A Brief Course of Quantum Theory

DONG-SHENG WANG

October 20, 2019

Contents

Introduction

Question 1. *What is 'quanta'?*

Quanta, or 'Quantum' is by no means a good name; however, there is no better one. Max Planck discovered the constant $h = 6.62606957(29) \times 10^{-34} J \cdot s$, which is treated as the element, or 'quanta', of *action*.

What does 'quantum' mean? This is a primary but difficult question. Intuitively, a new constant will lead to new principle of nature, which, unfortunately, is not so clear yet, but implicitly sets the foundation of quantum theory. Here I spell my answer first, then I will explain more later. Quantum theory is a theory that describes the *structure of motion* of objects. The structure of motion is the notion that differs from *motion*, *object*, and *structure of object*. Traditionally in physics and our common sense, motion is the movement in space and time, object is a 'thing' with mass or energy, e.g., a ball, an electron, or the earth, and structure of object is the details of the parts of object, could be geometrical, chemical, etc.

From classical mechanics we know *phase space*. We can view quantum theory as a generalization of it. It is a more powerful, complete version of phase space paradigm. Phase space is based on *conjugate variables*, such as position r and momentum p, and their product is nothing but the action. With a Hamiltonian $H(x, p)$, the Hamilton's equation is

$$\frac{\partial H}{\partial x} = -\dot{p}, \ \frac{\partial H}{\partial p} = \dot{x}. \tag{1}$$

This is, furthermore, equivalent to Lagrangian version by $L := \dot{x}p - H$ and the principle of least action.

In quantum theory, conjugate variables are not numbers anymore. Instead, they are collections of numbers, i.e., they are matrices, or, *operators*. The consequence is that, conjugate operators, also called observable, cannot be measured at the same time since they do not commute. This leads to the *complementarity principle* of Niels Bohr, or the *duality* property.

Why a matrix is needed to describe a 'thing'? Well, the short answer is, the thing is complicated. It cannot be well described by a single number or function. Let's see a most trivial thing: people wear clothes. He/she can wear different ones, he/she can wear the same one in different situations, he/she can choose one depending on

many factors, such as pocket, mood, etc. This complicated thing of wearing clothes, especially for females, has to be described by an exotic thing: an operator.

Question 2. *What is the primary quantum equation?*

In quantum physics, structure of the motion of one object, e.g., an electron, are described as a complex vector $|\psi\rangle$ in *Hilbert space*, and the evolution is represented by one operator acting on the state changing its magnitude and phase. Suppose the dynamics of the particle is driven by a Hamiltonian H, then the quantum equation is

$$i\hbar|\dot{\psi}\rangle = H|\psi\rangle, \tag{2}$$

with $\hbar = h/2\pi$ which appears much more unique than h itself. H is an operator, for which the eigenvalues are the possible energies. If the state $|\psi\rangle$ initially is expressed in the basis of the eigenstates of H, then the evolution of the state will include: the evolution of the coefficients of the eigenstates, which changes the relative populations (i.e. probabilities) and relative dynamical phases of the eigenstates, and the rotation of the eigenstates themselves which generates geometric phases. Let us name the above equation as *state equation*, since it provides the possible states of the particle, the probabilities for the particle to be in the states, and also the relative phases among the states. The state equation is originally discovered by Erwin Schrödinger for a special case of particles without internal degree of freedoms (d.o.f).

Briefly, the state equation provides the structure of motion of an object. As the result, quantum theory is a new kind of description of motion, which is different from other theories, including classical mechanics, wave mechanics etc.

Question 3. *Except $|\psi\rangle$ and H, why there are $\hbar$ and i?*

The obvious reason for i is that it transfers the effects of Hamiltonian into a phase, and the constant $\hbar$ servers as the action element. In the early age of quantum physics, $\hbar$ is more important. However, it is clear now the quantum state $|\psi\rangle$ (and also H) is more important.

There is a limit by $\hbar \to 0$, which shall implies there is no quantum effect anymore. This is the *correspondence principle* of Bohr. The disappearance of $\hbar$ means the ignorance of the structure of motion, e.g., averaging the detailed structure of motion, thus leading to classical results.

Question 4. *What is the magic of quantum coherence?*

By extending from numbers to operators, quantum theory brings us the most important concept: *quantum coherence*. Please do not mix it with other notions of coherence that you know. Quantum coherence is the origin of all these quantum 'magic': superposition, complementarity, uncertainty, entanglement, contextuality, duality, spin, etc.

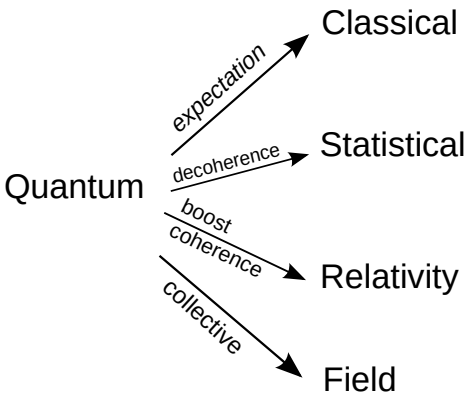

Figure 1: The quantum landscape. Quantum to classical: take average; quantum to statistics: take partial trace; quantum to special relativity: boost coherence, spacetime qubit; quantum to quantum field: take continuous limit.

The most essential consequence of quantum coherence is that it allows quantum theory to unify many theories. Quantum theory unifies classical mechanics, special relativity, statistical physics, electromagnetism, quantum field theory, and more. It reduces to them under certain limits, as shown in Fig. 1. Besides, there are many types of coherence:

- coherence in field or vacuum: revealed by quantum field.
- coherence in matter: usual quantum theory, e.g. atomic theory.
- coherence between matter and light: revealed by quantum optics.
- coherence between matter and vacuum: revealed by special relativity.

Question 5. *What is spin?*

In quantum theory, there is a thing called 'spin'. Spin is an inner d.o.f of a quantum object, or, particle. A macroscopic object many not have spin but no one proves this. Again, 'spin' is a bad name. It does not mean a particle is spinning. In math, there is a group called 'spin group' (also 'pin group'), making this term even more weird. I have to say, I do not know the real physical meaning of spin. It has to be explained by some 'post-quantum' theory. Elementary particles also have other features, such as charge, mass, color, weak charge etc, but spin is the most amazing one.

Spin is said to take integer or half-odd-integer values. It cannot be 1/3, say. However, this is not quite true. The observed value of spin can be any real values. The value of spin s actually determines the Hilbert space dimension $d = 2s + 1$. The dimension d can only be integers by definition (except some effective ones such as for anyons). Spin s can also be zero. A profound theorem from quantum field theory is that particles are classified into two classes: bosons for integer

spins, fermions for half integers. This is the *spin-statistics* theorem. A collection of bosons (fermions) obey Bose-Einstein (Fermion-Dirac) statistics, as generalizations of Boltzmann statistics.

Different spins obey slightly different state equations. Let's see how it describes free particles with spins.

- spin 0: Klein-Gordan equation.
- spin 1/2: Dirac equation.
- spin 1: Maxwell equation without source terms.
- higher spins: we will see these equations.

Question 6. *Is quantum theory self-consistent?*

The short answer is it must be. An amazing thing of quantum theory is that it is extendable, and the extended version is equivalent in a sense to the original one. In other words, it is 'closed' or 'self-consistent'. Here I show you two different kinds of extensions. First, a quantum theory can be further quantized to a higher level. This is because 'structure of motion' can be treated as motion, and then a higher-level structure can be introduced. If you could do so consistently, then you can get more and more complete understanding of the motion of an object. Second, a quantum theory can be 'de-quantized' (or randomized in an ensemble), but can be treated as part of a quantum theory again. This is the connection with statistical physics.

Question 7. *How to interpret quantum theory?*

What is 'interpretation'? In brief, it is different understandings or explanations of roughly the same phenomena (experiments) or mathematics. A theory needs interpretation since the mathematical symbols need meanings. Just like 'gravity', which was just a name and it took centuries for people to accept it, 'quantum' is also just a name. The power of quantum theory has been proven by the tremendous progress during the last century.

There is a subject known as 'interpretation of quantum theory' or 'quantum foundation'. Historically, this originates from confusions about some concepts, the role of quantum theory itself in physics, and also about philosophical implications of quantum theory. These include the concepts of wavefuction, Hilbert space, spin, measurement, etc, and the relation with other theories such as 'classical' ones, and notions of reality, causality, completeness, determinism, etc. Fortunately, nowadays we almost know what quantum theory is, as described above. Quantum theory is a type of *unification* theory, i.e., a more complete description of motion, that can include other theories as special cases. Quantum state, which certainly is physical, describes the structure (or scope) of the motion of objects.

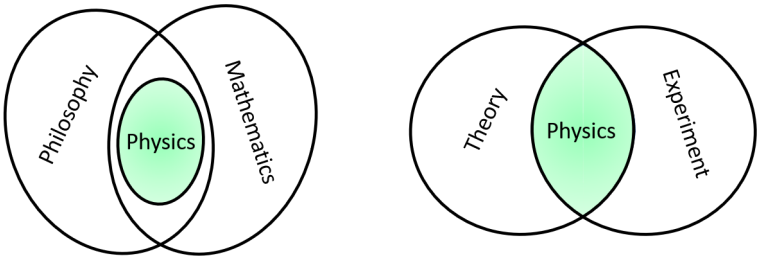

Figure 2: Left: Philosophy and Mathematics as abstraction of Physics. Right: Theory and Experiment overlap for Physics.

Although there are no need of more interpretations, quantum theory is not at all the 'ultimate' theory since there are still so many phenomena to understand, such as spin, gravity, anti-matter, holography, chaos, mind, and things we still do not know of. Certain post-quantum theories shall be able to explain at least one thing beyond the paradigm of quantum theory.

Question 8. *What features specify physics?*

Physics is not about knowledge or truth since lots of knowledge prove to be wrong. Physics is a way to quest nature about the nature (character) of nature. Mathematics is a modeling of nature, it constructs an abstract model which can be used by physicists to quest nature. Metaphysics, or philosophy, is also a modeling of nature but it uses the language of concepts and ideas, instead of symbols and equations, and it tells us what nature should be, and it teaches us how to understand why nature takes the current form. See Fig. 2, left panel.

Physicists are pragmatist: they use anything that they think are helpful, any concepts and models from math and philosophy. They use these tools to pursue the goal: they want to find new phenomena in nature.

Physicists are skepticism: they do not believe that the current theory is the eventual truth, and they always believe that there will be a better theory. They do not believe that they really do not believe in anything.

Mathematics and philosophy should not be pragmatist or skepticism because the following reason: their job is to do modelling, they want a better model than nature, a model that can explain nature, hence they have to believe that there exists such a better model.

However, in practice there is no clear or sharp boundaries between physics and math, or philosophy. Despite this, it is easy to tell them in front of experiments: after all, physics needs to explain experiments of matter in nature.

Question 9. *What are quantum physics?*

Physics is not just about 'theory'. Instead, physics contains both theory and experiment, and even applications. The theory and experiment are about *matter*, such

as atom, light, liquid, solid, earth, and the universe, instead of other things, such as people, animals, etc. However, theory and experiment only overlap to a certain degree. A theory may have application beyond physics in other areas, such as economics, and an experiment may need theory or concepts beyond physics, such as chemistry, biology, or even computer science. See Fig. 2, right panel. So we shall be open minded and treat physics, including quantum physics, as an *open system*, so it can fit into the whole world of science and build various connections.

We shall see what quantum physicists are doing. There are many, here are some:

- atom, molecule: structure of atoms, molecules, and reactions.
- cold atom, laser: cooling atoms, interaction with laser, condensation.
- condensed matter: quantum phase transition, materials.
- high energy: elementary particles and related.
- nano: 1D, 2D, 3D structures, graphenes, semiconductors.
- optics, laser: photons, non-linearity, photonic crystal, chips.
- quantum thermodynamics: open system, heat engine.
- quantum computing: gates, channels, entanglement.
- quantum gravity: duality, black hole, universe.
- quantum chaos: few-body dynamics.
- quantum magnetism: many-body systems of spins, topological phases.
- solid-state: phonons, interaction with light, pressure, etc.
- spintronics: control of spins, magnetism, devices, logic.
- superconductors: high-temperature, junctions.

Quantum theory is becoming more and more complicated since too many things are squeezed into it. This brief lecture notes would not cover all topics; instead, it aims for the opposite: we try to distill the most central concepts in quantum theory, and explain them in the most simple words. Once you grasp the essence, you will find quantum theory is amazing and you will be able to improve it.

D.-S. Wang
wdscultan@gmail.com
Date: October 20, 2019

Chapter 1

Qubits

A quantum bit, or 'qubit', is a two-level quantum system, which is the simplest yet nontrivial quantum systems. In this chapter, we will survey many concepts and theories just for the case of qubit. Although there are indeed some complications when generalizing to other cases, we find the qubit case captures lots of the essence.

1.1 Bit, pbit

A *bit* is an object with two states: 0 or 1. An example is a coin with a head and a tail state. A slight generalization of bit is a probabilistic bit, or *pbit*, which can be at state 0 or 1 with a probability. Namely, a state of a pbit can be written as $p0 + (1-p)1$ or a vector $(p, 1-p)$ for $p \in [0,1]$.

Question 10. *What is the meaning of* p*, as probability?*

This is not a new question, but we need to be careful of its *physical* meaning. It is a sense of ignorance, but there are usually two types.

1. First, it is a statistical weight for a collection of identical bits. Suppose there are N bits and there are pN of them with head state. So the state of the whole collection can be viewed as a pbit.

2. Second, it is a dynamical abstraction for the evolution of a bit. Say, the coin is thrown N times and pN of them with head state. So the state of the whole event can be viewed as a pbit.

These two point of views are equivalent under a certain condition, which is known as the *ergodic* condition: all points in phase space can be equally visited. Here we already encounter the concept of phase space, which we shall study more in Chapter 3. In addition, there is another type that is not common in physics but somewhere else. It is a subjective ignorance, say, due to the lack of knowledge of people. This relates to *Bayesian* inference. However, in physics we do not deal with the lack of

knowledge of people. The thing that interacts with an observed object is a physical device. So p shall be explained by physical terms.

1.2 Qubit

A qubit is a further generalization of a pbit from real vectors to complex vectors (states). We need the *Dirac notation*: $|\psi\rangle$ for a state ψ. As it is complex, we use $\langle\psi|$ for its conjugate. A qubit state can be written as

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \quad \alpha,\beta \in \mathbb{C}, \tag{1.1}$$

for $|0\rangle = (1,0)^t$, $|1\rangle = (0,1)^t$. Furthermore, we can use polar form for α and β and ignore a common phase factor, then

$$|\psi\rangle = a|0\rangle + be^{i\varphi}|1\rangle, \quad a,b,\varphi \in \mathbb{R}. \tag{1.2}$$

Next, we require that the state is *normalized* as it is in a projective Hilbert space (see Chapter 2), namely, $\langle\psi|\psi\rangle = 1$, we have $a^2 + b^2 = 1$. So a qubit state contains *two* real parameters.

For a Hilbert space, the inner product, also known as 'overlap', of any two states $|\psi_1\rangle$ and $|\psi_2\rangle$ is defined as $\langle\psi_1|\psi_2\rangle$. Two states are orthogonal if their overlap is zero. For dimension d, a set of mutually orthogonal state is an *orthonormal basis* such that any state can be written as a linear superposition of them. The states $|0\rangle$ and $|1\rangle$ form an orthonormal basis for a qubit.

Question 11. *Is there a geometric representation of a qubit?*

If we use the complex plane, then it can be viewed as two points. But this is not so satisfactory (although this might work for qudit cases; e.g., the Majorana representation). We need to use a single point to account for two parameters. It turns out we need a three-dimensional object, which is known as *Bloch sphere*, a 2-sphere embedded in $\mathbb{R}^3$, see Fig. 1.1.

The triangular parameters $a = \cos\theta/2$, $b = \sin\theta/2$. The factor of 2 is there since a rotation of $\theta = \pi$ leads to its orthogonal state instead of itself. We then have

$$|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\varphi}\sin\frac{\theta}{2}|1\rangle. \tag{1.3}$$

This is also known as a *pure* state.

Mixed states, or density matrices, are those points inside the Bloch sphere, so together they form the Bloch ball. Intuitively, a mixed state is a point whose distance to the centre of the ball is smaller than one. How to quantify this?

Mixed states cannot be written as vectors; instead they are matrices. A mixed state can be expanded with Pauli matrices $\{\sigma^i\} = \{\mathbb{1}, X, Y, Z\}$ $(i = 0,1,2,3)$ as

$$\rho = (\mathbb{1} + \vec{n}\cdot\vec{\sigma})/2, \tag{1.4}$$

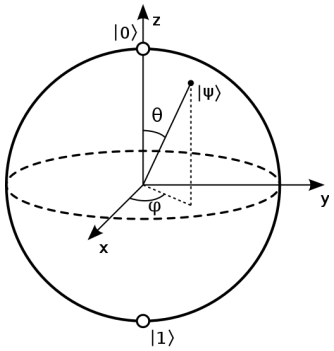

Figure 1.1: The Bloch sphere.

with

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.5}$$

It is trace one $\mathrm{tr}\rho = 1$, and the purity is $\mathrm{tr}\rho^2 = (1+|\vec{n}|^2)/2 < 1$. The vector $\vec{n}$ is also known as Bloch vector. For pure state, we have $n_x = \sin\theta\cos\phi$, $n_y = \sin\theta\sin\phi$, $n_z = \cos\theta$. For mixed state, we can denote $\vec{n} = r(n_x, n_y, n_z)$ for $r \in [0,1)$ as the length of the Bloch vector. We see that a mixed qubit state contains three real parameters!

1.3 Unitary evolution

As a matrix, a qubit state is

$$\rho = \begin{pmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{pmatrix}. \tag{1.6}$$

The diagonal elements ρ_{00} and ρ_{11} are called 'population' as they are the probability on states $|0\rangle$ and $|1\rangle$, and off-diagonal elements ρ_{01} and ρ_{10} are called 'coherence', and they satisfy

$$\rho_{00} + \rho_{11} = 1, \rho_{01} = \rho_{10}^*, \tag{1.7}$$

and $\rho_{00}\rho_{11} \geq \rho_{01}\rho_{10}$ due to positivity of ρ. For pure states, $\rho_{00}\rho_{11} = \rho_{01}\rho_{10}$.

Given a state, what shall we expect? Well, just like any object, you can apply drives (or forces) on it to see how it evolves. This is the subject of quantum dynamics which studies the evolution of quantum systems. The evolution could be invertible or not. As quantum states form Hilbert space, invertible evolution of them is described as *unitary operator* U so that

$$U^\dagger = U^{-1}, \tag{1.8}$$

and it requires $\det U = 1$ for simplicity.

When U takes the form $U = e^{itH}$, it is generated by a Hamiltonian H, which is hermitian. The Liouville–von Neumann equation is

$$i\dot{\rho} = [H, \rho]. \tag{1.9}$$

For pure states, it reduces to the state equation

$$i|\dot{\psi}\rangle = H|\psi\rangle. \tag{1.10}$$

Note we take $\hbar = 1$ for simplicity.

Question 12. *How to solve the above quantum equation?*

The solution takes the form

$$|\psi(t)\rangle = U|\psi(0)\rangle \tag{1.11}$$

for an initial state $|\psi(0)\rangle$. If H can be diagonalized $H = V\Lambda V^\dagger$, for the diagonal terms of Λ known as 'energy', then $U = Ve^{it\Lambda}V^\dagger$. Now we have encountered the first but most crucial problems in quantum theory: diagonalize matrices, which turn out to be very difficult in general! (and this basically motivates the subject of quantum computing.) However, there are enough solvable examples in physics with interesting phenomena, as discussed below.

