k t}|\phi_k\rangle = \hat{U}_0(t)|\psi(t)\rangle_{\mathrm{I}} \quad \Longleftrightarrow \quad |\psi(t)\rangle_{\mathrm{I}} = \sum_k c_k(t)|\phi_k\rangle.$$

When our Hamiltonian is only $\hat{H} = \hat{H}_0$, we saw above that the amplitude for $|\phi_k\rangle$ in the Schrödinger picture evolve as $d_k(t) = c_k e^{-i\omega_k t}$. The $c_k(t)$ in the equations above are constants and all of the time evolution of $|\psi(t)\rangle_{\mathrm{I}}$ has been absorbed into its definition, so that it too remains constant in time.

For a Hamiltonian that includes an interaction term $\hat{H} = \hat{H}_0 + \hat{V}(t)$, $c_k(t)$ are time-dependent coefficients and their time-dependence is due solely to the interaction term $\hat{V}(t)$.

Theorem 2.2 — Relation between Schrödinger Picture and Interaction Picture. The transformation between the Schrödinger picture and the interaction picture is mediated by the operator $\hat{U}_0(t) = \exp(-iH_0 t/\hbar)$. This relationship can be summarized as follows:

$$\begin{aligned} \hat{H} = \hat{H}_0 + \hat{V} \quad &\Longleftrightarrow \quad \hat{H}_{\mathrm{I}}(t) = \hat{U}_0^\dagger(t)\hat{V}\hat{U}_0(t), \\ |\psi(t)\rangle_{\mathrm{S}} = \sum_k d_k(t)|\phi_k\rangle \quad &\Longleftrightarrow \quad |\psi(t)\rangle_{\mathrm{I}} = \sum_k c_k(t)|\phi_k\rangle, \\ \hat{O}_{\mathrm{S}} \quad &\Longleftrightarrow \quad \hat{O}_{\mathrm{I}} = \hat{U}_0^\dagger(t)\hat{O}_{\mathrm{S}}\hat{U}_0(t), \end{aligned}$$

where $\hat{O}_{\mathrm{S}}$ and $\hat{O}_{\mathrm{I}}$ denote any quantum operator in the Schrödinger and interaction pictures, respectively, and we have defined

$$|\psi(t)\rangle_{\mathrm{S}} = \hat{U}_0(t)|\psi(t)\rangle_{\mathrm{I}}, \;\; d_k(t) = c_k(t)e^{-i\omega_k t}.$$

The interaction picture is particularly helpful when the interaction part of the Hamiltonian $\hat{V}(t)$ is not diagonal in the eigenbasis of $\hat{H}_0$. In such situations, this picture simplifies the dynamics by isolating the effect of the interaction term.

■ **Example 2.2 — Two-level System in the Interaction Picture (coefficients).** Consider a two-level quantum system with energies $\hbar\omega_0$ and 0 respectively. The system's time-independent Hamiltonian, $\hat{H}_0$, is given by:

$$\hat{H}_0 = \begin{pmatrix} \hbar\omega_0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Let's suppose that this system is driven by a time-dependent field with strength $\Omega(t)$. The interaction Hamiltonian, $\hat{V}(t)$, can then be expressed as:

$$\hat{V}(t) = \frac{\hbar}{2}\begin{pmatrix} 0 & \Omega(t) \\ \Omega(t) & 0 \end{pmatrix}.$$

The state of the system in the Schrödinger picture, $|\psi(t)\rangle_{\mathrm{S}}$, can be written as:

$$|\psi(t)\rangle_{\mathrm{S}} = d_0(t)|\phi_0\rangle + d_1(t)|\phi_1\rangle,$$

where $d_0(t)$ and $d_1(t)$ are time-dependent coefficients.