1.3.1 Rabi oscillation

For a qubit with a diagonal Hamiltonian H, the two states are eigenstates and will obtain phases under free evolution. Interesting thing would happen when off-diagonal terms are added. Rabi oscillation is an example, which can describe the dynamics of a two-level atom driven by a classical electromagnetic field.

The population of states and coherence will oscillate at Rabi frequency. This shows the periodic exchange of energy between the atom and the field.

What is the energy that is exchanged? They are carried by photons. The field can be quantized as a collection of photons. The Jaynes-Cummings Hamiltonian for a two-level atom coupled to a quantized electromagnetic field is given by

$$H = \omega(a^\dagger a + 1/2) + \omega_0\sigma^z/2 - i\Omega\sigma^x(a - a^\dagger). \tag{1.12}$$

With the rotating wave approximation (RWA), it becomes

$$H = \Delta\sigma^z - i\Omega(\sigma^+ a - \sigma^- a^\dagger), \tag{1.13}$$

with $|\Delta| = (\omega_0 - \omega)/2$ as the detuning, $\sigma^\pm = (\sigma^x \pm \sigma^y)/2$. The RWA ignores $\sigma_+ a^\dagger$ and $\sigma_- a$, i.e., excite or de-excite both the spin and the field which apparently do not preserve energy. The Rabi frequency is found by diagonalizing H to be $\sqrt{\Delta^2 + \Omega^2}$.

1.3.2 Periodic Hamiltonian

When Hamiltonian has parameter $r \in [0, R]$ such that it is periodic

$$H(0) = H(R), \tag{1.14}$$

there are interesting physical effects.

Question 13. *How special can a periodic Hamiltonian be?*

First, this is effectively a symmetry, denoted as G and

$$[G, H(r)] = 0. \tag{1.15}$$

Any state $\psi(r)$ will contain the parameter r. A common set of eigenstates $\{\psi(r)\}$ for G and $H(r)$ can be chosen and $\psi(r+R) = \lambda\psi(r)$. For instance, when r is space position x, G is the translation operator, R can be the lattice constant, and λ becomes e^{ikR} for Bloch wave-vector k. The interval $(-\pi/R, \pi/R]$ is known as the Brillouin zone. The set of states $\{\psi_{nk}(r)\}$ is known as Bloch bands. The relation between $E_n(k)$ and k is the dispersion relation. For a periodic Hamiltonian with time-reversal symmetry the energy bands are degenerate between k and $-k$, which is the Kramers degeneracy.

If the parameter is time t, then there is a special structure of its unitary evolution. Floquet theory shows that the time evolution operator $U(t)$ by $H(t)$ takes the form

$$U(t) = e^{itK}V(t) \tag{1.16}$$

for a hermitan operator K, a unitary operator $V(t)$ with $V(0) = \mathbb{1}$ and $V(t) = V(t+T)$. However, K is not easy to find. Special examples can be solved, e.g., for adiabatic evolution K can be approximated by $H(t)$ itself.

1.3.3 Parameter-dependent Hamiltonian

Given a Hamiltonian H, the spectrum is given by its eigenstates. The equation $i|\dot{\psi}(t)\rangle = H|\psi(t)\rangle$ can be solved. If the initial state $|\psi(0)\rangle$ is an eigenstate, then the evolution yields $|\psi(t)\rangle = e^{iEt}|\psi(0)\rangle$ with E as the energy of $|\psi(0)\rangle$. We see that the dynamics only generates a phase factor e^{iEt}.

For a time-dependent Hamiltonian, the physics becomes more interesting. First, the equation

$$i|\dot{\psi}(t)\rangle = H(t)|\psi(t)\rangle \tag{1.17}$$

cannot be easily solved in general. Note here $H(t)$ is not periodic. There does exist a formal solution, known as Dyson series (Chapter 2), which describes the solution as a sum of many terms, while you do not know how many terms you need to calculate, and these terms are hard to calculate. This method together with Feynman diagram is widely used in particle physics.

If the time-dependence of $H(t)$ is of special forms, we might hope to obtain solutions. Below we analyze an example to illustrate the idea. Consider a spin-1/2 in a magnetic field, while the field is time-dependent. The direction of $B(t)$ is changing smoothly. With polar coordinates, let's consider

$$\vec{B}(t) = B(\sin\theta\cos\omega t, \sin\theta\sin\omega t, \cos\theta), \tag{1.18}$$

i.e., the field strength B does not change, while it is rotating around the z-axis with an offset angle θ (see Fig. 1.1 but with state vector replaced by field vector). The rotation frequency ω is constant. Now, what is the form of $H(t)$ for the spin? It takes the form

$$H(t) = b\vec{B}(t)\cdot\vec{S} \tag{1.19}$$

for the three-component spin operator $\vec{S} = (\sigma^x, \sigma^y, \sigma^z)$. Here we ignore the details for the constant b, which depends on the physical particle that carries the spin.

The model $H(t)$ is simple enough so it can be exactly solved! First, with easy algebra it can be written as

$$H(t) = V^\dagger H' V, V := e^{i\omega t\sigma^z}, H' = e^{-i\theta\sigma^y}\sigma_z e^{i\theta\sigma^y}. \tag{1.20}$$

So the time-dependence is encoded in a unitary operator U which acts on a fiducial model H'. Note that in general, the time-dependence of a model $H(t)$ cannot be extracted as an external unitary operator U. Here it just happens to be the case.

Also the model we study is periodic: $H(T) = H(0)$ for $T = 2\pi/\omega$. This means that the eigenvalues and eigenstates will come back to themselves after T. However, there are additional phase factors in front of each eigenstate, as we will see below. The model can be exactly solved with solution

$$|\psi(t)\rangle = e^{-it\omega\sigma^z} e^{-it(H(0)-\omega\sigma^z)}|\psi(0)\rangle. \tag{1.21}$$

Now we define the Aharonov-Anandan dynamical phase

$$\alpha^d := \int_0^T \langle\psi(t)|H(t)|\psi(t)\rangle dt. \tag{1.22}$$

The total phase α is defined as

$$e^{-i\alpha} := \langle\psi(T)|\psi(0)\rangle, \tag{1.23}$$

and the geometric phase is $\gamma := \alpha - \alpha^d$. We find that

$$\gamma = 2\pi(1 + \frac{\omega}{\Omega} - \frac{b}{\Omega}\cos\theta), \tag{1.24}$$

for $\Omega := b\sqrt{1+\omega(-2\cos\theta+\omega/b)/b}$. In the limit that $\omega/b \ll 1$, the phase reduces to $\gamma = 2\pi(1-\cos\theta)$, which is actually known as Berry's *adiabatic* geometric phase.

Question 14. *How good is adiabatic evolution?*

The model above just happens to be exactly solvable. There are many models that are not the case. However, as hinted above, if the evolution is adiabatic, the solution can also be easily handled. When the field rotates slowly in the *adiabatic* regime, we can use the adiabatic solution. We can denote the instaneous eigenstates as $|k,t\rangle$

with $H(t)|k,t\rangle = E(k,t)|k,t\rangle$. Here the energy $E(k,t)$ are time-independent, so can be simply denoted as E_k, which is proportional to k. In this regime, if it starts on $|k,0\rangle$, the system will stay on its instaneous eigenstate $|k,t\rangle$ with the same k. There will be a tiny jump to other states, though. Given a set of instaneous eigenstates $\{|k,t\rangle\}$, the condition for the adiabatic approximation is that

$$\langle k,t|\partial_t|k',t\rangle \approx 0, \quad \forall k \neq k'. \tag{1.25}$$

This turns out to be necessary and sufficient for unitary evolution. This is also known as 'parallel transport'. The condition can also be written as

$$\frac{\langle k,t|\partial_t H(t)|k',t\rangle}{E_k(t) - E_{k'}(t)} \approx 0, \quad \forall k \neq k'. \tag{1.26}$$

For the spin model we study, this becomes the condition $\omega/b \ll 1$. The state $|k,T\rangle$ can be written as

$$|k,T\rangle = e^{-ib2\pi k/\omega} e^{-i\gamma_k}|k,0\rangle. \tag{1.27}$$

Note that, however, state $|k,T\rangle$ does not satisfy state equation exactly, since the components of other states are ignored. The phase

$$\gamma_k = k\int_0^{2\pi}\int_0^{\theta} \sin\theta' d\theta' d\varphi = 2\pi k(1-\cos\theta) \tag{1.28}$$

is known as Berry's adiabatic phase, or Berry phase. The factor γ_k/k is also known as 'solid angle', which is a geometric quantity. We see that, it depends on θ, so it is not topological.

The model above has the same cyclic period with the period of the Hamiltonian. This is not necessary. There are geometric phases for non-periodic Hamiltonian as long as there is a cycle in parameter space. The *cyclic condition* is more general than the *periodic condition* of a Hamiltonian (see Chapter 2).

Question 15. *How to measure the geometric phase?*

For spin-1/2 there are two states. When they both obtain geometric phases, we can measure the difference of them. Alternatively, we can embed the two-level system into a qudit, say, a three-level qutrit. This experiment has been done with a triplet. The idea is this: use the third level as a reference. Create some coherence between levels two and three. You can measure the coherence directly. You can also measure it after a cyclic evolution of the subspace of levels one and two. The difference of the two cases is just the geometric phase.

Geometric phases appear in many contexts. The parameter space for H can be variant, in general, it could be any space. There are geometric phase in solid-state physics, e.g., due to the periodicity of Brillouin zone in momentum space. The Aharonov-Bohm phase is an example of *topological* phase. The cycle is the loop

travelled by the electrons in space. The electron does not have to move adiabatically, but it assumes that the local field cannot be changed by the electron. The field is semi-classical and does not depends on the electron dynamics. The anyons in topological states of matter generate abelian and non-abelian geometric phases under braiding. Finally, for infinite-dimensional system, such as harmonic oscillator, there are also forms of geometric phases.

1.3.4 Quantum control

Quantum control, as a technique, is the extension of control technique (or theory) to the quantum regime. As the name suggests, it aims to control a quantum system for some purposes. Here we are interested in a single qubit: given the Bloch ball, how can we play with it?

Question 16. *What is the goal of quantum control?*

A class of question is about *controllability*: is it possible to get arbitrary desired state or unitary evolution given a set of control items (or schemes)? The controllability is roughly the same as *universality* in quantum computing, which aims to simulate arbitrary gates in a unitary group efficiently (see Chapter 4). Also note that controllability depends on the control schemes.

In the setting of Hamiltonian control, we have $H(\lambda) = H_0 + H_\lambda$, and the control term H_λ has a set of tunable parameters λ. The geometric phase can be viewed as an example of this. Here, in general, the goal is to use H_λ, which may stand for interaction of the qubit with external fields, to drive the qubit to desired states.

From algebra, it is easy to see that if $H(\lambda)$ can generate the algebra $su(2)$, then $e^{itH(\lambda)}$ can realize the whole group $SU(2)$. To generate $su(2)$, $H(\lambda)$ must contain non-commuting terms. For instance, $H_0 = \sigma^z$, $H_\lambda = \lambda\sigma^x$. By turning λ on and off, any unitary evolution can be achieved.

But note that, unitary evolution preserves purity; namely, it preserves the size of Bloch vector r. That is to say, the qubit can only move around on a sphere of radius r. In order to change r, we have to use incoherent control methods, which are described as non-unitary quantum channels. A surprising but reasonable fact is that, Lindblad dynamics is not Hamiltonian controllable. Instead, if measurement or non-unitary control can be used, then Lindblad dynamics becomes controllable. In fact, this becomes a universality or quantum simulation problem.

A different question we concern is how well we can control, and this is the task of *optimal control*. Optimal control aims to maximize some objective function or observable. This could be anything you want: such as time, speed, energy, space, stability, purity, coherence, etc. As an optimization problem, there are lots of optimization algorithms. Here we look at a primary one.

Question 17. *How fast can a quantum system evolve?*

This is to minimize the time, or maximize the speed of evolution. To answer this, we first need a notion of 'speed' for quantum dynamics, which we do not know yet. In classical mechanics, the speed is the rate of change of position x. In quantum theory position is an operator $\hat{x}$, its evolution is from the Heisenberg equation

$$i\dot{A} = [A,H] \tag{1.29}$$

when A is $\hat{x}$. We can treat expectation value as the classical version of a quantum operator (see Chapter 3), so the speed (or velocity) can be defined as $\langle\psi|i[H,\hat{x}]|\psi\rangle$ for a state $|\psi\rangle$. For a free particle $H = \hat{p}^2/2m$, we find the speed is $\langle\psi|\hat{p}|\psi\rangle/m$. For a harmonic oscillator, we find the same expression.

We could define speed or rate of change more generally as $\dot{A}$ for any observable A. Then the speed is simply $\langle\psi|i[H,A]|\psi\rangle$. Clearly if A commute with H, then speed is zero, which means A is preserved by the evolution. Yet even if $[H,A] \neq 0$, the speed can be zero since it depends on the state $|\psi\rangle$. We see that the speed of evolution is a joint property of both observable and state.

Another way to characterize speed is from the standard derivation and uncertainty relation (see Chapter 2). The standard derivation of A is defined as

$$\Delta A = \sqrt{\langle A^2\rangle - \langle A\rangle^2} \tag{1.30}$$

for $\langle A\rangle := \langle\psi|A|\psi\rangle$, which clearly depends on the state $|\psi\rangle$. The uncertainty relation for two noncommuting operators A and B can be derived easily

$$(\Delta A)^2(\Delta B)^2 \geq |\frac{1}{2}\langle\{A,B\}\rangle - \langle A\rangle\langle B\rangle|^2 + |\frac{1}{2i}\langle[A,B]\rangle|^2. \tag{1.31}$$

Now we want to include time t, which, however, is not an operator! Observe that in the uncertainty relation there is a term $\langle[A,B]\rangle$. If B is the Hamiltonian H, then this term is $\dot{A}$, which contains time t. If we carry out this we obtain

$$\tau_A \Delta H \geq \frac{1}{2}, \tag{1.32}$$

which is usually known as the time-energy uncertainty relation, if we identity ΔH as δE, the uncertainty of energy. $\tau_A := \frac{\Delta A}{|\frac{d}{dt}\langle A\rangle|}$ is the time needed to specify A within ΔA. Note this requires $\frac{d}{dt}\langle A\rangle \neq 0$, i.e., the value $\langle A\rangle$ has to change under the evolution, which, of course, shall be the starting point to study this.

It turns out the time-energy uncertainty relation is very useful. It can be used to prove lower bound for certain quantum computing algorithms, (e.g., Grover search), and it can also be used to design optimization algorithms for various control problems.

Control technique is a great toolbox for many things. We can use control technique to realize quantum gates, to reduce noise level, to measure observable precisely, which can be further used for other purposes.

1.4 Non-unitary evolution

1.4.1 Quantum channels

A qubit unitary evolution has 3 real parameters. From linear algebra, there is a very elegant formula for arbitrary qubit unitary

$$U(\gamma,\beta,\alpha) = e^{i\gamma\sigma^x} e^{i\beta\sigma^z} e^{i\alpha\sigma^x}. \tag{1.33}$$

Question 18. *How many parameters are needed to specify a qubit non-unitary evolution?*

It turns out a general qubit non-unitary evolution needs 12 real parameters. This is surprising since, on one hand, it does not need infinite many, and on the other hand, this number is not small. All this is due to the theory of quantum channels.

Non-unitary evolution have to preserve positivity of quantum states, and it turns out they have to be completely positive and trace-preserving (CPTP), known as quantum channels. For an arbitrary qubit quantum channel $\mathscr{E} : \mathscr{D}(\mathscr{H}_2) \to \mathscr{D}(\mathscr{H}_2) : \rho \mapsto \mathscr{E}(\rho)$, the Kraus operator-sum representation is

$$\mathscr{E}(\rho) = \sum_{i=0}^{r-1} K_i \rho K_i^\dagger \tag{1.34}$$

for the rank $r \leq 4$ and a set of Kraus operators $\{K_i\}$, which form a linearly independent set, and the trace-preserving condition is $\sum_i K_i^\dagger K_i = \mathbb{1}$. The set of Kraus operators is equivalent to an isometry

$$V = \sum_i K_i |i\rangle, \tag{1.35}$$

which satisfies $V^\dagger V = \mathbb{1}$, but $VV^\dagger \neq \mathbb{1}$.

As an isometry can be viewed as part of a unitary operator, we can 'dilate' a channel to a unitary operator U such that $K_i = \langle i|U|0\rangle$ for $\{|i\rangle\}$ an orthonormal basis state of the 'ancilla'. This forms the Stinespring dilation.

It turns out there are more representations of channels. Below are the standard ones.

- Choi state and process state representations. In Pauli basis $\{\sigma_i\}$, the process matrix χ, is defined with entries $\chi_{jk} = \sum_i \mathrm{tr}(K_i^\dagger \sigma_j)\mathrm{tr}(K_i \sigma_k)^*$. It holds $\mathscr{C} = U\chi U^\dagger$ for the basis transformation

$$U = \frac{\sqrt{2}}{2} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & -i & 0 \\ 0 & 1 & i & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix}, \tag{1.36}$$

with $U_{\alpha\beta} = \mathrm{tr}(\tau_\alpha^\dagger \sigma_\beta)$, from Pauli basis to Kronecker basis $\{\tau_\alpha\}$ (see Chapter 2). The Choi state is a four-by-four matrix that can have 16 parameters. With the trace-preserving condition, there are totally 12 independent real parameters.

- Dynamical representations. The affine map for a qubit channel is

$$\mathscr{T} = \begin{pmatrix} 1 & \mathbf{0} \\ \boldsymbol{t} & \boldsymbol{T} \end{pmatrix}, \tag{1.37}$$

which is a four-by-four real matrix with $\mathscr{T}_{ij} = \frac{1}{2}\mathrm{tr}[\sigma_i \mathscr{E}(\sigma_j)]$. In this representation, $\rho = \frac{1}{2}(\mathbb{1} + \boldsymbol{p}\cdot\boldsymbol{\sigma})$, and the channel is an affine map

$$\mathscr{T} : \boldsymbol{p} \mapsto \boldsymbol{p}' = \boldsymbol{T}\boldsymbol{p} + \boldsymbol{t}. \tag{1.38}$$

Geometrically, $\mathscr{E}$ maps the Bloch ball into an ellipsoid, with $\boldsymbol{t}$ the shift from the ball's origin and $\boldsymbol{T}$ a distortion matrix for the ball. The $\boldsymbol{T}$ matrix can be diagonalized from singluar-value decomposition

$$\mathscr{T} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ s_1 & \lambda_1 & 0 & 0 \\ s_2 & 0 & \lambda_2 & 0 \\ s_3 & 0 & 0 & \lambda_3 \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ \boldsymbol{s} & \boldsymbol{\Lambda} \end{pmatrix}, \tag{1.39}$$

with two rotations O_1 and O_2 such that $O_2\boldsymbol{s} = \boldsymbol{t}$, $O_2\boldsymbol{\Lambda}O_1 = \boldsymbol{T}$. As $SU(2)/\mathbb{Z}_2 \cong SO(3)$, the rotations O_1 and O_2 correspond to prior and posterior $SU(2)$ rotations U_1 and U_2, respectively. There are totally 12 independent parameters in $\mathscr{T}$, with six from the prior and posterior rotations, and six from the rest. Also the dynamical operator $\mathscr{D}$, known as 'transfer matrix', can be obtained as $\mathscr{D} = U\mathscr{T}U^\dagger$. The dynamics is $\mathrm{res}(\rho) \mapsto \mathscr{D}\mathrm{res}(\rho)$ for $\mathrm{res}(\rho) = (\rho_{00}, \rho_{01}, \rho_{10}, \rho_{11})^{\mathrm{t}}$.