The time evolution of these coefficients in the Schrödinger picture is governed by the Schrödinger equation:

$$i\hbar\frac{d}{dt}d_0(t) = \hbar\omega_0 d_0(t) + \frac{\hbar\Omega(t)}{2}d_1(t),$$
$$i\hbar\frac{d}{dt}d_1(t) = \frac{\hbar\Omega(t)}{2}d_0(t).$$

In the absence of the interaction term $\hat{V}(t)$, the differential equations governing the evolution of $d_0(t)$ and $d_1(t)$ simplify to:

$$i\hbar\frac{d}{dt}d_0(t) = \hbar\omega_0 d_0(t),$$
$$i\hbar\frac{d}{dt}d_1(t) = 0.$$

These equations can be solved easily:

$$d_0(t) = d_0(0)e^{-i\omega_0 t}$$
$$d_1(t) = d_1(0).$$

Now, let's transform this system into the interaction picture. The state in this picture, $|\psi(t)\rangle_{\mathrm{I}}$, can be written as:

$$|\psi(t)\rangle_{\mathrm{I}} = c_0(t)|\phi_0\rangle + c_1(t)|\phi_1\rangle,$$

where $c_0(t)$ and $c_1(t)$ are time-dependent coefficients that represent the probability amplitudes of the system being in the lower and upper energy level respectively. They are related to the Schrödinger picture coefficients by

$$d_0(t) = c_0(0)e^{-i\omega_0 t}$$
$$d_1(t) = c_1(0).$$

The time evolution of these coefficients in the interaction picture is governed by the Schrödinger equation:

$$i\hbar\frac{dc_0(t)}{dt} = \frac{\hbar\Omega(t)}{2}e^{i\omega_0 t}c_1(t),$$
$$i\hbar\frac{dc_1(t)}{dt} = \frac{\hbar\Omega(t)}{2}e^{-i\omega_0 t}c_0(t).$$

■

■ **Example 2.3 — Two-level System in the Interaction Picture (vectors and operators).** Consider a two-level system governed by a time-independent Hamiltonian, $\hat{H}_0 = \frac{\hbar\omega_0}{2}(\hat{\sigma}_z + \hat{1})$, and an interaction Hamiltonian $\hat{V}(t) = \frac{\hbar\Omega(t)}{2}\hat{\sigma}_x$, where $\Omega(t)$ is a time-dependent driving field.

The Schrödinger picture wavefunction is $|\psi(t)\rangle_S = d_0(t)|\phi_0\rangle + d_1(t)|\phi_1\rangle = \hat{U}_0(t)|\psi(0)\rangle_S$, with $\hat{U}_0(t) = \exp(-i\hat{H}_0 t/\hbar)$. In the absence of interaction, its coefficients evolve as $d_0(t) = d_0(0)e^{-i\omega_0 t}$ and $d_1(t) = d_1(0)$.

We then transfer to the interaction picture, where the wavefunction is $|\psi(t)\rangle_S = \hat{U}_0(t)|\psi(0)\rangle_I$. The interaction Hamiltonian in this picture becomes $\hat{V}_I(t) = \hat{U}_0^\dagger(t)\hat{V}(t)\hat{U}_0(t)$.

In the Schrödinger picture, the additional Hamiltonian representing the drive is written in terms of vectors as $\hat{V}(t) = \frac{\hbar\Omega(t)}{2}(|\phi_0\rangle\langle\phi_1| + |\phi_1\rangle\langle\phi_0|)$. Now, transitioning into the interaction picture, we have to compute the interaction Hamiltonian as $\hat{V}_I(t) = \hat{U}_0^\dagger(t)\hat{V}(t)\hat{U}_0(t)$. This gives us

$$\hat{V}_I(t) = \frac{\hbar\Omega(t)}{2}(e^{i\omega_0 t}|\phi_0\rangle\langle\phi_1| + e^{-i\omega_0 t}|\phi_1\rangle\langle\phi_0|).$$

This is the interaction picture Hamiltonian which gives us the same differential equations derived in example 2.2. ∎

2.2.5 Pictures related by unitary transformations

The Schrödinger and Interaction pictures as defined above are related to one another by the unitary transformation $\hat{U}_0(t)$. Other unitary transformation can also be used and often valuable for the computational and conceptual simplifications that they provide. It is therefore helpful to to derive the equations of motion and definition of states and operators for general unitary transformations.