We see that a qubit channel needs 12 real parameters. However, there is a little caveat: from the dilation form, the channel is dilated to a unitary on three qubits. This unitary needs $d^6 - 1 = 63$ real parameters. It is clear that there is a big gap there. The reason for such a mismatch is that there are more structures of channels that are not used yet.

Question 19. *What are the basic properties for the set of channels?*

The additional structures of channels is the *convexity* of the set of channels. The set of qubit channels is convex, and it is a convex body, i.e., it has infinite many *extreme points*. Just like the set of states, the extreme points are pure states which are rank one, and there are infinite many of them. Due to the convexity, a channel or state can be written as a convex sum of extreme points. For a state, the number of extreme points needed is the rank of the state. For instance, a qubit state needs at most two

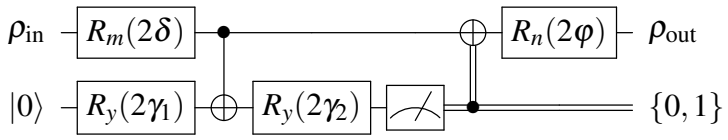

Figure 1.2: The rank-two qubit channel circuit. $R_y(2\gamma) = e^{-iY\gamma} = \mathbb{1}\cos\gamma - iY\sin\gamma$; the two angles are $2\gamma_1 = \beta - \alpha + \pi/2$ and $2\gamma_2 = \beta + \alpha - \pi/2$. The final operation is a classically controlled X operation.

pure states for such a decomposition, which may be the eigenstate decomposition. It turns out, there are extreme channels whose rank are higher than one. As a result, the convex decomposition of any qubit channel becomes nontrivial.

It turns out a channel is extreme iff the set $\{K_i^\dagger K_j\}$ is linearly independent. This implies that the rank of an extreme channel is at most d for qudit channel. How many parameters there are for a rank-d qudit channel? It turns out the number is $2d^2(d-1)$, which is 8 for the qubit case. It is clear to see that rank-two channels that are not extreme can be written as sum of rank-one channels, i.e., unitary operators. Furthermore, it is known that any qubit channel can be written as a sum of at most two rank-two channels

$$\mathscr{E} = p\mathscr{E}_1^g + (1-p)\mathscr{E}_2^g \tag{1.40}$$

for $\mathscr{E}_i^g$ $(i = 1,2)$ as rank-two channels. In general, a qudit channel can be written as a sum of a certain number of rank-d channels. From dilation, a rank-d channel is dilated to a d^2-dimensional unitary, which has $O(d^4)$ parameters, in the same order of the parameters in a channel (which is $d^4 - d^2$). As a result, a qubit channel can be realized as convex sum of two channels, each realized by a unitary acting on two qubits.

There is a simple *canonical form* of rank-two qubit channels $\mathscr{E}^g$. The two Kraus operators can be expressed as

$$F_0 = \begin{pmatrix} \cos\beta & 0 \\ 0 & \cos\alpha \end{pmatrix}, \quad F_1 = \begin{pmatrix} 0 & \sin\alpha \\ \sin\beta & 0 \end{pmatrix}, \tag{1.41}$$

up to pre- and post- unitary rotations on the qubit. The two Kraus operators can be realized by the unitary operator

$$U = \text{CNOT}M_{10}(\alpha,\beta), \tag{1.42}$$

where the multiplexor $M_{10}(\alpha,\beta) := R_y(2\beta) \oplus R_y(2\alpha)$, and the CNOT gate allows the ancilla as control. See Fig 1.2, with the pre- and post- rotations for basis transformation. There are in total 8 real parameters, agree with the parameter counting for rank-two qubit channels.

1.4.2 Thermodynamics

We know that when a system experience non-unitary evolution, its state can be mixed states. The non-unitary evolution may come from ignoring, e.g., tracing out a part of the whole system. This can describe the situation of thermodynamics when the traced out part is a 'big' bath or environment.

A bath is described by some macroscopic emergent variables, such as temperature, pressure, volume, entropy etc and we do not care of its microscopic details. For a bath at temperature β, a system at equilibrium with it will be at state

$$\rho = e^{-\beta H} \tag{1.43}$$

for H as the Hamiltonian of the system, renormalized by interactions with the bath. Here $\beta = \frac{1}{T}$, for T as the usual symbol for temperature. We can see the state ρ is a Boltzmann distribution in the basis of eigenstates of H. If $H = \sum_i E_i |\psi_i\rangle\langle\psi_i|$, then $\rho = \sum_i e^{-\beta E_i} |\psi_i\rangle\langle\psi_i|$. But note the eigenstates of H may be hard to find in practice, and the values $e^{-\beta E_i}$ are hard to measure, too.

Here we are more interested in how to use quantum thermal process for some purposes, e.g., to harvest energy. This is the subject of quantum heat or refrigeration engines. An engine, in general, is a great device that can convert or transfer energies. We surely know Carnot engine, Otto engine, Maxwell demon, and other ones, so here let us see how to build a quantum engine via just a single qubit.

Question 20. *How to use a qubit as an energy-harvest engine?*

First note that the energy is the expectation value of H on a state $|\psi\rangle$ as $E = \langle\psi|H|\psi\rangle$. On mixed state it can be written as $E = \sum_i p_i E_i$. Now the change of energy contains two terms

$$\delta E = \sum_i p_i \delta E_i + \sum_i \delta p_i E_i. \tag{1.44}$$

The two terms are nothing but the work and heat changes, respectively, since work refers to changes of energy levels, while heat refers to distribution over a set of energy levels. For a qubit, both of them can be easily done.

In the usual engine model, there are hot bath, cold bath, and worker (work substance). For a refrigerator, the worker extracts energy from the cold bath and dispose to the hot bath. For a quantum refrigerator, we shall use the qubit as the worker, and assume the two baths are each at thermal equilibrium. The qubit interacts with the two baths alternatively. From the point of view of the qubit, it just undergoes a sequence of non-unitary processes described by master equation or quantum channel. See the refrigerator model and its quantum circuit in Fig. 1.3.

During a cycle, the qubit first contacts with the cold bath, absorb energy, and thermalize to state $\rho_1 = e^{-\beta_c H}$. The qubit is said to be also at temperature β_c. Now we need to modify H to $\tilde{H}$ such that the effective temperature T_1 is higher than T_h, i.e., $\beta_1 < \beta_h$. This can be done by tuning the energy gap Δ of the qubit to a bigger

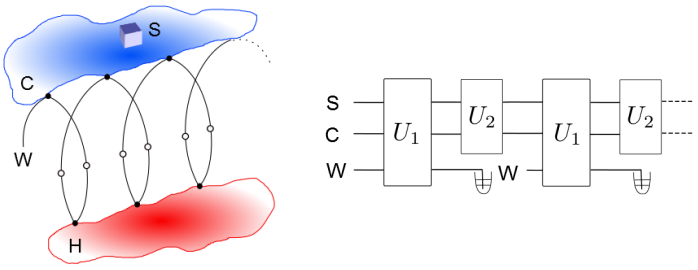

Figure 1.3: The refrigerator model and its quantum circuit.

value $\tilde{\Delta}$. Then the qubit contacts with the hot bath, dismiss energy, and thermalize to state $\rho_2 = e^{-\beta_h \tilde{H}}$. This complete the energy transfer from the cold bath to the hot bath, yet external work is done on the qubit, just like the compression of the work gas in the classical refrigerator. To start the next cycle, $\tilde{H}$ needs to be changed back to H so that the effective temperature T_2 is lower than T_c, and then the qubit can absorb energy from the cold bath again.

The external work on the qubit is required according to the second law of thermodynamics. The reverse process of the refrigerator is the heat engine, which generates work from the heat flow between the hot and cold baths. The fundamental fact is that the coefficient of performance ε of refrigerator is smaller than the classical one

$$\varepsilon = \left(\frac{\tilde{\Delta}}{\Delta} - 1\right)^{-1} \leq \left(\frac{\beta_c}{\beta_h} - 1\right)^{-1}, \tag{1.45}$$

since $1 < \frac{\beta_c}{\beta_h} \leq \frac{\tilde{\Delta}}{\Delta}$. Conversely, for Carnot heat engine it requires $\frac{\beta_c}{\beta_h} \geq \frac{\tilde{\Delta}}{\Delta} > 1$ (note $\tilde{\Delta}$ is the original gap), then the efficiency η is also bounded

$$\eta = 1 - \frac{\Delta}{\tilde{\Delta}} \leq 1 - \frac{\beta_h}{\beta_c}. \tag{1.46}$$

It seems the quantum case is no better than classical ones. This is not true. The simple engines above do not really employ quantum features yet since it only uses thermal process. There are states that do not have classical analog, such as squeezed states, which turn out to be powerful to surpass the classical bounds. When larger quantum system is being used as the worker, it turns out entanglement among subsystems is also powerful.

1.5 Observable and Measurement

Question 21. *What can be observed on a qubit?*

We see that above a qubit has a state which can evolve. So at least we can try to observe its state. It turns out there are also a lot of other quantities, or observable, that

can be observed. If the qubit has a Hamiltonian, then the energy is an observable. If it has spin, then the value of spin can be measured. In quantum theory, observable, denoted by A, are hermitian operators acting on a Hilbert space with

$$A^\dagger = A. \tag{1.47}$$

An observable does not have to be invertible or unitary. However, for qubit, the three Pauli operators are both unitary and hermitian! This is very special and useful, and does not generalize to qudits.

What is observed is usually the expectation value $\mathsf{tr}(A\rho)$ on a state ρ. This generalizes some expressions in thermodynamics via partition function. Suppose at time t we need to measure $\mathsf{tr}(A\rho(t))$. When $\rho(t) = U\rho(0)U^\dagger$, and from $\mathsf{tr}(AB) = \mathsf{tr}(BA)$, we find $\mathsf{tr}(A\rho(t)) = \mathsf{tr}(A(t)\rho(0))$ for

$$A(t) := U^\dagger A U. \tag{1.48}$$

This fact is important: it says the evolution effect can be attributed to the observable A, while the state does not evolve. Also note this only holds when the value of A is observed. This turns out to be fundamental in quantum theory, leading to the 'Heisenberg picture'. In Heisenberg picture, observable evolves while state does not; in 'Schrödinger picture', state evolves while observable does not. They are equivalent on the observation level since the same thing shall be observed! The Schrödinger picture is usually employed, though.

1.5.1 Projective measurement

Now, how to perform a measurement? As you can see, a measurement extracts some values, i.e., classical quantities, from the final state, so it cannot be unitary. That is to say, it has to be a quantum channel. However, there is a central difference between channel and measurement. In measurement, the result of each Kraus operator K_i is recorded. If you treat each Kraus operator (or some of them) as a not trace-preserving channel, then a measurement is a collection of them and in total it is trace-preserving; and this is sometimes known as 'quantum instrument'. Therefore, a measurement will lead to a set of classical records, probabilities, and a set of corresponding states. If some of results can be thrown away, this is called a 'post-selection', which proves to be computationally powerful.

There are many kinds of measurements. The simplest kind is known as projective measurement, which is made up by a set of projectors. Denote a projective measurement as $\mathscr{M} = \{P_i\}$, for $P_i = |\psi_i\rangle\langle\psi_i|$. The states $|\psi_i\rangle$ might be eigenstates of some observable A with $A|\psi_i\rangle = a_i|\psi_i\rangle$, and a_i are the eigenvalues. The value $\mathsf{tr}(A\rho) = \sum_i a_i p_i$ for $p_i = \langle\psi_i|\rho|\psi_i\rangle$ as probabilities. Here note that p_i may not be the eigenvalues of ρ as in general A and ρ do not commute, so do not share a set of eigenstates.

Question 22. *How to perform a projective measurement?*

Let's look at an example of the measurement of photon polarization. The photon polarization has two orthogonal states: horizontal $|H\rangle$ and vertical $|V\rangle$ states. Any polarization is a superposition (or mixture) of the two. How it is measured? Well, this is simple: just let photons pass through a polarization beam splitter, and they will split into two bunches, then use photon detectors to detect photons in each bunch. For a state $|\psi\rangle = \cos\theta|H\rangle + \sin\theta|V\rangle$, the value θ can be measured. However, after the measurement the photons are gone. If we want to keep the states after the measurement, we need something more advanced. In quantum optics, this is a quantum 'non-demolition' measurement. This can be done by coupling the photons to another quantum system, the so-called ancilla, then the measurement on ancilla will induce measurement on the photons. This is a kind of 'indirect measurement'. Note 'non-demolition' does not mean the state is not disturbed or measured.

We know that a quantum channel can be realized by its dilation and tracing out the ancilla. So the measurement on the ancilla can realize measurement on a system. A projective measurement, with no doubt, can also be realized in this way. The beam splitter is a direct measurement since it projects (or collapses) each photon onto either $|H\rangle$ or $|V\rangle$ state. We see that a projective measurement can be realized either directly or indirectly. A projective measurement is also known as a 'sharp' measurement.

1.5.2 Non-commuting set of observable

The set of projectors in a projective measurement can be viewed as the eigenstates of an observable A. Furthermore, it is also the eigenstates of any other observable B with $[A,B]=0$. On a state ρ, the value $\mathsf{tr}(A\rho)=\sum_i a_i p_i$ and $\mathsf{tr}(B\rho)=\sum_i b_i p_i$ as above. A joint, or simultaneous measurement of a set of observable $\{O_i\}$ means that the values of O_i can be read out simultaneously. One does not have to measure them sequentially. What are measured are the probabilities p_i. The values a_i and b_i are pre-calculated by hand. In a sense, we can say that any observable that takes the set of projectors in a projective measurement is measured simultaneously.

Question 23. *But, how to measure non-commuting observable simultaneously?*

Apparently, this can not be done. The uncertainty relation tells us that product of standard deviations on a state ρ is lower bounded

$$\Delta A \Delta B \geq \mathsf{tr}(\rho[A,B])/2 \tag{1.49}$$

due to the noncommuting part $[A,B]$. Note the lower bound is state-dependent. The equality holds when *the actions of A and B on ρ are equivalent*. It does not require A and B commute. But if we prefer a state-independent setting, then they have to commute to be measurable simultaneously.

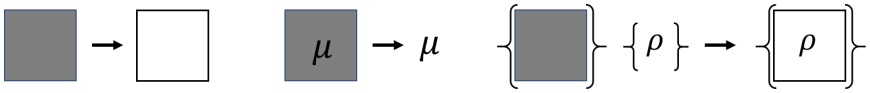

Figure 1.4: Quantum tomography (left), estimation (middle), and discrimination (right).

The standard deviations describes statistical imprecision, not disturbance to a system. The physical content of the relation is that: if ΔA is upper bounded, then ΔB will be lower bounded. It is a 'logical' relation, and it does not refer to joint or sequential measurement of A and B.

With no surprise, we can measure noncommuting observable sequentially, in either order. But after the first measurement, the state is modified to be a set of post-measurement states. A disturbance on B due to measurement of A can be defined, and an 'error-disturbance' tradeoff relation can be derived. This means there is a tradeoff on sequential measurement. Yet the relation could be product or sum of these terms, and the relations are not unique.

It is better to measure noncommuting observable in parallel; namely, prepare identical sample states and perform measurements independently.

1.5.3 Tomography, Estimation, Discrimination

Given a state or process that we do not know yet, how to know what it is? This is a black or white box problem. This is the task of tomography, estimation, discrimination or others, generalizing the goal of measurement.

Question 24. *Quantum system, a black or white box?*

For tomography, the given operator is usually a black box, i.e., we do not know any information of it, probably except the dimension of the system. When some information is known, e.g., a formula of state as a function of some parameters, then estimation technique can be used instead of the expensive tomography. In addition, sometime we do not need to know the full information of the unknown operators; instead we may need to make an assignment, namely, discriminate some operators according to a certain rules. See Fig. 1.4 for illustration of the three tasks.

State tomography is to determine the parameters of an unknown state. For d dimension, there are d^2-1 real parameters for a mixed state. The scheme is to make a set of measurements on the state, and use the obtained probabilities to determine the state. For a POVM $\{M_i\}$, $p_i = \mathrm{tr}(\rho M_i)$ (see Chapter 2). This set of equality is enough to determine ρ if there are d^2-1 or more p_i. Use vectorization form $\mathrm{tr}(\rho M_i) = \langle\rho|M_i\rangle$ for $|\rho\rangle = \mathrm{res}(\rho) = \rho\otimes\mathbb{1}|\omega\rangle$, $|\omega\rangle = \sum_i |ii\rangle$. Let the vector $\vec{p} = (p_i)$, the matrix $M = (|M_i\rangle)$, then $M|\rho\rangle = \vec{p}$, which can be solved with

$$|\rho\rangle = (M^t M)^{-1} M^t \vec{p} \tag{1.50}$$

provided M^tM is invertible, which can be chosen.

State tomography can be used for process tomography, represented by the process state χ. The idea is as follows. Choose an operator basis $\{E_n\}$ and a set of pure states $\{|\psi_j\rangle\}$ which is informational-complete. For d dimensional system, there are d^2 E_n and $|\psi_j\rangle$. A matrix B can be determined with $E_m|\psi_j\rangle\langle\psi_j|E_n^\dagger = \sum_k B_{mnjk}|\psi_k\rangle\langle\psi_k|$ and it shall have inverse. Any channel can be written as $\mathscr{E}(\rho) = \sum_{mn}\chi_{mn}E_m\rho E_n^\dagger$ for the process state χ. Now send each state $|\psi_j\rangle$ to a given channel to obtain $\mathscr{E}(|\psi_j\rangle) = \sum_k c_{jk}|\psi_k\rangle\langle\psi_k|$, and c_{jk} can be determined from state tomography. Then we can find $\chi_{mn} = \sum_{jk} B^{-1}_{mnjk}c_{jk}$. Tomography is expensive since the resource scales with d^4, which might be exponential with the system size. This is basically the same with state tomography of the Choi state $\mathscr{C}$.

Quantum estimation, or metrology, uses a dedicated (set of) observable to extract the value of unknown parameters. The measurement of the observable will disturb the system, so we expect the parameters cannot be exactly determined. This is shown by the quantum Cramer-Rao bound, proved via Schwarz inequality. For a state ρ_μ depending on an unknown parameter μ, define symmetric logarithmic derivatives L satisfying

$$\partial_\mu\rho_\mu = \{L,\rho_\mu\}/2. \tag{1.51}$$

Assume the derivative ∂_μ exists, and assume μ could be zero, denote $\rho_0 := \rho$. L 'drives' the dependence of ρ on μ. Define $\Omega(\cdot) = \{\rho,\cdot\}/2$, then $\Omega(L) = \partial_\mu\rho_\mu|_{\mu=0}$. Define an overlap between operators $(A,B)_\rho := \mathrm{tr}[A^\dagger\Omega(B)]$. Then $(A,A)_\rho = (\Delta A)^2$ as the standard derivation. Quantum Fisher information is defined as

$$F(\rho) := (L,L)_\rho, \tag{1.52}$$

and it is a property of both the state and an observable. The Cramer-Rao bound is

$$(\Delta A)^2 = (A,A)_\rho \geq \frac{1}{F(\rho)} \tag{1.53}$$

for A as 'locally unbiased estimator' $\partial_\mu\mathrm{tr}(\rho_\mu A)|_{\mu=0} = 1$. The bound is proved from $(A,A)_\rho(L,L)_\rho \geq |(A,L)_\rho|^2 = 1$ and $\mathrm{tr}(\rho A) = 0$. From the measured value of A, the unknown parameter μ can be calculated. The bound shows that μ can not be calculated exactly.