Theorem 2.3 — Relation between Schrödinger Picture and a Frame Defined by a General Unitary Transformation. The transformation between the Schrödinger picture and a frame defined by a general unitary transformation is provided by the operator $\hat{U}_\mathrm{T}(t)$. This connection is captured in the following equations:

$$\begin{aligned}
\hat{H}(t) &\iff \hat{H}_\mathrm{I}(t) = \hat{U}_\mathrm{T}^\dagger(t)\hat{H}(t)\hat{U}_\mathrm{T}(t) - i\hbar\hat{U}_\mathrm{T}^\dagger(t)\partial_t\hat{U}_\mathrm{T}(t), \\
|\psi(t)\rangle_\mathrm{S} &\iff |\psi(t)\rangle_\mathrm{T} = \hat{U}_\mathrm{T}^\dagger(t)|\psi(t)\rangle_\mathrm{S}, \\
\hat{O}_\mathrm{S} &\iff \hat{O}_\mathrm{T} = \hat{U}_\mathrm{T}^\dagger(t)\hat{O}_\mathrm{S}\hat{U}_\mathrm{T}(t),
\end{aligned}$$

where $\hat{O}_\mathrm{S}$ and $\hat{O}_\mathrm{T}$ represent any quantum operator in the Schrödinger and transformed pictures, respectively.

Proof. The proof follows directly from discussion in 2.2.2, but we will provide an alternate proof here that starts with the Schrödinger equation.

We start with the Schrödinger equation for state evolving under the influence of a possibly time-dependent Hamiltonian:

$$i\hbar\partial_t|\psi(t)\rangle_\mathrm{S} = \hat{H}(t)|\psi(t)\rangle_\mathrm{S}.$$

We denote our general unitary transformation, $\hat{U}_\mathrm{T}(t)$ (T stands for transformation). This transformation defines a new frame of reference where hopefully the physics, which is equivalent to that described in any other frame, is a bit easier to analyze. Acting on a state in the new basis, the transformation maps every vector in the new frame to its corresponding Schrödinger picture state:

$$|\psi(t)\rangle_\mathrm{S} \equiv \hat{U}_\mathrm{T}(t)|\psi(t)\rangle_\mathrm{T}.$$

Using this definition, we can see what the original equation implies:

$$\begin{aligned}
i\hbar\partial_t|\psi(t)\rangle_\mathrm{S} &= \hat{H}(t)|\psi(t)\rangle_\mathrm{S} \\
\Rightarrow i\hbar\partial_t\left(\hat{U}_\mathrm{T}(t)|\psi(t)\rangle_\mathrm{T}\right) &= \hat{H}(t)\left(\hat{U}_\mathrm{T}(t)|\psi(t)\rangle_\mathrm{T}\right) \\
\Rightarrow i\hbar\left(\left[\partial_t\hat{U}_\mathrm{T}(t)\right]|\psi(t)\rangle_\mathrm{T} + \hat{U}_\mathrm{T}(t)\partial_t|\psi(t)\rangle_\mathrm{T}\right) &= \hat{H}(t)\left(\hat{U}_\mathrm{T}(t)|\psi(t)\rangle_\mathrm{T}\right)
\end{aligned}$$

	Schrödinger Picture	Interaction Picture	Heisenberg Picture
Hamiltonian	$\hat{H} = \hat{H}_0 + \hat{V}(t)$	$\hat{H}_I(t) = \hat{V}_I(t)$ $\hat{V}_I(t) = \hat{U}_0^\dagger(t)\hat{V}(t)\hat{U}_0(t)$	$\hat{H}_H = 0$
States	$\hat{\rho}(t) = \hat{U}(t)\hat{\rho}(0)\hat{U}^\dagger(t)$	$\hat{\rho}_I(t) = \hat{U}_I(t)\hat{\rho}_I(0)\hat{U}_I^\dagger(t)$	$\hat{\rho}(t) = \hat{\rho}(0)$
Observables	$\hat{O}(t)$	$\hat{O}_I(t) = \hat{U}_0^\dagger(t)\hat{O}(t)\hat{U}_0(t)$	$\hat{O}_H(t) = \hat{U}^\dagger(t)\hat{O}(t)\hat{U}(t)$ $\hat{O}_H(t) = \hat{U}_I^\dagger(t)\hat{O}_I(t)\hat{U}_I(t)$
Propagators	$i\hbar\partial_t\hat{U}(t) = \hat{H}(t)\hat{U}(t)$	$i\hbar\partial_t\hat{U}_I(t) = \hat{H}_I(t)\hat{U}_I(t)$ $\hat{U}(t) = \hat{U}_0(t)\hat{U}_I(t)$	$\hat{U}_H(t) = \hat{1}$
$\hat{U}_\mathrm{T}(t)$	$\hat{1}$	$\hat{U}_0(t) \equiv \exp\left(-\frac{i\hat{H}_0 t}{\hbar}\right)$	$\hat{U}(t)$