If N samples are used in parallel, then the bound is improved to be

$$(\Delta A)^2 \geq \frac{1}{\sqrt{N}F(\rho)}. \tag{1.54}$$

If entanglement is used among the samples, then the bound is improved to be

$$(\Delta A)^2 \geq \frac{1}{NF(\rho)}. \tag{1.55}$$

For instance, if $\rho_\mu = e^{-i\mu H}\rho e^{i\mu H}$ and ρ is pure, then $\Delta A \geq \frac{1}{4\Delta H}$, which is a special form of time-energy uncertainty relation. Note Fisher information is not a standard derivation of some observable, so quantum Cramer-Rao bound is not a special case of uncertainty relation, although both of them depends on Schwartz inequality. Both of them show a trade-off between the amount of information extracted, and the disturbance to the system.

For discrimination, we need to tell a target from a set of possible states by a measurement, and we wish to maximize the success probability. We cannot make sure for each event we can tell a target with certainty since the states are not orthogonal in general.

There are different strategies. You may maximize the average success probability, this is known as minimum-error scheme; you may want to tell each state with a corresponding event with certainty once detected, otherwise you say nothing, this is known as unambiguous scheme; you may want to maximize some other success probability, and in general this is called maximal confidence scheme. Depending on the objective function (i.e., success probability), the measurement scheme may differ significantly. This game is a manifest of the interference feature of quantum system.

1.6 The power of qubit

People want to understand the difference between quantum systems and classical ones. For instance, in what sense a qubit is better than a cbit or pbit? At least we can say there is coherence for qubit: states can be superposed and can interfer due to non-orthogonality, observable can be non-commuting, evolution can be unitary which evolves coherence. Computer scientists want to make the difference exact: what kinds of problems can quantum computers, which are collective states of many qubits, solve far more efficient than classical computers? We will analyze this later, and for now we will see several examples to illustrate the power of qubits, without entanglement.

1.6.1 Black-box encoding

There is a saying that a qubit cannot transmit more than a bit. This is based on the *Holevo bound*. Suppose Alice wants to transmit a random variable $X = \{x, p_x\}$, and she encodes $x \mapsto \rho_x$ and send the collection of ρ_x to Bob. To reveal X, Bob has to measure ρ_x and his measurement is modelled as a POVM $\{E_y\}$, which is free to choose. His outcome is modelled as a random variable $Y = \{y, p_y\}$ for $p_y = \mathrm{tr}(E_y\rho)$ with $\rho = \sum_x p_x\rho_x$. From state discrimination, we know Bob cannot distinguish non-orthogonal ρ_x perfectly. No matter what POVM he uses, the mutual information is upper bounded

$$I(X:Y) \leq \chi_H \leq H(X) \tag{1.56}$$

for $\chi_H := S(\rho) - \sum_x p_x S(\rho_x)$. For von Neumann entropy, $\chi_H = H(x)$ if the support of ρ_x are orthogonal. So it seems we shall encode x into orthogonal states. The mutual information $I(X:Y) = H(X) - H(Y|X)$ is smaller than $H(X)$ since $H(Y|X)$ is nonnegative. $H(Y|X)$ is zero when X and Y are completely correlated. Now we can see that if X is a bit of information, then Y at most is also a bit.

Question 25. *Shall we use clever encoding into quantum states?*

The answer shall be yes. For Holevo bound, the quantum states are treated as black boxes: the bits are encoded into the states themselves instead of their amplitudes. The input state is treated as a mixed state ρ without any detailed structures. Instead of treating quantum states as black boxes, we could treat them as 'white boxes' and encode information via the amplitudes. This applies to quantum computing, such as quantum simulation. For quantum simulation, a big advantage compared with classical simulation is that quantum system will be encoded as qubits which saves a lot of memory space. Product of unitary matrices can be simulated by applying unitary gates, while on classical computers, matrices multiplication are not very efficient.

1.6.2 Quantum key distribution

Quantum key distribution (QKD) is an example of better encoding, than the black box ones, that a key can be established among two parties, Alice and Bob. Any Eve who attempts to eavesdrop information will disturb the quantum states, and that can be detected. If the disturbance is above a threshold, then the key is not trustable, and they can try again another time, probably when Eve gets asleep.

Question 26. *How many qubit states are needed for QKD?*

In the famous BB84 protocol, two bases are used to encode a bit. 0 can be encoded as $|0\rangle$ or $|+\rangle$, 1 can be encoded as $|1\rangle$ or $|-\rangle$. So only 4 qubit states are being used, yet these states is an informational complete set. It is assumed that Alice and Bob have a classical authenticated channel to communicate. Also Alice will record the time of each state being send, i.e., these states are not mixed together. Alice generates a random bit, 0 or 1, then choose a basis from the two, then send the state. Alice then repeats the process.

Bob does not know the basis for each send. He choose a basis at random, and record his measurement results. After this, they communicate over the public classical channel. Only the bits that are of the same basis to them will be kept, which is half on average, leaving half the bits as garbage.

Eve, in the middle, could act like Bob. So it is easy to see the probability for Eve to know a bit is 1/4 (Both Eve and Bob need to be correct). The probability that Eve can be detected is 1/4. For n bits transmission, the probability that Eve can be detected is $1 - (\frac{3}{4})^n$. Furthermore, Eve could be far more better than this, and there

are all kinds of attacks she can do. But eventually, she will be detected and a secure key can still be established no matter how small the rate is.

1.6.3 Measure classical values

Here we show the power of qubits for the read of classical values. Suppose we need to measure the value of the integral of a classical field

$$I = \int \phi(x)dx, \tag{1.57}$$

e.g., a magnetic field or some flux value. We will see that quantum protocol can measure this value exponentially precise. The model is to use N bits or qubits, while there is no 'talk' among bits or qubits, or other resources. The bits or qubits interact with the field one at a time.

For the classical model, it allows a bit to flip depends on I. It assumes a Markovian process for each bit, and the probability for a flip is

$$p = 1 - e^{-\lambda I}, \tag{1.58}$$

for a certain adjustable parameter λ. Then the task is to estimate the value of p. For N samples, the precision is of the order $1/\sqrt{N}$.

For the quantum model, the interaction between the field and a qubit is unitary. As the field is 'big', the backaction on the field can be neglected. For the case $I = m\alpha$ for an integer m and a chosen value $\alpha \propto 2^{-N}$, m can be written as a binary number and each digit in it can be read and encoded in a qubit state. The basic method is as follows. All initial qubit state is $|0\rangle$. Choose an interaction H such that it rotate the qubit by π for $I = \alpha$. So if m is even, its state remains, and if m is odd, it becomes $|1\rangle$. The first qubit interact with H, and the nth interact with $\frac{1}{2^{n-1}}H$. The last digit of m is encoded in the first qubit, and in the next step the preceding digit is encoded etc. There are some details depending on the digit to be 0 or 1, but at the end the precision is in the order $1/2^N$.

Question 27. *Where does the quantum advantage come from?*

The primary reason for the quantum advantage is coherence: the quantum interaction can be unitary so interference can make some quantum 'leap'. The above model, with qubits or bit interact with an adversary (the field), while no direct interaction among themselves, is a simple model of one-way *Turing machine* (see Chapter 4). It is one-way since a bit or qubit cannot interact with the adversary twice. For a common Turing machine, a bit or qubit may interact with the adversary multiple times.

1.6.4 Leggett–Garg inequalities

When we say a system evolves, what does this mean? In Heisenberg picture, this means observable can change in time. An observable $A(t_1)$ at time t_1 may not commute with $A(t_2)$ at time t_2. So we can apply uncertainty relation in this setting and obtain some temporal inequalities. This is the subject of Leggett–Garg inequalities.

Question 28. *How to obtain information of a system at different times?*

Given a system which evolves, now we define n measurements at times t_l for $l = 1, 2, ...n$. We need correlations between two measurements. An immediate problem is that for any two t_i and t_j there shall be no other measurements. How to deal with this request? You may imagine there are many samples of the system, then we can perform these measurements at different times on different samples. The cost is that we have make sure these samples shall be identical. Another method is that, instead of using strong projective measurement, we can use weak measurement that only disturb the system slightly while obtain a fuzzy (or 'weak') value of observable (see Chapter 2). Both of these methods work, while the inequalities for weak measurement have to be derived separately, and it depends on the form of weak measurements.

We assume multiple samples and projective measurement, implemented indirectly, so the measurement will not introduce new ingredients. The Leggett–Garg quantity is

$$K_n = \sum_{l=2}^{n} C_{n(n-1)} - C_{n1}, \tag{1.59}$$

for two-time average correlation coefficient $C_{ij} = \sum_{r=1}^{N} Q_r(t_i)Q_r(t_j)/N$ for N trials, and $|Q(t)| \leq 1\ \forall t$. For instance, $K_3 = C_{21} + C_{32} - C_{31}$. The classical bound is $-n \leq K_n \leq n-2$ for odd $n \geq 3$, $-n+2 \leq K_n \leq n-2$ for even $n \geq 4$.

Quantum system will violate the Leggett–Garg bound. We need to redefine the correlation as

$$C_{ij} := \frac{1}{2}\mathrm{tr}\{Q(t_i), Q(t_j)\}. \tag{1.60}$$

For qubit, Q can be written as $Q(t_i) = \vec{a}_i \cdot \vec{\sigma}$, and then $C_{ij} = \vec{a}_i \cdot \vec{a}_j$. The Leggett–Garg quantity becomes

$$K_n = \sum_{l=1}^{n-1} \cos\theta_l - \cos(\sum_l \theta_l) \tag{1.61}$$

for $\theta_l = \arg(\vec{a}_l \cdot \vec{a}_{l+1})$. K_n is maximal when $\theta_l = \pi/n$, i.e., equal distance, and $\max K_n = n\cos\pi/n$, which is not bounded for $n \to \infty$.

Violations of the inequality is due to the non-commutativity of the operators. The commutator is $[Q_i, Q_j] = 2i\vec{\sigma} \cdot (\vec{a}_i \times \vec{a}_j)$, which is bigger when the angle between them is bigger. Each operator Q_i is described by a vector $\vec{a}_i$, and for the collection of them, the absolute average violation of the commutations is maximal when all the angles are the same and sum to π mod 2π.

1.6.5 Contextuality

In Leggett–Garg setting, the correlations of noncommuting operators are measured. This could be hard. There is a setting which only involves measurements of commuting operators, and this is the test for contextuality.

Question 29. *What is a context?*

Let us look at the famous Peres–Mermin square

$$\begin{pmatrix} A & B & C \\ a & b & c \\ \alpha & \beta & \gamma \end{pmatrix} = \begin{pmatrix} Z_1 & Z_2 & Z_1Z_2 \\ X_2 & X_1 & X_1X_2 \\ Z_1X_2 & X_1Z_2 & Y_1Y_2 \end{pmatrix} \tag{1.62}$$

and there are nine Pauli operators. The central fact is that, the observable within each row and column commute, hence form a 'context'. A context can be defined as the common eigenvectors (or projectors) of commuting observable. The eigenvalues of each operator can only be 1 or -1 for qubits. If you want to simulate this square classically by the assignment of a definite value 1 or -1 to each operator, you will find that it is not possible. This fact is coined 'quantum contextuality'. It is state-independent since no underlying state needs to be fixed. To quantify this feature, an inequality suffices. It is easy to see

$$\langle ABC\rangle + \langle abc\rangle + \langle \alpha\beta\gamma\rangle + \langle Aa\alpha\rangle + \langle Bb\beta\rangle - \langle Cc\gamma\rangle \leq 4 \tag{1.63}$$

for classical system, while it is 6 for any quantum state. State-dependent inequalities can also be defined, like Bell inequality (see Chapter 2).

The heuristic meaning of a context is just a classical setting in which everything can have definite values without conflict. For classical system, all observable commute, so there is only one context; or, we say, classical system is 'non-contextual'. For quantum system, as observable are operators, hence can form various contexts; so we say quantum system is contextual. In order to have consistent observation (of values) on a quantum system, we need a context. The observation in one context is not informationally complete; in order to do so, many contexts are necessary. For instance, for state tomography of a qubit, the projectors $|0\rangle\langle 0|$, $|1\rangle\langle 1|$, $|+\rangle\langle +|$, $|-\rangle\langle -|$ suffice. It have two contexts: the first two form a context, and the last two form the other one. These two contexts are used for QKD, by the way! The quantum power is due to the non-orthogonality of states, i.e., the existence of quantum coherence. It is the same to say it is due to contextuality. We also see that, as the Leggett–Garg inequality reveals, observable can be noncommuting due to evolution. As such, quantum dynamics is contextual (or coherent) and contextuality (or coherence) can be changed by the evolution.

These inequalities are useful for experimentalists to determine if their system has quantum features, at least. However, it does not guarantee their quantum systems are fully coherent with no classical ingredient, e.g., decoherence. The direct measure of

the coherence of a quantum system is by decoherence processes or measures, such as dephasing time, relaxation time etc. Those time scales are standard measures for the quality of a quantum system, such as a qubit used in quantum computers.

1.7 Zoo of qubits

Here we list various kinds of qubits people have studied, but this is far from a complete list, and there surely will be more.

Question 30. *What defines a good qubit?*

A qubit is just a two-level quantum system, but a good qubit is more than this. A good qubit shall be easy for control, to interact with other systems, have long coherence time, and support quantum gates on it and allow entangling operations among qubits, and also easy for measurements. With all these requirements, it becomes nontrivial to find good qubits.

Below we list popular ones, topological ones, and other qubits that relatively need more study. Only their names, encodings, decoherence, gates (and examples) are shown. Besides, there are qubits designed from error-correction codes, such as stabilizer codes that we will study later. Those are more software-oriented, while the qubits here are more physical or hardware-oriented.

Popular qubits:

- Superconducting charge qubit: (also transmon and xmon) even or odd number of Cooper pairs; charge fluctuation noise; gates from Hamiltonian dynamics
- Superconducting flux qubit: up or down flux through a loop; flux fluctuation noise; gates from Hamiltonian dynamics
- Superconducting phase qubit: lower energy levels of un-harmonic oscillator; thermal or phase noises; gates from Hamiltonian dynamics
- Trapped ion: lower energy electronic states; thermal noises; qubit gates from laser coupling; entangling gates from phonon and laser coupling
- Photon: polarization, angular momentum, path; photon counting noise; gates from linear optical devices such as beam splitter
- Cold atom: atomic states; thermal noises; gates from laser coupling
- Quantum dots: electron spin states; spin or charge fluctuation; qubit gates from external field coupling; entangling gates from coupling
- NV center or NMR: nuclear spin states; spin fluctuation; qubit gates from external magnetic field coupling; entangling gates from coupling

Topological qubits:

- TOP qubit: topological ground state degeneracy; high energy local noise; qubit gates by ('Wilson') loops of operators; entangling gates by measurement or coupling; e.g., toric code
- SPT/edge qubit: edge state degeneracy; high energy local noise; qubit gates by symmetry on bulk and edge; entangling gates by measurement or coupling; e.g., valence-bond solid edge qubits
- SPT/SSB qubit: ground state degeneracy from symmetry breaking; high energy local noise; qubit gates by SPT and SSB orders; entangling gates by measurement or coupling; e.g., valence-bond solid qubits
- SET qubit: topological ground state degeneracy; high energy local noise; qubit gates by ('Wilson') loops of operators or SPT orders; entangling gates by measurement or coupling; e.g., fractional quantum Hall liquid
- TOP/edge qubit: edge state degeneracy; high energy local noise; qubit gates by ('Wilson') lines of operators; entangling gates by measurement, coupling, or braiding; e.g., toric code with edge or hole
- TOP/anyon qubit: (this include SET/anyon qubit) fusion space of anyonic excitations; high energy local noise; qubit gates by braiding; entangling gates by braiding; e.g., Majorana zero-mode, Ising anyon

More exotic qubits:

- GKP qubit: two sectors of spectrum of a harmonic oscillator; shift of states; gates by operators of creation and annihilation operators
- Cat-code qubit: superposition of coherent states; photon number change or phase noise; gates by operators of creation and annihilation operators
- Breather qubit: (include SC Rhombus chain qubit, 0-π qubit) low spectrum of breather in sine-Gordon theory; high energy noise; gates by external control
- spin-cluster qubit: ground states of FM or AF spin chain; local spin noise; qubit gates by external field; entangling gates by exchange interaction coupling at the edges, or by spin-waves
- Magnon qubit: magnon excitation of FM or AF spin chain; local spin noise; qubit gates by external field; entangling gates by collision, spin-waves, or external field
- Andreev qubit: low spectrum of modes in quantum point-contact (QPC) defects in SC loop; high energy noise; gates by coupling to external control

There shall be more on the lists.

Chapter 2

Basic Formalism

In this chapter, we generalize the setting of qubit in the former chapter to generic cases. There are various ways to formalize quantum theory, such as $*$-algebra, operational, categoric, path-integral, information-theoretic approaches. These approaches, sometime in the name of 'interpretation', are not minimal due to a bag of addon that are not easy tools. Here we present a minimal approach of modern quantum theory, which is based on Hilbert spaces and operators on them.

2.1 Hilbert space

Let us start from the idea of space and operator. In math, space is nothing but a set. The space we use is known as linear space or vector space, which is a space with linear property. Here, 'vector' is not a good name. It does not mean a real vector or arrow, instead, it refers to property of linear superposition of elements in the space, just like linear superposition of 'vectors' we learn from Newtonian mechanics.

The Banach space $\mathfrak{B}$ is the complete normed vector space. By 'complete', it means all Cauchy sequences converge in it, where Cauchy sequence is a sequence whose elements become arbitrarily close to each other as the sequence progresses. A norm $||\cdot||$ is a function that assigns a strictly positive length or size to all elements in a space. For instance, the trace-norm $||\cdot||_1$ of an operator $A \in \mathfrak{B}$ is defined as $||A||_1 := \mathsf{tr}\sqrt{A^\dagger A}$. A linear map $T : \mathfrak{B} \to \mathfrak{B}$ is called a contraction iff $||T(A)|| \leq ||A||$, $\forall A \in \mathfrak{B}$. The resolvent operator of T is defined as

$$R_T(\lambda) := (\lambda \mathbb{1} - T)^{-1}, \lambda \in \mathbb{C}. \tag{2.1}$$

If $R_T(\lambda)$ does not exist, then $\lambda \in \sigma(T)$, with $\sigma(T)$ as the spectrum of T, λ is called eigenvalues of T; if $R_T(\lambda)$ does exist, then $\lambda \in \rho(T)$, with $\rho(T)$ as the resolvent set of T.

The linear operators involved in physics is mainly *normal* operator, with the property $TT^\dagger = T^\dagger T$, while there are exception, e.g., the creation and annihilation opera-

tor satisfying $[a, a^\dagger] = 1$, field operators, and Kraus operators. There are three basic types of normal operators in quantum mechanics:

- unitary operators U: $UU^\dagger = \mathbb{1}$, $U^\dagger = U^{-1}$;
- hermitian operators: $A = A^\dagger$;
- positive operators: hermitian operators ρ with all eigenvalues $\lambda(\rho) \geq 0$.