Table 2.1: Relations between Schödinger, Interaction, and Heisenberg pictures.

leading to

$$i\hbar\partial_t|\psi(t)\rangle_\mathrm{T} = \underbrace{\left(\hat{U}_\mathrm{T}^\dagger(t)\hat{H}(t)\hat{U}_\mathrm{T}(t) - i\hbar\hat{U}_\mathrm{T}^\dagger(t)\left[\partial_t\hat{U}_\mathrm{T}(t)\right]\right)}_{\hat{H}_\mathrm{I}(t)}|\psi(t)\rangle_\mathrm{T}.$$

■

Note that by setting $\hat{H}(t) = \hat{H}_0 + \hat{V}$ and choosing the transformation $\hat{U}_\mathrm{T}(t) = \hat{U}_0(t) \equiv \exp\left(-\frac{i\hat{H}_0 t}{\hbar}\right)$, we arrive at the interaction picture as discussed in section 2.2.4. This is due to the fact that $-i\hbar\hat{U}_\mathrm{T}^\dagger(t)\partial_t\hat{U}_\mathrm{T}(t) = -\hat{H}_0$, while $\hat{U}_\mathrm{T}^\dagger(t)\hat{H}(t)\hat{U}_\mathrm{T}(t) = \hat{H}_0 + \hat{U}_\mathrm{T}^\dagger(t)\hat{V}\hat{U}_\mathrm{T}(t)$.

We arrive at the Heisenberg picture by setting $\hat{U}_\mathrm{T}(t)$ to be the unitary evolution operator from the Schödinger picture, generated by the full Hamlitonian $\hat{H}$.

We summarize the resulting relationship between the three pictures in table 2.1.

2.2.6 Examples

In the following example, we examine the system that we considered in the introductory paragraph of this chapter. A two level system is driven by an external field oscillating at the frequency close to its transition.

■ **Example 2.4 — Periodic Driving of a Two-Level System.** Consider a two-level system with states $|0\rangle$ and $|1\rangle$. For this system, the Hamiltonian consists of two parts: a time-independent part, $\hat{H}_0$, and a time-dependent driving term, $\hat{V}(t)$.

1. **The Two-Level System:**

   $$\hat{H}_0 = \hbar\omega_1|1\rangle\langle 1|$$

   Where:
   - ω_1 is the energy difference between the two states.
2. **The Driving Term:** The system is subjected to an external drive given by:

   $$\hat{V}(t) = \hbar\Omega[|1\rangle\langle 0| + |0\rangle\langle 1|]\cos(\omega_d t)$$

   Where:
   - Ω is the Rabi frequency which determines the strength of the drive.
   - ω_d is the driving frequency.

3. **The Total Time-Dependent Hamiltonian:**

$$\hat{H}(t) = \hat{H}_0 + \hat{V}(t)$$

4. **Interaction Picture Hamiltonian:** By moving to the interaction picture, the Hamiltonian becomes:

$$\hat{H}_I(t) = \hat{U}_0^\dagger(t)\hat{V}(t)\hat{U}_0(t)$$

With:

$$\hat{U}_0(t) = \exp[-i\hat{H}_0 t/\hbar] = |0\rangle\langle 0| + e^{-i\omega_1 t}|1\rangle\langle 1|$$

Though this is a perfectly valid frame to work in, we will find it even *more* convenient to move into a slightly different frame, one that oscillates at the drive frequency.

5. **Moving to the Frame of the Drive:** We define a transformation

$$\hat{U}_T(t) = \exp[-i\omega_d t|1\rangle\langle 1|] = |0\rangle\langle 0| + e^{-i\omega_d t}|1\rangle\langle 1|$$

that rotates with the drive. The new Hamiltonian $\hat{H}_\mathrm{I} = \hat{U}_\mathrm{T}^\dagger(t)\hat{H}\hat{U}_\mathrm{T}(t) - \hbar\omega_d|1\rangle\langle 1|$ results in:

$$\hat{H}_\mathrm{I} = -\hbar\delta|1\rangle\langle 1| + \hbar\Omega[e^{i\omega_d t}|1\rangle\langle 0| + |0\rangle\langle 1|e^{-i\omega_d t}]\cos(\omega_d t)$$

Where $\delta = \omega_d - \omega_1$ is the detuning between the drive and the system.

6. **Making the Rotating Wave Approximation (RWA):** In the RWA, rapidly oscillating terms are ignored. Observe the terms in $\hat{H}_\mathrm{T}$ that oscillate rapidly.

$$\begin{aligned}\hat{H}_I &= -\hbar\delta|1\rangle\langle 1| + \frac{\hbar\Omega}{2}\left[|1\rangle\langle 0| + |0\rangle\langle 1|\right] + \frac{\hbar\Omega}{2}\left[e^{2i\omega_d t}|1\rangle\langle 0| + e^{-2i\omega_d t}|0\rangle\langle 1|\right] \\ &= \hat{H}_{TI} + \hat{H}_{TD}\end{aligned}$$

where the time-independent ($\hat{H}_{TI}$) and time-dependent ($\hat{H}_{TD}$) parts of the transformed Hamiltonian are defined as:

$$\begin{aligned}\hat{H}_{TI} &\equiv -\hbar\delta|1\rangle\langle 1| + \frac{\hbar\Omega}{2}\left[|1\rangle\langle 0| + |0\rangle\langle 1|\right] \\ \hat{H}_{TD} &\equiv \frac{\hbar\Omega}{2}\left[e^{2i\omega_d t}|1\rangle\langle 0| + e^{-2i\omega_d t}|0\rangle\langle 1|\right]\end{aligned}$$

We make the rotating wave approximation by noting that when $\Omega \ll \omega_d$, $\hat{H}_{TD}$ only minutely affects the system's evolution since the rapidly oscillating part averages out. This gives us

$$\hat{H}_I \approx \hat{H}_{TI} \equiv -\hbar\delta|1\rangle\langle 1| + \frac{\hbar\Omega}{2}\left[|1\rangle\langle 0| + |0\rangle\langle 1|\right]$$

We see that by moving into the frame of the drive, we can eliminate rapidly oscillating terms and obtain a time-independent Hamiltonian – a significant simplification of our equations of motion. ■

The two following examples require some familiarity with the Harmonic Oscillator. You may wish to come back to these two after Chapter **??**.

■ **Example 2.5 — Linear Driving of a Harmonic Oscillator.** Consider a spring-block system that is subject to an external driving force $F(t)$. In quantum mechanics, our harmonic oscillator will have a Hamiltonian, $\hat{H}_0$, and an external driving term, $\hat{V}$.

1. **The Free Harmonic Oscillator:**

$$\hat{H}_0 = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega_o^2\hat{x}^2$$

Where:
- $\hat{p}$ is the momentum operator.
- m is the mass of the particle.
- ω_o is the natural frequency of the oscillator.
- $\hat{x}$ is the position operator.

2. **The Driving Term:** The system is driven by an external force which is periodic in time:

$$\hat{V}(t) = -F(t)\cdot\hat{x} = 2F_\omega\cos(\omega_d t)\hat{x}$$

Where we chose $F(t) = -2F_\omega\cos(\omega_d t)$ with
- F_ω is the amplitude of the external force.
- ω_d is the driving frequency.
- The factor of 2 is for normalization convenience.

3. **The Total Time-Dependent Hamiltonian:**

$$\hat{H}(t) = \hat{H}_0 + \hat{V}(t)$$

4. **Expression in terms of Creation and Annihilation Operators:**
As shown in Chapter **??**, it is often convenient to express the position and momentum operators in terms of creation and annihilation operators $\hat{a}^\dagger$ and $\hat{a}$ defined as:

$$\hat{x} = x_{\text{zpf}}(\hat{a}^\dagger + \hat{a})$$
$$\hat{p} = ip_{\text{zpf}}(\hat{a}^\dagger - \hat{a})$$

Where:
- x_{zpf} is the standard deviation of the oscillator's position when in its ground state; $x_{\text{zpf}} = \sqrt{\hbar/2m\omega_o}$.
- p_{zpf} is the standard deviation of the oscillator's momentum when in its ground state; $p_{\text{zpf}} = \sqrt{\hbar m\omega_o/2}$.