Observable is hermitian, state is positive, and evolution is essentially unitary. Note that the combination of the creation and annihilation operators are also observable.

Hermitian operators can be added together, unitary operators can be multiplied together. A unitary operator can be written as the exponent of a certain hermitian operator (for finite dimension cases). For infinite dimension case, the set of unitary operators as exponent of hermitian operators is dense in the set of all unitary operators.

Every Hilbert space is a Banach space. The converse is not always true; not every Banach space is a Hilbert space since Hilbert space requires a *distance*, not just a norm, to be defined. Furthermore, as quantum states are usually normalized, we shall reduce Hilbert space to projective Hilbert space by identifying vectors that differ up to a certain total non-zero complex number. Usually we just ignore the term 'projective'.

2.1.1 States

A d-dimentional pure quantum state $|\psi\rangle \in \mathscr{H}$ can be expressed as

$$|\psi\rangle = \sum_{i=0}^{d-1} c_i |i\rangle, \tag{2.2}$$

for $\{|i\rangle\}$ a basis of the Hilbert space $\mathscr{H}$, $d \in \mathbb{Z}^+$, $d \geq 2$, and the normalization condition $\sum_{i=0}^{d-1} |c_i|^2 = 1$. The distance between two pure states $|\psi\rangle$ and $|\varphi\rangle$ is not measured uniquely. A common one is the fidelity $F = |\langle\varphi|\psi\rangle|$, also known as 'overlap' or 'inner product'.

Question 31. *How many parameters there are in a pure state?*

There are $2d$ real parameters for the d coefficients, while the normalization reduces one. Yet there is another condition: the global phase of a state is meaningless, so the total number of d.o.f of a pure state is $2d-2$.

More general quantum state is known as density operator ρ, also called mixed state, which is a semidefinite positive trace class operator acting on a Hilbert space $\mathscr{H}$, and $\rho \geq 0$, $\mathrm{tr}\rho = 1$. Denote the set of linear operators acting on $\mathscr{H}$ as $\mathscr{L}(\mathscr{H})$, and the set of density operators acting on $\mathscr{H}$ as $\mathscr{D}(\mathscr{H})$, and clearly $\mathscr{D}(\mathscr{H}) \subset$

$\mathscr{L}(\mathscr{H})$. Sometimes we use the shorthand $\mathscr{D}$ for $\mathscr{D}(\mathscr{H})$, also we use $\mathscr{D}_d$ to indicate that the underlying Hilbert space is of dimension d. The space $\mathscr{L}(\mathscr{H})$ is also a Hilbert space with respect to Hilbert-Schmidt inner product

$$\mathsf{tr}(A^\dagger B),\ \forall A,B \in \mathscr{L}(\mathscr{H}). \tag{2.3}$$

A basis for $\mathscr{L}(\mathscr{H})$ is a linearly independent spanning set, denoted as $\{M_\alpha\}_{\alpha=0}^{d^2-1}$, which is an operator basis instead of vector basis. An orthogonal basis is a basis with $\mathsf{tr}(M_\alpha^\dagger M_\beta) = 0$ for $\alpha \neq \beta$. Furthermore, every finite-dimensional inner product space has an *orthonormal* basis satisfying

$$\mathsf{tr}(M_\alpha^\dagger M_\beta) = d\delta_{\alpha\beta}. \tag{2.4}$$

Note our definition of an orthonormal basis involves the coefficient d, which may be absent following other conventions.

For a clean representation of density operator $\rho \in \mathscr{D}$, there exists such a basis satisfying (i) $M_0 = \mathbb{1}$, (ii) $\mathsf{tr}M_\alpha = 0$ for $\alpha \neq 0$, (iii) $\mathsf{tr}(M_\alpha^\dagger M_\beta) = 0$ for $\alpha \neq \beta$, which is called a trace-free and trace-orthogonal basis, and, for simplicity, termed as *canonical* basis here. In a canonical basis, a density operator ρ can be written as

$$\rho = \frac{1}{d}\left(\mathbb{1} + \sum_{\alpha=1}^{d^2-1}\sqrt{\frac{d(d-1)}{\mathsf{tr}(M_\alpha^\dagger M_\alpha)}}p_\alpha M_\alpha\right). \tag{2.5}$$

The parameters p_α form the Bloch *polarization vector* $\boldsymbol{p} := (p_1,\ldots,p_{d^2-1})$ with $\|\boldsymbol{p}\| = 1$ for pure state and $\|\boldsymbol{p}\| < 1$ for mixed state. If each M_α is hermitian, $p_\alpha \in \mathbb{R}$. It is clear to see that a mixed state contains d^2-1 real parameters.

We can see that a canonical basis may be orthonormal or not. The Kronecker basis $\{|i\rangle\langle j|\}$ is orthonormal yet not canonical. Here $|i\rangle\langle j|$ is a matrix with a single entry 1 in ith row and jth column, and all others 0. The merit of this basis is that it is related to the Choi-Jamiołkowski isomorphism $\mathscr{J} : \mathscr{L}(\mathscr{H}) \to \mathscr{H}\otimes\mathscr{H}$, from which $\mathscr{J} : |i\rangle\langle j| \mapsto |i,j\rangle$. The action of this isomorphism on matrices is the same with the reshaping operation, which is defined as

$$\mathrm{res}A := (a_{11},\ldots,a_{1m},\ldots,a_{m1},\ldots,a_{mm})^{\mathrm{t}} \tag{2.6}$$

for an $m\times m$ matrix $A = [a_{ij}]$ with elements a_{ij}. With the inverse of reshaping operation, the matrix A can be obtained from the vector res$A$, i.e., $\mathrm{res}^{-1}(\mathrm{res}A) = A$. In addition, a vectorization operation can be defined such that $\mathrm{vec}A = \mathrm{res}A^{\mathrm{t}}$.

For a qubit, an example of orthonormal as well as canonical basis is the Pauli basis $\{\sigma^i\} = \{\mathbb{1},X,Y,Z\}$ $(i = 0,1,2,3)$. A qubit state $\rho = \frac{1}{2}(\mathbb{1} + \boldsymbol{p}\cdot\boldsymbol{\sigma})$ is represented by a real polarization vector $\boldsymbol{p}$ for $\boldsymbol{\sigma} := (X,Y,Z)$. In Kronecker basis $\{\tau_j\} = \{|0\rangle\langle 0|,|0\rangle\langle 1|,|1\rangle\langle 0|,|1\rangle\langle 1|\}$ $(j = 0,1,2,3)$ a qubit state ρ is represented as a vector

$$\rho \mapsto \mathrm{res}\rho = (\rho_{00},\rho_{01},\rho_{10},\rho_{11})^t. \tag{2.7}$$

For higher-dimensional cases the Kronecker basis carries over easily, while there are different generalizations of Pauli basis. The Pauli basis is both hermitian and unitary, i.e. self-invertible. There are two well-known bases that are both canonical and orthonormal, one is Gell-Mann basis, which is hermitian, and the other is Heisenberg-Weyl basis, which is unitary.

Let E_{jk} denote $|i\rangle\langle j|$, then the generalized Gell-Mann matrices for $SU(n)$ are defined as

$$X_{ij} = \frac{1}{2}(E_{ij} + E_{ji}),\ 1 \le i < j \le n, \tag{2.8a}$$

$$Y_{ij} = \frac{-i}{2}(E_{ij} - E_{ji}),\ 1 \le i < j \le n, \tag{2.8b}$$

$$Z_j = \frac{1}{\sqrt{2j(j-1)}}\left(\sum_{\ell=1}^{j-1} E_{\ell\ell} - (j-1)E_{jj}\right),\ 2 \le j \le n. \tag{2.8c}$$

These traceless matrices are usually denoted as $\{\lambda^i\}$, or $\{t^i = \lambda^i/2\}$, with $\mathrm{tr}(t^i t^j) = \frac{1}{2}\delta_{ij}$. They satisfy

$$[t^i, t^j] = i f_{ijk} t^k \tag{2.9}$$

$$\{t^i, t^j\} = \frac{1}{N}\delta_{ij} + d_{ijk} t^k \tag{2.10}$$

for $f_{ijk} = -2i\mathrm{tr}([t^i,t^j]t^k)$, $d_{ijk} = 2\mathrm{tr}(\{t^i,t^j\}t^k)$ as anti-symmetric and symmetric structure constants, respectively. Einstein summation rule is assumed. It also hold $t^i t^j = \frac{1}{2N}\delta_{ij}\mathbb{1} + \frac{1}{2}h_{ijk}t^k$, for $h_{ijk} = d_{ijk} + i f_{ijk}$, and $t^s t^a t^s = -\frac{1}{2N}t^a$. For $SU(2)$, λ^i are Pauli matrix, f_{ijk} is ε_{ijk}, $d_{ijk} = 0$, and

$$\varepsilon_{ijk}\varepsilon_{imn} = \delta_{mj}\delta_{kn} - \delta_{nj}\delta_{km}, \tag{2.11}$$

$$\varepsilon_{ijk}\varepsilon_{ijn} = 2\delta_{kn}. \tag{2.12}$$

The Heisenberg-Weyl basis $\{M_{jk}\}$ for a qudit system is specified by

$$X_j = \sum_{i=0}^{d-1} |i\rangle\langle i+j|, \quad Z_k = \sum_{l=0}^{d-1} \omega^{lk}|l\rangle\langle l| \pmod d, \tag{2.13}$$

for $M_{jk} = X_j Z_k$, and $\omega = e^{i2\pi/d}$, $j,k \in \{0,\ldots,d-1\}$.

The two operators X_1 and Z_1 are the generators for the so-called Heisenberg-Weyl group $\mathscr{G}_{\mathrm{HW}}$ containing group element $\{\omega^{i-jk}M_{jk}\}$ with degree d and order d^3. The center of $\mathscr{G}_{\mathrm{HW}}$ is $\{\omega^{i-jk}\mathbb{1}\}$. The two generators do not commute

$$X_1 Z_1 = \omega Z_1 X_1, \quad X_1^d = \mathbb{1}, \quad Z_1^d = \mathbb{1}. \tag{2.14}$$

The X_1 is sometimes called the 'shift' operator, and Z_1 is called the 'clock' operator, and they are related by a Hadamard operator W, understood as discrete Fourier

transform

$$W := \frac{1}{\sqrt{d}} \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega^{d-1} & \omega^{2(d-1)} & \cdots & \omega^{(d-1)^2} \\ 1 & \omega^{d-2} & \omega^{2(d-2)} & \cdots & \omega^{(d-1)(d-2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega & \omega^2 & \cdots & \omega^{d-1} \end{pmatrix}, \tag{2.15}$$

and $X_1 = WZ_1W^\dagger$. The eigenvectors of Z_1 are the standard computational basis $|l\rangle$ with eigenvalues ω^l such that $Z_1|l\rangle = \omega^l|l\rangle$. The eigenvectors of X_1 are $W|l\rangle$ with eigenvalues ω^l such that $X_1W|l\rangle = \omega^l W|l\rangle$. As $|\langle l|W|l'\rangle| = 1/d$, their eigenbasis of X_1 and Z_1 are *mutually unbiased*.

The Heisenberg-Weyl basis $\{M_{jk}\}$ is orthonormal

$$\mathsf{tr}\left(M_{jk}^\dagger M_{j'k'}\right) = d\delta_{jj'}\delta_{kk'}. \tag{2.16}$$

The following commutation relations hold

$$X_jZ_k = Z_kX_j\omega^{jk}, \tag{2.17}$$

$$M_{jk}M_{j'k'} = M_{j'k'}M_{jk}\omega^{jk'-kj'}, \tag{2.18}$$

$$M_{jk}M_{j'k'} = M_{j+j',k+k'}\omega^{-kj'}. \tag{2.19}$$

In this basis, a state is represented as

$$\rho = \frac{1}{d}\left(\mathbb{1} + \sqrt{d-1}\sum_{\alpha=1}^{d^2-1} p_\alpha M_\alpha\right), \tag{2.20}$$

with $M_\alpha \equiv M_{jk}$, $p_\alpha \equiv p_{jk}$. $p_{jk}^* = p_{-j-k}\omega^{-jk} \in \mathbb{C}$.

Question 32. *Pure vs. mixed state, which is more fundamental?*

We see that a mixed state has one order more parameters than a pure state. How this occurs? First, we can use the eigenvalue decomposition

$$\rho = \sum_i p_i|\psi_i\rangle\langle\psi_i| \tag{2.21}$$

for $\sum_i p_i = 1$, and $\langle\psi_i|\psi_j\rangle = \delta_{ij}$. We find the parameter counting agrees:

$$d^2 - 1 = d - 1 + d(2d-2) - d(d-1). \tag{2.22}$$

On the right-hand side, $d-1$ is for the probabilities p_i, $d(2d-2)$ is the number of real parameters in the eigenstates, $d(d-1)$ is the number of (real) orthogonality constraints.

We could count parameters in a different way. A rank-r d-dimensional hermitian matrix has $2dr - r^2$ real parameters. A full rank d-dimensional mixed state will have d^2-1 real parameters, while the minus one is due to the trace condition. A pure state is a rank-1 mixed state, so it has $2d-2$ parameters, agree with our former result.

2.1.2 Norm

We consider bounded linear operators in $\mathscr{L}(\mathscr{H})$. A family of norms is known as the *Schatten p-norm*

$$\|T\|_p := \left[\mathsf{tr}\left(\left(T^\dagger T\right)^{p/2}\right)\right]^{1/p}, \quad p \geq 1,\ \forall T \in \mathscr{L}(\mathscr{H}). \tag{2.23}$$

This norm includes some commonly used norms as special cases.

- Trace norm ($p = 1$): $\|T\|_{\mathsf{tr}} := \|T\|_1 = \mathsf{tr}\sqrt{T^\dagger T}$.
- Operator (or spectral) norm ($p = \infty$): $\|T\| = \sigma_{\max}(T)$, $\sigma_{\max}$ denotes the largest singular value. It can also be defined as $\|T\| := \sup_{|\psi\rangle} \|T|\psi\rangle\|, \forall |\psi\rangle \in \mathscr{H}$, and $\||\psi\rangle\| = 1$.
- Hilbert-Schmidt (or Frobenius) norm ($p = 2$): $\|T\|_F := \|T\|_2 = \sqrt{\mathsf{tr}(T^\dagger T)}$.

The Schatten p-norm has many useful properties.

- Positive semidefiniteness: $\|T\|_p \geq 0$ with $\|T\|_p = 0$ if $T = 0$.
- Positive scalability: $\|\alpha T\|_p = |\alpha| \cdot \|T\|_p$ for $\alpha \in \mathbb{C}$.
- The triangle inequality: $\|T_1 + T_2\|_p \leq \|T_1\|_p + \|T_2\|_p$.
- Unitary invariance: $\|T\|_p = \|UTV\|_p$ for unitary operator $U, V \in \mathscr{L}(\mathscr{H})$.
- Decreasing in p: $\|T\|_p \geq \|T\|_q$ for $1 \leq p \leq q \leq \infty$.
- Sub-multiplicative under composition: $\|T_1 T_2\|_p \leq \|T_1\|_p \|T_2\|_p$.
- Hölder's inequality: For $p, q, r \geq 1$, $1/p + 1/q \leq 1/r$, then $\|T_1 T_2\|_r \leq \|T_1\|_p \|T_2\|_q$.

For operator norms, there exists one crucial property that any two norms $\|\cdot\|_p$ and $\|\cdot\|_q$ are equivalent iff

$$r\|\cdot\|_p \leq \|\cdot\|_q \leq s\|\cdot\|_p,\ \exists r, s \in \mathbb{R}^+. \tag{2.24}$$

The following inequalities are important for quantum computation tasks

$$\|T\|_F \leq \|T\|_1 \leq \sqrt{r}\|T\|_F,\ \|T\| \leq \|T\|_F \leq \sqrt{r}\|T\|, \quad r := \mathrm{rank} T. \tag{2.25}$$

As the case of operator norm, norm for 'superoperator', which acts on operators, can also be well defined. Based on the Schatten p-norm, the *induced Schatten ($q \to p$)-norm* is defined as

$$\|\Pi\|_{q\to p} := \max_T \frac{\|\Pi T\|_p}{\|T\|_q}, \quad p, q \geq 1,\ \forall T \in \mathscr{L}(\mathscr{H}),\ \forall \Pi \in \mathscr{L}(\mathscr{L}(\mathscr{H})). \tag{2.26}$$

However, one unpleasant property of this norm is that for $1 \leq p < 2$, $\|\Pi\|_{1\to p}$ is not stable, i.e., $\|\Pi\|_{1\to p} \neq \|\Pi \otimes \mathbb{1}\|_{1\to p}$ with $\mathbb{1}$ acting on another space. For the case $p = q = 1$, a stabilized norm, or called completely-bounded norm, known as diamond norm, is defined as

$$\|\Pi\|_\diamond := \|\Pi \otimes \mathbb{1}\|_{1\to 1} = \max_T \|(\Pi \otimes \mathbb{1})T\|_{\mathsf{tr}}, \tag{2.27}$$

for $\mathbb{1} \in \mathscr{L}(\mathscr{L}(\mathscr{K}))$, $T \in \mathscr{L}(\mathscr{H} \otimes \mathscr{K})$, and $\dim(\mathscr{K}) \geq \dim(\mathscr{H})$. The main properties of the diamond norm include:

- Sub-multiplicative under composition: $\|\Pi_1\Pi_2\|_\diamond \leq \|\Pi_1\|_\diamond \|\Pi_2\|_\diamond$.
- Multiplicative under tensor product: $\|\Pi_1 \otimes \Pi_2\|_\diamond = \|\Pi_1\|_\diamond \|\Pi_2\|_\diamond$.
- Chain property: $\|\Pi_1\Pi_2 - \Pi_1'\Pi_2'\|_\diamond \leq \|\Pi_1 - \Pi_1'\|_\diamond + \|\Pi_2 - \Pi_2'\|_\diamond$.
- Unitary invariance: $\|\Pi\|_\diamond = \|\mathscr{U}\Pi\mathscr{V}\|_\diamond$ for unitary operators $\mathscr{U}, \mathscr{V} \in \mathscr{L}(\mathscr{L}(\mathscr{H}))$.

One important application of diamond norm is to quantify the distance between quantum channels. In this case, the domain over which the optimization is taken is $\mathscr{D}(\mathscr{H})$. Furthermore, the maximum is achieved for a pure state following from the convexity of $\mathscr{D}(\mathscr{H})$.

2.1.3 Dynamics

Given a Hamiltonian H, the unitary evolution takes the form

$$U = e^{itH}, \tag{2.28}$$

but this only holds for time-independent Hamiltonian. When $H(t)$ depends on time t, we know that U can not be written in this form. Instead, we use the propagator $U(t,s)$ with

$$\begin{aligned} U(t,s) &:= \mathscr{T} e^{\int_s^t H(\tau)d\tau} \\ &= \mathbb{1} + \sum_{m=1}^{\infty} \int_s^t \int_s^{t_1} \cdots \int_s^{t_{m-1}} H(t_1)H(t_2)\cdots H(t_m)dt_{m\to 1}. \end{aligned} \tag{2.29}$$

This is known as Dyson expansion. This can also be formally generalized to the non-unitary case by replacing H with bounded operator L depending on time. However, this expansion is extremely difficult to deal with, and it causes lots of problem in particle physics.