Then for the quantum harmonic oscillator, we find[4]

$$\hat{H}_0 = \hbar\omega_o\hat{a}^\dagger\hat{a}$$
$$\hat{V}(t) = 2F_\omega\cos(\omega_d t)x_{\text{zpf}}(\hat{a}^\dagger + \hat{a})$$

Using the identity $\cos(\omega_d t) = \frac{1}{2}\left(e^{i\omega_d t} + e^{-i\omega_d t}\right)$, we express the driving term as:

$$\hat{V}(t) = F_\omega x_{\text{zpf}}(\hat{a}^\dagger + \hat{a})(e^{i\omega_d t} + e^{-i\omega_d t})$$

5. **Going into the frame of the drive field:** Our goal is to attempt to remove the time-dependence in the equations above – this would greatly simplify their analysis. It is generally not possible to find a time-independent description for a system that contains a time-dependent force. But in the case of the harmonically driven oscillator, we can succeed to within an approximation described below. First we move into a rotating frame by choosing the correct transformation:

$$\hat{U}_T(t) = \exp(-i\omega_d\hat{a}^\dagger\hat{a}t).$$

[4] Actually $\hat{H}_0 = \hbar\omega_o(\hat{a}^\dagger\hat{a} + 1/2)$, but we drop the constant offset of $\hbar\omega_o/2$ in the energy for simplicity.

Notice, that this is slightly different than $\hat{U}_0$ which we use to go the interaction picture. Applying this transformation to our Hamiltonian, we find[5]:

$$\hat{H}_{\mathrm{I}}(t) = \hat{U}_{\mathrm{T}}^{\dagger}(t)\hat{H}(t)\hat{U}_{\mathrm{T}}(t) - i\hbar\hat{U}_{\mathrm{T}}^{\dagger}(t)\partial_t\hat{U}_{\mathrm{T}}(t)$$

which gives us

$$\hat{H}_I(t) = \hat{H}_0 - \hbar\omega_d\hat{a}^{\dagger}\hat{a} + F_{\omega}x_{\mathrm{zpf}}(\hat{a}^{\dagger}e^{i\omega_d t} + \hat{a}e^{-i\omega_d t})(e^{i\omega_d t} + e^{-i\omega_d t})$$

.
Expanding the driving term in $\hat{H}_T(t)$, we have:

$$\begin{aligned} &F_{\omega}x_{\mathrm{zpf}}(\hat{a}^{\dagger}e^{i\omega_d t} + \hat{a}e^{-i\omega_d t})(e^{i\omega_d t} + e^{-i\omega_d t}) \\ &\quad = F_{\omega}x_{\mathrm{zpf}}(\hat{a}^{\dagger} + \hat{a}) + F_{\omega}x_{\mathrm{zpf}}(\hat{a}^{\dagger}e^{2i\omega_d t} + \hat{a}e^{-2i\omega_d t}) \end{aligned}$$

Now, splitting $\hat{H}_I(t)$ into time-independent and time-dependent parts, we find:

$$\begin{aligned} \hat{H}_{TI} &= \hat{H}_0 - \hbar\omega_d\hat{a}^{\dagger}\hat{a} + F_{\omega}x_{\mathrm{zpf}}(\hat{a}^{\dagger} + \hat{a}) \\ \hat{H}_{TD}(t) &= F_{\omega}x_{\mathrm{zpf}}(\hat{a}^{\dagger}e^{2i\omega_d t} + \hat{a}e^{2i\omega_d t}) \end{aligned}$$

6. **Making the rotating wave approximation (RWA):** The crux of the RWA lies in observing that certain terms in the Hamiltonian will oscillate rapidly, and under some conditions will effectively average to zero over time and hence can be neglected. Looking at $\hat{H}_{TD}(t)$, we notice the terms are oscillating at twice the driving frequency. On the other hand, the rest of the Hamiltonian in this picture is generating time evolution at frequencies on the order of $|\omega_d - \omega_0|$ and $|F_{\omega}x_{\mathrm{zpf}}|$. If $2\omega_d$ is much faster than $|\omega_d - \omega_0|$ and $|F_{\omega}x_{\mathrm{zpf}}|$, then its effect will be averaged out and we can ignore it. This is called the RWA.
Within the RWA, the Hamiltonian is approximated as:

$$\hat{H}_I \approx \hbar(\omega_o - \omega_d)\hat{a}^{\dagger}\hat{a} + F_{\omega}x_{\mathrm{zpf}}(\hat{a}^{\dagger} + \hat{a})$$

■

■ **Example 2.6 — Parametric Driving of a Harmonic Oscillator.** Consider a parametric oscillator where the oscillation frequency is modulated, for example, by changing the spring constant periodically in time.