Even when there is no time-dependence, the unitary U can be hard to solve when H contains several non-commuting terms. First of all, the exponential of an operator A is defined as

$$e^A := \sum_{n=0}^{\infty} \frac{A^n}{n!}. \tag{2.30}$$

If two operators L_1 and L_2 satisfy $[L_1, L_2] \neq 0$, then

$$e^{L_1+L_2} = \lim_{n\to\infty} (e^{L_1/n} e^{L_2/n})^n, \tag{2.31}$$

which can be generalized as $e^{\sum_k L_k} = \lim_{n\to\infty} (\prod_k e^{L_k/n})^n$. This formula is known as the Lie-Trotter product formula, and plays a central role in quantum simulation of Hamiltonian dynamics. The Trotter formula results error $O(t^2)$.

Smaller error can be achieved by the Suzuki formula, which is a symmetric and systematic iterative form and can reduce the error to $O(t^{1+2\chi})$, χ can be arbitrary. However, when χ is arbitrarily big, there will be e^{χ} number of gates in Suzuki sequence, so one optimal value of χ can be chosen to make the number of gates optimally small.

To simulate U for $H = \sum_{\lambda=1}^{\Lambda} H_\lambda$ by Trotter-Suzuki formula $\tilde{U} = [U_\chi(\tau)]^r$, with $t = r\tau$ as evolution time, and $s := 1/(4 - 4^{1/(2p-1)})$, $1 < p \leq \chi$, define

$$U_1(\tau) = \prod_{\lambda=1}^{\Lambda} u_\lambda(\tau/2) \prod_{\lambda=\Lambda}^{1} u_\lambda(\tau/2), \tag{2.32}$$

$$U_p(\tau) = \left[U_{p-1}(s\tau)\right]^2 U_{p-1}((1-4s)\tau) \left[U_{p-1}(s\tau)\right]^2 \tag{2.33}$$

for $u_\lambda(\tau/2) := e^{-iH_\lambda \tau/2}$. The operator-norm distance is

$$\|U - \tilde{U}\| \in O\left(\frac{t^{2\chi+1}}{r^{2\chi}}\right). \tag{2.34}$$

For Hamiltonian evolution simulation, it requires $\|U - \tilde{U}\| \leq \varepsilon$. So values of r and χ need to be chosen. The value of r is firstly fixed to make the error smaller than the bound ε, then χ is fixed to make the number of gates optimal. Denote $h \equiv \|H\|$, we have $r = \left\lceil \frac{2\Lambda\chi(5/3)^{\chi-1}(ht)^{1+1/2\chi}}{\varepsilon^{1/2\chi}} \right\rceil$, and $\chi = \left\lceil \sqrt{\frac{\log_{25/3}(\Lambda ht/\varepsilon)}{2}} \right\rceil$. The number of gate in Trotter-Suzuki sequence is

$$N \in O\left(\Lambda^{2+1/2\chi} (ht)^{1+1/2\chi} / \varepsilon^{1/2\chi}\right). \tag{2.35}$$

However, the scaling is not efficient with respect to $\frac{1}{\varepsilon}$. There are methods to reduce it to $\log \frac{1}{\varepsilon}$.

To see the dynamical effect, as we have discussed in Chapter 1, there are Heisenberg and Schrödinger pictures. The Heisenberg picture has closer relation with classical mechanics since it describes the dynamics

$$i\hat{A}(t) = [A(t), H] \tag{2.36}$$

of observable $A(t)$. The Ehrenfest theorem states that the classical observable is the expectation value of an observable $A(t)$ on the underlying state $|\psi\rangle$ or ρ, which does

not evolve in the Heisenberg picture. For the unitary evolution, $A(t) = U^\dagger A(0) U$ for $U = e^{-itH}$. The expectation value $\mathsf{tr}(\rho A(t)) = \mathsf{tr}(\rho(t) A(0))$ as if the state evolves $\rho(t) = U \rho U^\dagger$.

Note 'picture' is like a choice of reference frame in classical mechanics; e.g., we are free to choose whether the earth moves around the sun, or vice versa. By comparison, a basis is the analog of a coordinate of a space. There is also another picture called 'interaction picture' that applies to cases when H contains special interaction terms between systems. The study of this can be found in many places, so we will move on to the next topic.

2.1.4 Measurement

In order to detect what's going on in a quantum evolution, we have to perform measurements. Intuitively, measurement will convert quantum operators to numbers. This is true and described as positive operator-valued measure (POVM), which is a set $\{E_i\}$ for $\sum_i E_i = \mathbb{1}$, $p_i = \mathsf{tr}(E_i \rho) \geq 0$, and $\sum_i p_i = 1$. The final state after measurement is ignored.

Question 33. *What is the goal of measurement?*

We find there are at least three of them:

- to prepare states;
- to verify a known, partially known, or unknown process, such as tomography, estimation, metrology;
- to extract values of observable as an answer to a question.

There are many classifications of measurements. Here is a binary one: sharp and non-sharp ones. A *sharp* measurement is a set of projectors, described as projector-valued measure (PVM), usually known as *direct* measurement (but can still be realized indirectly). We have studied this in Chapter 1. A *non-sharp* measurement is a POVM, also known as *indirect* measurement. Some direct measurement indeed destroy the final state. While indirect measurement will keep the final state.

A POVM can be realized by a PVM on a larger space from dilation theorems. Say, couple a system to an ancilla, and PVM on the ancilla realizes POVM on the system. Each effect $E_i = K_i^\dagger K_i$ for a Kraus operator K_i from a channel, then usually the final states are kept.

There are many kinds of POVM. A slight generalization of PVM is the IC-POVM, with IC means informational complete, which contains rank-one projectors that are not orthogonal. It is often used for state tomography. Also SIC-POVM is IC-POVM with symmetry property on the overlaps of projectors.

Question 34. *Can a quantum measurement be close to identity?*

The answer is yes, and the disturbance is truly small. Here we study the weak measurement, which is a type of POVM that is close to identity. In terms of Kraus operators, each of them is close to identity or zero operator. Weak measurements are interesting and important, so here we analyze it in details.

Below is a usual model for weak measurement induced by a Hamiltonian. The system s and apparatus a starts from product state $|\psi_s\rangle|\psi_a\rangle$, and interact by $H = \nu O_s \otimes O_a$. In general the interaction can be $H = \sum_i \nu_i O_{s;i} \otimes O_{a;i}$, but for simplicity we only need one term. The interaction strength ν is small compared with some intrinsic feature of the system. For $U = e^{-itH}$, the final state is

$$U|\psi_s\rangle|\psi_a\rangle \approx (\mathbb{1} - it\nu O_s \otimes O_a)|\psi_s\rangle|\psi_a\rangle. \tag{2.37}$$

If we trace out s, then the apparatus state approximates $e^{-it\nu\alpha O_a}|\psi_a\rangle$ for $\alpha = \langle\psi_s|O_s|\psi_s\rangle$. If we trace out a, then the system state approximates $e^{-it\nu\beta O_s}|\psi_s\rangle$ for $\beta = \langle\psi_a|O_a|\psi_a\rangle$.

We want to see how to approximate U by a simpler one. The values of α and β may be big, so reducing U to a smaller one on the system or apparatus is not good. It turns out the better approximation is a multiplexer, i.e., a controlled operation.

If we intend to measure the system in an orthonormal basis $\{|t_s\rangle\}$ such that $\sum_t |t_s\rangle\langle t_s| = \mathbb{1}$, then

$$U|\psi_s\rangle|\psi_a\rangle \approx \sum_t \langle t_s|\psi_s\rangle|t_s\rangle e^{-it\nu\varpi_s O_a}|\psi_a\rangle. \tag{2.38}$$

This generates a bit of entanglement. This means that U can be written as

$$U \approx \sum_t |t_s\rangle\langle t_s| \otimes e^{-it\nu\varpi_s O_a}. \tag{2.39}$$

Note that this form depends on the initial state! Here ϖ_s is a so-called 'weak value'

$$\varpi_s := \frac{\langle t_s|O_s|\psi_s\rangle}{\langle t_s|\psi_s\rangle}, \tag{2.40}$$

given that $\langle t_s|\psi_s\rangle \neq 0$. Similarly, if we intend to measure the apparatus in an orthonormal basis $\{|q_a\rangle\}$, we get $U \approx \sum_q |q_a\rangle\langle q_a| \otimes e^{-it\nu\varpi_a O_s}$ with $\varpi_a := \frac{\langle q_a|O_a|\psi_a\rangle}{\langle q_a|\psi_a\rangle}$, given that $\langle q_a|\psi_a\rangle \neq 0$. The approximate form of U is block-diagonal. In general, i.e., not weak interaction, U can be expanded with sine-cosine decomposition and we know that U is not block-diagonal. Only when the interaction is weak so that the sine (cosine) terms are close to zero (one), U is approximately block-diagonal.

We see that if the apparatus is projected onto $|q_a\rangle$ with success probability approximately $|\langle q_a|\psi_a\rangle|^2$, the system state is $e^{-it\nu\varpi_a O_s}|\psi_s\rangle$. The Kraus operators are $K_q = \langle q_a|\psi_a\rangle e^{-it\nu\varpi_a O_s}$ forming a quantum channel. If the system is projected onto $|t_s\rangle$ with success probability approximately $|\langle t_s|\psi_s\rangle|^2$, the apparatus state is $e^{-it\nu\varpi_s O_a}|\psi_a\rangle$. The Kraus operators are $K'_t = \langle t_s|\psi_s\rangle e^{-it\nu\varpi_s O_a}$ forming a complementary quantum channel. Note that ϖ_s or ϖ_a may not be a real number! So each Kraus operator is only approximately unitary. See Fig. 2.1.

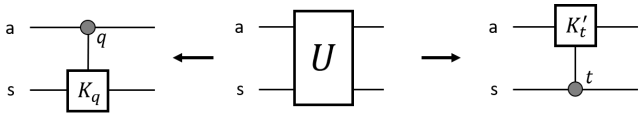

Figure 2.1: The model of weak measurement and weak value. The gray dots are the control operations in the multiplexers.

The above analysis generalizes to mixed states. For initial state $\rho_s \otimes \rho_a$, after the interaction if s is traced out, then the apparatus state approximates $e^{-it\nu\alpha O_a}\rho_a e^{it\nu\alpha O_a}$ for $\alpha = \mathsf{tr}(O_s\rho_s)$. If a is traced out, then the system state approximates $e^{-it\nu\beta O_s}\rho_s e^{it\nu\beta O_s}$ for $\beta = \mathsf{tr}(O_a\rho_a)$.

In full, the final state takes the form

$$\rho_f = \sum_{qq'} |q_a\rangle\langle q'_a|\langle q_a|\rho_a|q'_a\rangle e^{-it\nu\varpi_a^L O_s}\rho_s e^{it\nu\varpi_a^R O_s} \tag{2.41}$$

for

$$\varpi_a^L := \frac{\langle q_a|O_a\rho_a|q'_a\rangle}{\langle q_a|\rho_a|q'_a\rangle},\ \varpi_a^R := \frac{\langle q_a|\rho_a O_a|q'_a\rangle}{\langle q_a|\rho_a|q'_a\rangle}, \tag{2.42}$$

and $\varpi_a^L \neq \varpi_a^R \in \mathbb{C}$ in general. This means the U is approximated by two controlled-operations U_L and U_R such that

$$U(\rho_s \otimes \rho_a)U^\dagger \approx U_L(\rho_s \otimes \rho_a)U_R. \tag{2.43}$$

If the apparatus is projected onto $|q_a\rangle$ with success prob $\langle q_a|\rho_a|q_a\rangle$, then the system state is $e^{-it\nu\varpi_a O_s}\rho_s e^{it\nu\varpi_a^* O_s}$ for $\varpi_a = \frac{\langle q_a|O_a\rho_a|q_a\rangle}{\langle q_a|\rho_a|q_a\rangle}$. This reduces to the pure state form when ρ_a is pure.

The above can be generalized to mixed post-selected states. Note that projector has a special property: the state after the projection is specified by the projector itself. General Kraus operator does not have such a property as it leads to $K\rho K^\dagger$ which depends on the input state ρ. The projective measurement can be viewed as a measure-prepare procedure: first measure a state by a projector, then prepare the projected state. This could be generalized to mixed state case and this is the so-called 'entanglement-breaking' channel $\mathscr{E}(\rho) = \sum_k \rho_k \mathsf{tr}(\rho F_k)$ with a set of states $\{\rho_k\}$ and a POVM $\{F_k\}$. The Kraus operators take the form $\sqrt{\rho_k}|m\rangle\langle n|\sqrt{F_k}$ for orthonormal basis $\{|m\rangle\}$ and $\{|n\rangle\}$. With it on the system, the final state takes the form

$$\rho_f = \sum_k \mathsf{tr}(\rho_s F_k)\rho_k e^{-it\nu\varpi_s^L O_a}\rho_a e^{it\nu\varpi_s^R O_a} \tag{2.44}$$

for

$$\varpi_s^L := \frac{\mathsf{tr}(F_k O_s\rho_s)}{\mathsf{tr}(F_k\rho_s)},\quad \varpi_a^R := \frac{\mathsf{tr}(F_k\rho_s O_s)}{\mathsf{tr}(F_k\rho_s)}, \tag{2.45}$$

which is a general form of weak value.

2.2 Quantum notions

The weak value above is an example of 'quantum notions', which has proven to be powerful or useful. By 'quantum notions', we mean measures on quantum states and/or observable that quantify the amount of certain 'resources' which can be proven to be powerful. There are three classes of notions:

- only on state: entanglement, entropy, etc
- only on observable: state-independent contextuality, etc
- both on state and observable: uncertainty (include state-dependent contextuality), weak value, etc

2.2.1 Information and entropy

Question 35. *Are information and entropy the same?*

Information is a special kind of observable that it is only a function of the state itself. General observable are defined quite independent of what a state is. In physics, information is quantified via 'entropy'. Entropy plays important roles in many areas, such as thermodynamics and communication. Entropy of a state is the *minimal* amount of information required to reproduce a state, given some promises. The promise is a 'translation book'; e.g., the bits 'encode' an apple, or a thermal state. The promise specifies the context and assigns physical or practical meanings, otherwise the information (entropy) is just one abstract description. Entropy is the 'lack of information'. Entropy describes the complexity of a state. The more disorder it is, the bigger the entropy is.

Similar with the role of partition for entanglement, entropy should be defined so that it can be measured, i.e., in operational way. We will see that the von Neumann entropy is a lower bound of it.

A state is labelled as ρ, which can be determined by tomography. If partial information is promised of ρ, then a full tomography is not necessary and other methods such as estimation can be used. In physics, there usually is a context for ρ, such as a preparation (or measurement) scheme, a coupling to other system by Hamiltonian, i.e., ρ is not a black box. To convert an operator to a classical value, there must be an operational process or a context. Such a classical value is said to be 'operational' or 'contextual'. Entropy shall be contextual: its definition needs a preparation (or measurement) scheme. In classical mechanics, measurement is not very significant, while in quantum theory measurement disturbs the target nontrivially.

The minimal axioms for an information measure are:

1. non-negative (lower bounded) $H(\rho) \geq 0$.
2. additive for product states $H(\rho \otimes \sigma) = H(\rho) + H(\sigma)$.

Note that it does not need to be upper bounded. For infinite-dimensional system, a measure may not be upper bounded.

There are many information measures satisfying the above properties, e.g., Shannon entropy, von Neumann entropy, Renyi entropy, etc. There are also measures violating the above properties, in particular, the additivity, and one such example is the Tsallis entropy.

Shannon entropy $H(p_i)$ of probability distribution $\{p_i\}$ is

$$H(p_i) = -\sum_i p_i \log p_i. \tag{2.46}$$

The von Neumann entropy, denoted uniquely by

$$S(\rho) = -\mathrm{tr}(\rho \log \rho) = -\sum_i \lambda_i \log \lambda_i, \tag{2.47}$$

is the Shannon entropy in the eigenbasis $\{|i\rangle\}$ for $\rho = \sum_i \lambda_i |i\rangle\langle i|$. The contextual, or 'fine-grained', entropy

$$H(x,\rho) = -\sum_x p_x \log p_x \geq S(\rho) \tag{2.48}$$

in a basis $\{|x\rangle\}$ is the Shannon entropy of probability distribution $H(p_x)$ for $p_x = \langle x|\rho|x\rangle$. $S(\rho)$ is the lower bound of other contextual entropies $H(x,\rho)$.

Shannon entropy also satisfies:

- concave.
- upper bounded by $\log d$ for system dimension d.
- conditional entropy chain rule: $H(Y|X) := H(XY) - H(X)$ is nonnegative, and equal to zero for completely correlated X and Y.

Question 36. *Is the contextual entropy upper bounded?*

The contextual entropy is potentially not upper bounded. The reason is as follows. The decomposition of a state ρ can be very complicated: it can be a convex sum (or superposition) of an unbounded number of other states. This complicated (i.e., 'fine-grained') decomposition will increase the contextual entropy. In practice, we probably would not encounter such a scenario as physicists prefer simplicity very much rather than being cumbersome.

Given two sets of bases, $\{|x\rangle\}$ and $\{|y\rangle\}$, the Shannon entropy $H(x,\rho)$ and $H(y,\rho)$ shows tradeoff by Maassen–Uffink entropic inequality

$$H(x,\rho) + H(y,\rho) \geq -\log \max_{x,y} |\langle x|y\rangle|^2 + S(\rho), \tag{2.49}$$

while the von Neumann entropy $S(\rho)$ of ρ is often omitted. The equality holds for a pure state that serves as a common element of the two sets. It appears as state-independent since the von Neumann entropy of pure states vanish. But the contextual entropy of a pure state is nonzero, in general. The basis is better to be a PVM; otherwise, POVM or channels will introduce additional feature to the state.

2.2.2 Entanglement

When there is a partition, the shared coherence among subsystems leads to nontrivial correlation among these parts. The quantum correlation is usually referred as 'entanglement'.

There can be direct sum and direct product of Hilbert spaces, and this provides the partition, $\mathscr{I}$ and $\mathscr{J}$, for

$$\mathscr{H} = \bigoplus_{i\in\mathscr{I}} \bigotimes_{j\in\mathscr{J}} \mathscr{H}_{ij}. \tag{2.50}$$

We emphasis that entanglement is defined given the partition set. The partition should be physically accessible so the entanglement can be easily measured.

Entangled states are defined in a negative way: a state is entangled if it is not separable. A separable state takes the form

$$\rho = \sum_k \sum_i p_{ki} \otimes_j \rho_{kij}. \tag{2.51}$$

A proper measure of entanglement E is called a 'monotone' with conditions

1. it is positive $E \geq 0$.
2. it is additive for tensor product $E(\rho_1 \otimes \rho_2) = E(\rho_1) + E(\rho_2)$.
3. it is upper bounded; the upper bound is $\log_2 d$ for Hilbert space dimension d.
4. it decreases under entanglement-degrading operations, e.g., SLOCC.

Note SLOCC stands for stochastic local operation and classical communication. The last condition is a bit redundant, which is clear to see.

Some additional conditions

1. it is convex for mixing $E(\sum_i p_i\rho_i) \leq \sum_i p_i E(\rho_i)$.
2. it reduces to entanglement entropy for pure state case.

Note the entanglement entropy of a pure state $|\psi\rangle$ is $S(\rho_A)$ for ρ_A as a reduced state of it, given such a partition.

Entanglement is a *partial order* on the set of states. A single measure is not enough to provide a total order of all states. There are many measures that can be found in literature, including concurrence, negativity, mutual information, entanglement of entropy, distillation, or formation, etc. There are also witness of entangled states, entanglement for identical particles or relativistic ones, and dynamics of entanglement under quantum channels. We would not study these topics in details.

Entanglement is applied in various settings, e.g., quantum communication, thermodynamics, phase transition and topological order, computing and codes, holography and gravity.

Question 37. *What is the relation between entanglement and quantum field theory?*

The relation is complicated. They are different frameworks. For entanglement, there are short-distance and long-distance ones. In many-body system, e.g., a quantum computer, the local property and relations among parts are important, while the global features are difficult to extract. While quantum field theory aims to describe global long-distance features, ignoring the local and relational features. That is, field theory and many-body theory are kind of complementary to each other. If you think the internet as a complicated many-body system, the interaction among local nodes (computers) are important, while the global features of the whole internet are difficult to see.

2.2.3 Uncertainty

A quantum system carries lots of non-commuting operators. How can these operators be consistent since they do not commute? Well, non-commuting does not mean inconsistency, instead it refers to a sort of intrinsic structure of a quantum system, such as coherence, entanglement, etc. Here we analyze uncertainty.

Question 38. *Given a quantum system, what is uncertain of it?*

The uncertain thing is about some observable, once other non-commuting observable becomes certain, i.e., it is a tradeoff, duality, or complementary feature.

Uncertainty is a property depending on both states and observable. It includes the so-called nonlocality as a special case. In particular, the quantum bound on Bell inequality follows from uncertainty relation.

The uncertainty relation is a *logical* relation, i.e., it does not necessarily relate to practical situations, such as noise and disturbance in practical measurement. It can be derived from the Cauchy-Schwarz inequality.

For a system s composed with two subsystems s_A and s_B, the Hilbert space is bipartite $\mathscr{H} = \mathscr{H}_A^n \otimes \mathscr{H}_B^m$, n and m are the dimensions. Suppose observable A_α and A_β act on s_A, observable B_α and B_β act on s_B. We require $[A_\alpha, A_\beta] \neq 0$, $[B_\alpha, B_\beta] \neq 0$, and $[A_i, B_j] = 0$ for $i, j = \alpha, \beta$. Define two new observable on subsystem s_B as $P_\alpha = B_\alpha + B_\beta$, $P_\beta = B_\alpha - B_\beta$, then $[P_\alpha, P_\beta] \neq 0$, $[A_i, P_j] = 0$.

The uncertainty relation is about variance and covariance. The variance of an operator T is defined as

$$\mathrm{Var}(T) = \langle (\Delta T)^2 \rangle = \langle T^2 \rangle - \langle T \rangle^2, \tag{2.52}$$

The covariance between T_α and T_β are

$$\begin{aligned} \mathrm{Cov}(T_\alpha, T_\beta) &= \langle T_\alpha T_\beta \rangle - \langle T_\alpha \rangle \langle T_\beta \rangle, \\ \mathrm{Cov}(T_\beta, T_\alpha) &= \langle T_\beta T_\alpha \rangle - \langle T_\alpha \rangle \langle T_\beta \rangle. \end{aligned} \tag{2.53}$$

Note the expectation $\langle\cdot\rangle$ is done on the corresponding state.

The Cauchy-Schwarz inequality takes the form

$$|\langle f|g\rangle|^2 \leq \langle f|f\rangle\langle g|g\rangle, \tag{2.54}$$

for two arbitrary state vectors $|f\rangle$ and $|g\rangle$ in Hilbert space. Let $\langle f|f\rangle = \mathrm{Var}(A_\alpha)$, $\langle g|g\rangle = \mathrm{Var}(A_\beta)$, $|\langle f|g\rangle|^2 = \mathrm{Cov}(A_\alpha,A_\beta)\mathrm{Cov}(A_\beta,A_\alpha)$, then

$$\mathrm{Cov}(A_\alpha,A_\beta)\mathrm{Cov}(A_\beta,A_\alpha) \leq \mathrm{Var}(A_\alpha)\mathrm{Var}(A_\beta), \tag{2.55}$$

which can also be expressed as the uncertainty relation

$$\Delta A_\alpha \Delta A_\beta \geq \sqrt{\left(\tfrac{1}{2}\langle\{A_\alpha,A_\beta\}\rangle - \langle A_\alpha\rangle\langle A_\beta\rangle\right)^2 + \left(\tfrac{1}{2i}\langle[A_\alpha,A_\beta]\rangle\right)^2}, \tag{2.56}$$

for $\Delta A_i = \sqrt{(\Delta A_i)^2}$, and the Heisenberg uncertainty relation is

$$\Delta A_\alpha \Delta A_\beta \geq \left|\frac{1}{2i}\langle[A_\alpha,A_\beta]\rangle\right|. \tag{2.57}$$

With the same method, on the whole system we find

$$\mathrm{Cov}(A_i,P_j)^2 \leq \mathrm{Var}(A_i)\mathrm{Var}(P_j), \tag{2.58}$$

which becomes

$$\langle A_\alpha P_\alpha\rangle + \langle A_\beta P_\beta\rangle \leq \sqrt{\mathrm{Var}(A_\alpha)\mathrm{Var}(P_\alpha)} + \sqrt{\mathrm{Var}(A_\beta)\mathrm{Var}(P_\beta)} + \langle A_\alpha\rangle\langle P_\alpha\rangle + \langle A_\beta\rangle\langle P_\beta\rangle, \tag{2.59a}$$

and

$$\langle A_\alpha P_\alpha\rangle + \langle A_\beta P_\beta\rangle \geq -\sqrt{\mathrm{Var}(A_\alpha)\mathrm{Var}(P_\alpha)} - \sqrt{\mathrm{Var}(A_\beta)\mathrm{Var}(P_\beta)} + \langle A_\alpha\rangle\langle P_\alpha\rangle + \langle A_\beta\rangle\langle P_\beta\rangle, \tag{2.59b}$$

where we denote $\langle A_\alpha P_\alpha\rangle \equiv \langle A_\alpha \otimes P_\alpha\rangle$, etc. Introduce the CHSH operator $\mathfrak{B} := A_\alpha P_\alpha + A_\beta P_\beta$, then the above two inequalities provides the upper bound (2.59a) and lower bound (2.59b) for the expectation value of $\mathfrak{B}$.

It is clear that inequality (2.59) applies to all bipartite quantum states, including entangled and separable ones, thus it is not supposed to be employed to witness entanglement directly; on the contrary, it specifies the correlation between observable of the two subsystems under a special state. The equality "=" holds for some particular state and observable.

For dichotomic operators A_i and B_i, and $\langle A_i\rangle = 0$, $\langle B_i\rangle = 0$, $\langle P_i\rangle = 0$, $A_i^2 = \mathbb{1}$, $B_i^2 = \mathbb{1}$, we have

$$|\langle\mathfrak{B}\rangle| \leq \sqrt{\langle A_\alpha^2\rangle\langle P_\alpha^2\rangle} + \sqrt{\langle A_\beta^2\rangle\langle P_\beta^2\rangle} = \sqrt{2+\lambda} + \sqrt{2-\lambda} \leq 2\sqrt{2}, \tag{2.60}$$

where $\langle A_\alpha^2\rangle = \langle A_\beta^2\rangle = 1$, $\lambda \equiv \langle\{B_\alpha,B_\beta\}\rangle$. The inequality above is the Tsirelson bound. We can verify that the bound is saturated as $2\sqrt{2}$ for the singlet state by operators $A_\alpha = Z\otimes\mathbb{1}$, $A_\beta = X\otimes\mathbb{1}$, $B_\alpha = -\frac{\sqrt{2}}{2}\mathbb{1}\otimes(Z+X)$, $B_\beta = \frac{\sqrt{2}}{2}\mathbb{1}\otimes(Z-X)$, where X,Y,Z are Pauli operators.

2.3 Geometric phases

We have studied the basics of geometric phases in Chapter 1. We know that geometric phases arise from time-dependent or parameter-dependent Hamiltonian evolution, which are required to be of particular forms, e.g., adiabatic or cyclic, in order to define meaningful geometric phases.

Question 39. *What is the essence of geometric phases?*

Geometric phases can be defined without Hamiltonian. What we need is a one-parameter family of states $|\psi_r\rangle$ for $r \in [0,R]$ such that the states vary smoothly and $|\psi_R\rangle = e^{i\gamma}e^{i\delta}|\psi_0\rangle$ for dynamical phase δ and geometric phase γ. The smoothness can be defined via the overlap $\langle\psi_{r-1}|\psi_r\rangle$ or the derivative $\partial_r|\psi_r\rangle$. The geometric phase is defined as

$$\gamma := i\int \langle\psi_r|\partial_r|\psi_r\rangle dr. \tag{2.61}$$

Depending on the manifold of the parameter, the geometric phase can depend on the geometry (such as solid angle) or the topology (such as genus). The geometric phase might be topological if the parameter space is topologically nontrivial, i.e., the homotopy group is nontrivial. The change of parameters can be adiabatic or cyclic, but this is not required. The cyclic adiabatic case is known as Berry phase, the cyclic non-adiabatic case is Aharonov-Anandan phase, and the non-cyclic case is not widely used, though. Berry phase applies to more general systems since the adiabatic evolution is easy to handle, while Aharonov-Anandan phase applies to some special systems that are easy to solve. The smooth change can be driven by a Hamiltonian, which does not have to be periodic. For periodic Hamiltonian, the geometric phase is easy to obtain due to Bloch theorem and Floquet theory.

2.3.1 Aharonov-Anandan phase

For non-adiabatic evolution, it is hard to solve the state equation. However, if we *assume* there exists cyclic solution, then we can define cyclic non-adiabatic geometric phase, which is known as Aharonov-Anandan (AA) phase.

If $\psi(t)$ is a solution of the state equation

$$i\dot{\psi}(t) = H(t)\psi(t) \tag{2.62}$$

then we define AA dynamical phase

$$\alpha^d(t) := \int_0^t \langle\psi(t')|H(t')|\psi(t')\rangle dt' \tag{2.63}$$

and the AA-lift state

$$\tilde{\psi}(t) = e^{i\alpha^d(t)}\psi(t), \tag{2.64}$$

which satisfies

$$i\dot{\tilde{\psi}}(t) = [H(t) - \langle\psi(t)|H(t)|\psi(t)\rangle]\tilde{\psi}(t), \tag{2.65}$$

with $\tilde{\psi}(0) = \psi(0)$. The term $\langle\psi(t)|H(t)|\psi(t)\rangle$ is a shift of the spectrum of $H(t)$, and it does not have nontrivial dynamical effects, while it is crucial for the definition of AA phase. Physically, the state $\tilde{\psi}(t)$ eliminates the dynamical phase since $\langle\tilde{\psi}(t)|\partial_t|\tilde{\psi}(t)\rangle = 0$.

A state is *cyclic* if $|\langle\psi(T)|\psi(0)\rangle| = 1$ after time T. The total phase $\alpha(T)$ is defined as

$$e^{-i\alpha(T)} := \langle\psi(T)|\psi(0)\rangle, \tag{2.66}$$

and the geometric phase is $\gamma(T) := \alpha(T) - \alpha^d(T)$. It turns out $\tilde{\psi}(t)$ can be written as

$$|\tilde{\psi}(t)\rangle = e^{i\gamma(t)}|\phi(t)\rangle \tag{2.67}$$

and $\phi(0) = \phi(T) = \psi(0)$. The state $\phi(t)$ is called a closed-lift of $\psi(t)$. As the result,

$$|\psi(t)\rangle = e^{-i\alpha^d(t)}e^{i\gamma(t)}|\phi(t)\rangle, \tag{2.68}$$

and $|\psi(T)\rangle = e^{-i\alpha^d(T)}e^{i\gamma(T)}|\phi(0)\rangle$. The AA geometric phase can be obtained from derivative of Eq. (2.67) as

$$\gamma(T) := i\int_0^T \langle\phi(t)|\partial_t|\phi(t)\rangle dt. \tag{2.69}$$

Note $|\phi(t)\rangle$ is not the exact solution of the original state equation and can be hard to find. The cyclic condition is no easier than the adiabatic condition, but we will see it is powerful and it is general than the periodic condition of H. When the dynamical phase is zero, the evolution phase is purely geometrically.

2.3.2 Applications

Here we would like to present examples of geometric phase in the many-body systems.

A diaelectric material contains a collection of positive and negative charges, but in total neutral. The polarization can be changed by external electric field. The polarization itself is multi-valued, but its change is single-valued. The change is in terms of geometric phase and localized Wannier functions, and can be measured.

With localized Wannier functions $w_n(r)$ for band n, the average position of the electrons, called Wannier center $\bar{r}_n$ is $\bar{r}_n = \int w_n^*(r) r w_n(r) dr$. With $r = -i\partial_k$, it holds

$$\bar{r}_n \propto \oint_{\mathrm{BZ}} e^{-ika}\langle u_{nk}|\partial_k|u_{nk}\rangle dk, \tag{2.70}$$

where the integral is in the Brillouin zone. This is in the form of geometric phase! Its value is not directly measurable; instead, difference between phases can be measured, and that is why only the change of polarization is physical and can be measured.

In strongly-correlated magnetism, there are states that are formed with singlets. A singlet is formed by a pair of spin-1/2. In total the state does not have magnetization. We can also define a certain macroscopic polarization although it should be of the spin operators. It takes the form

$$P = \frac{1}{L}\sum_n nS_n^z \tag{2.71}$$

for a 1D system of size L with periodic boundary condition. It turns out $2\pi P$ is a geometric phase, and the phase factor $e^{i2\pi P}$ can be generated by an external field that is of radial direction to the system. Namely, it can be done by inserting a line of charge along the center of the hole encircled by the periodic system. As well, the phase cannot be measured directly. Instead, its derivative can be measured, or the difference between two values on two states can be measured. For instance, it can be measured by Thouless pump of spinon ('spin current'), which is the spin analog of electric current.

The quantized Hall conductivity of quantum Hall states is another example of Berry phase. On a 2D sample with directions x and y, let $\theta_i = 2\pi\Phi_i/\Phi_0 \in [0, 2\pi)$, $i = x, y$ with flux quanta Φ_0, the cross Hall conductivity is

$$\sigma_{xy} = -e^2 \int d\theta_x d\theta_y F_{xy} \tag{2.72}$$

for $F_{xy} = \partial_{\theta_y} A_x - \partial_{\theta_x} A_y$ as Berry curvature, or field strength, and Berry connection $A_i = -i\langle G|\partial_{\theta_i}|G\rangle$ for a ground state $|G\rangle$. The value $\int d\theta_x d\theta_y F_{xy}$ is nothing but the first Chern number. The parameter space is formed by the two fluxes along the two directions. For a torus, the fluxes are quantized. In transport experiment, which can measure changes of σ_{xy}, we can see a number of plateaux due to the quantized Chern number (also see Chapter 5).

2.3.3 Generalizations

The geometric phase can be generalized to the setting when there are several degenerate states as 'non-abelian geometric phases'. The degeneracy condition is necessary to avoid dynamical effects. The 'non-abelian' geometric phases are not actually 'phases'; instead they are unitary operators or gates, called 'holonomy'. For a set of orthonormal states $|\psi_m\rangle$ with parameter r, the matrix-element of the holonomy is

$$\gamma_{mn} = -\int dr\langle\psi_m|\partial_r|\psi_n\rangle. \tag{2.73}$$

This in particular is important for holonomic quantum computing and anyonic quantum computing. The holonomic quantum computing is usually realized in small systems, and gates via holonomy are realized by tuning parameters in Hamiltonian. The degeneracy of states has to be fine-tuned. The anyonic quantum computing is realized

in topological phases of matter with non-abelian anyons, and gates via holonomy are realized by braiding of anyons. The degeneracy of states arises from the fusion space of anyons.

Another important generalization is geometric phases for mixed states under (non-)unitary evolution. Here we lay out the ideas. There are at least three ways to treat mixed states similarly with pure states:

- a mixed state ρ can be viewed as a mixture of its eigenstates $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$;
- ρ can be purified to a state $|\Psi\rangle = \sum_i \sqrt{p_i}|\psi_i\rangle|i\rangle$ with an ancilla;
- ρ can be mapped to a state $|\rho\rangle = (\rho \otimes \mathbb{1})|\omega\rangle$ from channel-state duality.

The geometric phase can be formally defined as $\arg\langle\Psi(t)|\Psi(0)\rangle$, i.e., the phase of the overlap between the initial state and the final state, which is a kind of average of the geometric phases $\arg\langle\psi_i(t)|\psi_i(0)\rangle$ for each eigenstate. However, it is difficult to obtain the eigenstates of a mixed state! The problem becomes a bit easier when the dynamics is specified by a channel or a Lindblad equation (see the next section), in which cases the geometric phases can take different forms.

2.4 Quantum channels

Here we provide a systematic study of quantum channels. We already know quantum channels are the generalization of unitary evolution. We also know that quantum channel can describe measurement. But besides quantum channels, there are also other generalizations of the quantum equation. The most notable one is the Lindblad equation. We will show how to derive it.

In quantum theory, the positivity of the linear map T means that the density matrix ρ stays as a density matrix after the map, i.e., $T(\rho)$ is positive. However, positivity is not enough, since there can be entangled state, so the map is further required to be *completely positive* (CP). A map T is CP iff the spectrum of $T \otimes \mathbb{1}\rho_{SS'}$ is positive for all $\rho_{SS'}$ as the composite state of the system S and another arbitrary system S'. For instance, the partial transpose is not CP since it does not preserve the positivity of certain entangled states.

2.4.1 Representations

Quantum channel can be represented in different ways. But note here 'representation' is not like that for groups. Instead, it means a channel can be described or 'presented' in various ways. We will see the following forms:

- Stinespring dilation
- Kraus operator-sum decomposition

- Choi state and process state
- Affine form and dynamical form

The Stinespring factorization (or dilation) theorem and Choi's theorem show that the general form of any CP map $\mathscr{E}$ is

$$\mathscr{E}(\rho) = \sum_i K_i \rho K_i^\dagger, \tag{2.74}$$

where $K_i \in \mathscr{L}(\mathscr{H})$. This formula is also known as Kraus operator-sum decomposition, with K_i called Kraus operators. When the trace preservation (TP) is satisfied, $\sum_i K_i^\dagger K_i = \mathbb{1}$, the map is CPTP, or a quantum channel. Note that the map does not depend on the formula of the density matrix ρ.

The operator $V = \sum_i K_i |i\rangle$ is an isometry with $V^\dagger V = \mathbb{1}$, which can be embedded in a unitary operator

$$U = \begin{pmatrix} K_0 & \cdot & \cdots & \cdot \\ K_1 & \cdot & \cdots & \cdot \\ \vdots & \vdots & \cdots & \vdots \end{pmatrix} \tag{2.75}$$

as its first block column. It is easy to see $V = U|0\rangle$, and $K_i = \langle i|U|0\rangle$. This means that the ancillary environment E is set at state $|0\rangle$ initially.

Conversely, a channel can be built up given a coupling between S and E. It turns out, even when the environment E has infinite numbers of d.o.f, the effects on the system can be only represented by several terms as above. That is to say, quantum channel is an effective description of the underlying process. There is one primary assumption: the initial state of the whole system is a product state $\rho_S(0) \otimes \rho_E(0)$. Suppose $\rho_E(0) = \sum_i^N p_i |i\rangle\langle i|$. The unitary operation is U. Then, the system state at arbitrary time t is

$$\begin{aligned} \rho_S(t) &= \mathrm{tr}_E(U \rho_S(0) \otimes \rho_E(0) U^\dagger) \\ &= \sum_i^N \sum_e^N p_i \langle e|U \rho_S(0) \otimes |i\rangle\langle i|U^\dagger|e\rangle \\ &= \sum_i^N \sum_e^N K_{ei} \rho_S(0) K_{ei}^\dagger, \end{aligned} \tag{2.76}$$

where the Kraus operator is defined as $K_{ei} \equiv \sqrt{p_i}\langle e|U|i\rangle$, with $\sum_i^N \sum_e^N K_{ei}^\dagger K_{ei} = \mathbb{1}$. Usually, for simplicity, we set the environment at pure state initially, labeled as $|0\rangle$, then the Kraus operator is reduced to $K_e = \langle e|U|0\rangle$.