1. **The Parametrically Driven Hamiltonian:**

$$\begin{aligned} \hat{H}(t) &= \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega_o^2(1 + \epsilon(t))\hat{x}^2 \\ \epsilon(t) &= \epsilon_0\cos(2\omega_d t) \end{aligned}$$

Here:
- $\epsilon(t)$ represents the time-dependent modulation of the spring constant.
- ϵ_0 is the amplitude of the modulation.
- $2\omega_d$ is the frequency of the parametric drive.

[5] We use the operator identity $\exp(-i\phi\hat{a}^{\dagger}\hat{a})\hat{a}\exp(i\phi\hat{a}^{\dagger}\hat{a}) = e^{i\phi}\hat{a}$ which is discussed in Chapter **??**.

2. **Hamiltonian in terms of Creation and Annihilation Operators:** Using the definitions above for the Hamiltonian and the drive term, and the relation between the position and momentum operators, and the creation and annihilation operators, we find:

$$\begin{aligned}\hat{H}(t) &= \frac{p_{\text{zp}}^2}{2m}(\hat{a}^\dagger - \hat{a})^2 + \frac{1}{2}m\omega_o^2(1+\epsilon(t))x_{\text{zp}}^2(\hat{a}+\hat{a}^\dagger)^2 \\ &= \hbar\omega_o\hat{a}^\dagger\hat{a} + \frac{1}{4}\hbar\omega_o\epsilon(t)(\hat{a}+\hat{a}^\dagger)^2 \\ &= \hbar\omega_o\hat{a}^\dagger\hat{a} + \frac{1}{4}\hbar\omega_o\epsilon_0\cos(2\omega_d t)(\hat{a}+\hat{a}^\dagger)^2 \\ &= \hbar\omega_o\hat{a}^\dagger\hat{a} + \frac{1}{8}\hbar\omega_o\epsilon_0(e^{2i\omega_d t}+e^{-2i\omega_d t})(\hat{a}+\hat{a}^\dagger)^2\end{aligned}$$

   Now we define $\beta \equiv \frac{1}{8}\omega_o\epsilon_0$, and find:

$$\hat{H}(t) = \hbar\omega_o\hat{a}^\dagger\hat{a} + \hbar\beta\left(e^{-2i\omega_d t}\hat{a}^{\dagger 2} + e^{2i\omega_d t}\hat{a}^2\right) + (\text{other terms})$$

   Here:
   - $\beta \equiv \frac{1}{8}\omega_o\epsilon_0$ is a parameter related to the amplitude of the parametric drive.

3. **Transforming to a Rotating Frame of Reference:**
   We use a transformation $\hat{U}_T(t) = \exp\left(-i\omega_d\hat{a}^\dagger\hat{a}t\right)$, which puts us in the frame of reference of the drive:

$$\begin{aligned}\hat{H}_I(t) &= \hat{U}_T^\dagger(t)\hat{H}(t)\hat{U}_T(t) - i\hbar\hat{U}_T^\dagger(t)\frac{d\hat{U}_T(t)}{dt} \\ &= \hbar(\omega_o-\omega_d)\hat{a}^\dagger\hat{a} + \hbar\beta\left(\hat{a}^{\dagger 2}+\hat{a}^2\right) + \hat{U}_T^\dagger(t)(\text{other terms})\hat{U}_T(t)\end{aligned}$$

4. **Simplification to a Time-Independent Hamiltonian within RWA:**
   Note that in the interaction picture, if we drop the "other terms", the Hamiltonian becomes time-dependent. Dropping these other terms may be justified within the rotating wave approximation resulting in:

$$\hat{H}_I \approx \hbar(\omega_o-\omega_d)\hat{a}^\dagger\hat{a} + \hbar\beta(\hat{a}^{\dagger 2}+\hat{a}^2)$$

∎