Question 40. *Are Kraus operators physical?*

Yes they are, but their effects are averaged as a channel. We may assign a physical process to each Kraus operator, but this may not be the real process in practice. In

practice, we do not know the actual representation and basis of the environment. Interestingly, there is a *unitary freedom* or *gauge redundancy* of it, which refers to the fact that different sets of orthonormal basis $\{|\ell\rangle\}$ of E can be chosen. For a unitary operator W acting on E after the interaction U, the channel will be expressed by a set of operators $M_\ell = \langle w_\ell|U|0\rangle$ for $\langle w_\ell| = \langle \ell|W$. The set $\{M_\ell\}$ and $\{K_\ell\}$ represent the same channel since

$$\sum_\ell M_\ell \rho M_\ell^\dagger = \sum_{\ell ab} \langle w_\ell|a\rangle\langle b|w_\ell\rangle K_a \rho K_b^\dagger = \sum_\ell K_\ell \rho K_\ell^\dagger. \tag{2.77}$$

The unitary freedom can be generalized to a CPTP freedom, that is, if W is substituted by a CPTP channel $\mathscr{F}$, represented by a set of Kraus operators $\{F_i\}$, the dynamics on the system is still the same with the original channel $\mathscr{E}$. After the action of both $\mathscr{F}$ and U, the channel takes the form

$$\mathscr{E}(\rho) = \sum_{ab} B_{ab} \rho B_{ab}^\dagger \tag{2.78}$$

for $B_{ab} = \sum_\ell K_\ell \langle a|F_b|\ell\rangle$, and

$$\begin{aligned}\sum_{ab} B_{ab}\rho B_{ab}^\dagger &= \sum_{\ell kab} \langle a|F_b|\ell\rangle\langle k|F_b^\dagger|a\rangle K_\ell \rho K_k^\dagger \\ &= \sum_{\ell kb} \langle k|F_b^\dagger F_b|\ell\rangle K_\ell \rho K_k^\dagger = \sum_\ell K_\ell \rho K_\ell^\dagger.\end{aligned} \tag{2.79}$$

The Choi state form is from the Choi-Jamiołkowski isomorphism $\mathscr{J} : \mathscr{D}(\mathscr{H}) \to \mathscr{H} \otimes \mathscr{H}$ and also $\mathscr{J} : \mathscr{L}(\mathscr{D}(\mathscr{H})) \to \mathscr{L}(\mathscr{H} \otimes \mathscr{H})$, which maps an operator $\mathscr{E} \in \mathscr{L}(\mathscr{D}(\mathscr{H}))$ into a quantum state, called Choi state $\mathscr{C} \in \mathscr{D}(\mathscr{H} \otimes \mathscr{H})$. This isomorphism is also known as the quantum channel-state duality, and the Choi state $\mathscr{C}$ is the dual of the channel $\mathscr{E}$. The Choi state takes the form

$$\mathscr{C} := \mathscr{E} \otimes \mathbb{1}(|\omega\rangle\langle\omega|), \tag{2.80}$$

with bipartite maximally entangled state $|\omega\rangle := \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |i,i\rangle$. Here the Choi state is normalized, and the un-normalized version is also used in some cases. One celebrated property is that the condition of complete positivity is equivalent to the positive semidefiniteness of the Choi state, i.e., $\mathscr{C} \geq 0$.

The Kraus operators relate to the eigenvectors of Choi state by reshaping. For instance, for a unitary operator U, the Choi state is a pure state $|\psi_U\rangle = (U \otimes \mathbb{1})|\omega\rangle = \mathrm{res}U/\sqrt{d}$. Given the state $|\psi_U\rangle$, the unitary operator can be recovered by $\sqrt{d}\mathrm{res}^{-1}|\psi_U\rangle = U$. On the other hand, given the set of Kraus operators, the Choi state can be derived as

$$\mathscr{C} = \frac{1}{d} \sum_i \mathrm{res}K_i \ (\mathrm{res}K_i)^\dagger = \sum_i (K_i \otimes \mathbb{1})\,|\omega\rangle\langle\omega|\left(K_i^\dagger \otimes \mathbb{1}\right). \tag{2.81}$$

In addition, the Choi state can also be equivalently defined as $\mathscr{C} = \mathbb{1} \otimes \mathscr{E}(|\omega\rangle\langle\omega|)$, and then the relation with Kraus operators becomes $\mathscr{C} = \frac{1}{d}\sum_i \mathrm{vec}K_i(\mathrm{vec}K_i)^\dagger$.

The Choi state (2.80) is actually a representation in the tensor-product Kronecker basis $\{|i\rangle\langle j|\otimes|i\rangle\langle j|\}$. The process state $\mathscr{S}$, which is usually denoted as χ, is a matrix equivalent to Choi state $\mathscr{C}=U\mathscr{S}U^\dagger$ by a basis transformation $U=[u_{\alpha\beta}]$ from generalized Pauli basis $\{\sigma_\beta\}$ to Kronecker basis $\{\tau_\alpha\}$ with

$$u_{\alpha\beta}=\mathsf{tr}(\tau_\alpha^\dagger\sigma_\beta). \tag{2.82}$$

This can be shown as follows. Given the Kraus operators $\{K_i\}$ for a channel and the generalized Pauli basis $\{\sigma_\beta\}$, each Kraus operator is a combination $K_i=\sum_\beta \mathsf{tr}(K_i^\dagger\sigma_\beta)\sigma_\beta$. Then $\mathscr{E}(\rho)=\sum_{\alpha\beta}S_{\alpha\beta}\sigma_\alpha(\rho)\sigma_\beta$, with $S_{\alpha\beta}=\sum_i\mathsf{tr}(K_i^\dagger\sigma_\alpha)\mathsf{tr}(K_i\sigma_\beta)^*$. In the generalized Pauli basis, the channel can be represented as $\mathscr{S}=[S_{\alpha\beta}]$.

Question 41. *Can Choi states be acted upon by something?*

The answer should be yes since states should be acted upon by channels. These operators are known as 'quantum combs', which are also CP and map channels to channels. If we imagine a channel as a quantum circuit, then a quantum comb can be viewed as a quantum program that can alter the circuit to another one. At this point, we do not discuss this further.

In a canonical and orthonormal basis, denoted as $\{\sigma_i\}$, a quantum state can be written as $\rho=\frac{1}{d}(\mathbb{1}+\sum_i p_i\sigma_i)$. As a result, the state ρ can be represented by the vector $\boldsymbol{p}:=(p_i)$. Examples of canonical and orthonormal basis include the Pauli basis for qubit case and generalized Pauli bases for qudit case. With $\rho\mapsto\boldsymbol{p}$, a quantum channel maps to an affine map

$$\mathscr{E}\mapsto\mathscr{T}=\begin{pmatrix}1 & \mathbf{0}\\ \boldsymbol{t} & \boldsymbol{T}\end{pmatrix},\ \mathscr{T}_{ij}=\frac{1}{d}\mathsf{tr}\left[\sigma_i\mathscr{E}(\sigma_j)\right], \tag{2.83}$$

which contains the *shift* vector $\boldsymbol{t}$ and *distortion* matrix $\boldsymbol{T}$. The affine map is $\mathscr{T}:\boldsymbol{p}\mapsto\boldsymbol{T}\boldsymbol{p}+\boldsymbol{t}$. The parameters in the affine map $\mathscr{T}$ are constrained by the complete positivity condition of a quantum channel. This means that quantum channels do not correspond to arbitrary affine maps. In the Gell-Mann basis, $p_i\in\mathbb{R}$, hence $\mathscr{T}_{ij}\in\mathbb{R}$. It is clear to see that there are indeed d^4-d^2 real parameters.

The $\mathscr{T}$ operator is defined in an orthonormal and canonical basis, while a so-called dynamical operator, or 'transfer matrix', denoted as $\mathscr{D}$, can be defined in the Kronecker basis $\{\tau_i\}$. Clearly, $\mathscr{D}=U\mathscr{T}U^\dagger$ for U from Eq. (2.82). Also the $\mathscr{D}$ operator takes the form

$$\mathscr{D}=\frac{1}{d}\sum_i K_i\otimes K_i^*. \tag{2.84}$$

It can be derived from $\mathscr{D}_{ij}=\frac{1}{d}\mathsf{tr}[\tau_i^\dagger\mathscr{E}(\tau_j)]$. It is equivalent to the Choi state by reshuffling $\langle ik|\mathscr{D}|jl\rangle=\langle ij|\mathscr{C}|kl\rangle$. In this form, a quantum state ρ is represented by its reshaping $\mathsf{res}\rho$, and the dynamics is $\mathsf{res}\rho\mapsto\mathscr{D}\mathsf{res}\rho$. For $K_\ell=\sum_{ij}k_{ij}^\ell|i\rangle\langle j|$, then

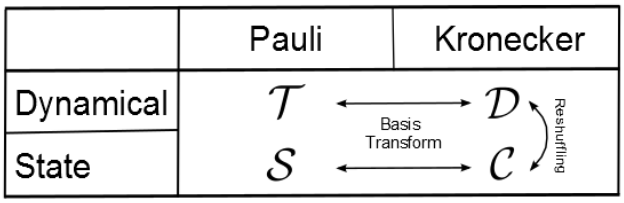

Figure 2.2: The representations of quantum channel. In the table, "Pauli" represents the Pauli basis or the generalized Pauli basis, and "Kronecker" represents the Kronecker basis, and there exists a basis transformation between the two bases. "Dynamical" means that the affine form $\mathscr{T}$ and dynamical operator $\mathscr{D}$ are operators for the dynamics of a quantum channel, and "State" means that a quantum channel is represented by a quantum state, the Choi state $\mathscr{C}$, which relates to $\mathscr{D}$ by the reshuffling operation, or the process state $\mathscr{S}$.

$\mathscr{D} = \sum_{\ell ijkl} k_{ij}^{\ell} \bar{k}_{kl}^{\ell} |ik\rangle\langle jl|$. Ignoring the coherence part the matrix

$$\mathscr{D}' = \sum_{\ell ij} k_{ij}^{\ell} \bar{k}_{ij}^{\ell} |ii\rangle\langle jj| \tag{2.85}$$

is stochastic (not doubly) as $\sum_{\ell i} |k_{ij}^{\ell}|^2 = 1$, and can be treated as the stochastic version of $\mathscr{D}$. Note $|ii\rangle$ can be simply viewed as an encoding of $|i\rangle$, same for $\langle jj|$. For the unitary case, $\mathrm{res} : U\rho U^{\dagger} \mapsto (U \otimes U^*)\mathrm{res}\rho$, $\mathscr{D}'_{\mathscr{U}} = \sum_{ij} |u_{ij}|^2 |i\rangle\langle j|$ of a unitary operator U is doubly stochastic, and also orthostochastic. The stochastic version of a random unitary channel is also doubly stochastic.

Representations of quantum channels are summarized in Fig. 2.2.

Question 42. *Can the forms of quantum channels be slightly extended?*

The answer is yes. There are cases that Kraus operators depend on the initial state, and these channels are designed for the particular physical contexts. Here are two examples.

- When there are initial correlation between system S and environment E: We can enlarge the environment by a fiducial system S', and use a SWAP operation, then the unitary evolution U is replaced by

  $$W = U_{SE} SWAP_{SS'} (\rho_S \otimes \rho_{S'E}), \tag{2.86}$$

  while $\rho_{S'E} = \rho_{SE}$, and $\rho_S = \mathrm{tr}\rho_{SE}$. Now tracing out $\rho_{S'E}$ will lead to a Kraus operator form of a channel.

- Given any two states ρ, σ, there exists a *replacement* channel that maps between them. Say, let $\{|\phi_i\rangle\}$ be a basis, $K_{ij} := \sqrt{q_j}|j\rangle\langle\phi_i|$ for $\sigma = \sum_j q_j |j\rangle\langle j|$, then $\mathscr{E}(\rho) = \sum_{ij} K_{ij}\rho K_{ij}^{\dagger} = \sigma$. This can be applied to the following setting: for

$\mathscr{E}(t_1,t_0)[\rho(0)] = \sum_i K_i(t_1)\rho(0)K_i^\dagger(t_1)$, also $\mathscr{E}(t_2,t_0)[\rho(0)] = \sum_i K_i(t_2)\rho(0)K_i^\dagger(t_2)$, then there exists

$$\mathscr{E}(t_2,t_1)[\rho(t_1)] = \sum_i K_i(t_2,t_1,\rho(t_1))\rho(t_1)K_i^\dagger(t_2,t_1,\rho(t_1)). \tag{2.87}$$

Contrary to the unitary evolution $U(t_2,t_0) = U(t_2,t_1)U(t_1,t_0)$, the map $\mathscr{E}(t_2,t_0) \neq \mathscr{E}(t_2,t_1)\mathscr{E}(t_1,t_0)$, which means quantum channel is generically *indivisible*.

2.4.2 Lindblad equation

We now study Lindblad equation, which describes continuous-time open-system dynamics. The idea of 'open' does not refer to the geometric boundary condition; instead, based on statistical mechanics, it means there can be energy, particle, or information exchange between the system and the outside world, termed as 'environment'. By definition, unitary evolution is for closed system.

Question 43. *Where does Markov sit in the framework of quantum channel?*

Generally, quantum channel describes both Markovian and non-Markovian processes, where Markovian corresponds to a semigroup generated by a map. The *one parameter semigroup* T_t, with the parameter $t \in \mathbb{R}$, satisfies

$$\begin{aligned} T_tT_s = T_{t+s}, \forall t,s \quad &\cdots \quad \text{semigroup divisibility} \\ \lim_{t\to+0} T_t = \mathbb{1}, \quad &\cdots \quad \text{continuity} \\ (T_t \otimes \mathbb{1}) \geq 0, \quad &\cdots \quad \text{complete positivity} \\ \mathsf{tr}(T_t\rho) = \mathsf{tr}\rho, \quad &\cdots \quad \text{trace preservation} \end{aligned}$$

Referring to the property of continuity, introduce operator L satisfying $\dot{T}_t = LT_t$, $L = \dot{T}_t|_{t=0}$, then L is called the *generator* of the semigroup. In classical theory, the counterpart of density matrix is the probability (distribution) vector of a random variable $X(t)$. A classical process is Markovian iff at time t_n, the value $X(t_n)$ only depends on $X(t_{n-1})$ at time t_{n-1}. Markovian is the same as infinite divisibility.

Question 44. *How to derive Lindblad equation?*

Lindblad equation describes Markovian quantum channels. There are several ways to derive the Lindblad equation, next we present one general way from the Kraus operator-sum forms. Given a set of Kraus operators $\{K_e\}$, in a canonical basis $\{F_i\}$, we have

$$K_e = \sum_i F_i \mathsf{tr}(F_i^\dagger K_e). \tag{2.88}$$

Then the system state after the map can be written as

$$\rho(t) = \sum_{i,j=1}^{N^2} S_{ij}(t)F_i\rho(0)F_j^\dagger, \tag{2.89}$$

for the process state $S(t) = [S_{ij}]$. From

$$\dot{\rho}(t) = \lim_{\tau\to 0} \frac{\rho(t+\tau) - \rho(t)}{\tau} = \lim_{\tau\to 0} \frac{\mathscr{E}(t+\tau,t) - \mathbb{1}}{\tau}\rho(t), \tag{2.90}$$

we define the Liouville operator

$$\mathscr{L} \equiv \lim_{\tau\to 0} \frac{\mathscr{E}(t+\tau,t) - \mathbb{1}}{\tau}. \tag{2.91}$$

After some algebra, we find

$$\mathscr{L}\rho(t) = -i[H,\rho(t)] + \sum_{i,j=1}^{N^2-1} G_{ij}\left(F_i\rho(t)F_j^\dagger - \frac{1}{2}\{F_j^\dagger F_i, \rho(t)\}\right), \tag{2.92}$$

with

$$H = \frac{1}{2i\sqrt{d}}\left(\sum_{i=1}^{d^2-1} G^*_{id^2}F_i^\dagger - G_{id^2}F_i\right), \tag{2.93}$$

and $G_{id^2} = \lim_{\varepsilon\to 0} S_{id^2}(\varepsilon)/\varepsilon$, $i = 1,\ldots,d^2-1$. Eq. (2.92) is the non-diagonal form of Lindblad equation. The matrix G, also known as Gorini-Kossakowski-Sudarshan (GKS) matrix, is positive then can be diagonalized by unitary transformation U, $UGU^\dagger = \mathrm{diag}(\gamma_1,\gamma_2,\ldots,\gamma_{d^2-1})$. The basis transfers as $F_i = \sum_{k=1}^{d^2-1} U_{ki}L_k$, then we arrive at the diagonal form of Lindblad equation

$$\dot{\rho}(t) = \mathscr{L}\rho(t) = -i[H,\rho(t)] + \sum_{k=1}^{N^2-1} \gamma_k\left(L_k\rho(t)L_k^\dagger - \frac{1}{2}\{L_k^\dagger L_k, \rho(t)\}\right). \tag{2.94}$$

The solution is $\rho(t) = e^{-it\mathscr{L}}\rho(0)$. The dynamics includes two parts: the work done on the system, represented by Hamiltonian H which contains Lamb-shift terms, and the heat transfer represented by jump operators L_k. The coefficients γ_k relates to the correlation function of the environment and represents the relaxation rates.

2.4.3 Markovianity

There also exists a slight generalization of Lindblad equation, called the 'master equation', which could be non-Markovian but still a special case of quantum channels. A general form of Master equation is

$$\dot{\rho}(t) = \mathscr{L}\rho(t) = -i[H,\rho(t)] + \int_0^t \mathscr{K}\rho(\tau)d\tau, \tag{2.95}$$

and $\mathscr{K}$ is a 'memory kernel'. For (Born-) Markovian case, $\mathscr{K}\rho(\tau) = \mathscr{K}\delta(t-\tau)\rho(\tau)$.

Another popular formalism in chemistry is the Nakajima-Zwanzig equation, which deals with system-bath correlation, in particular. This method uses the projector $\mathscr{P}$ which generates the uncorrelated part, and $\mathscr{Q}$ which generates the correlated part of the whole system, namely

$$\begin{aligned} \mathscr{P}\rho &= \mathrm{tr}_E(\rho)\otimes\rho_E = \rho\otimes\rho_E, \\ \mathscr{Q}\rho &= (\mathbb{1}-\mathscr{P})\rho, \end{aligned} \tag{2.96}$$

where the projector satisfies $\mathscr{P}^2=\mathscr{P}$, $\mathscr{Q}^2=\mathscr{Q}$, $\mathscr{P}\mathscr{Q}=\mathscr{Q}\mathscr{P}=0$. Suppose the Hamiltonian is written as

$$H = H + H_E + V, \tag{2.97}$$

then in the interacting picture,

$$\widetilde{\rho}(t) = e^{i(H+H_E)t}\rho(t)e^{-i(H_S+H_E)t}, \tag{2.98}$$

$$\widetilde{V}(t) = e^{i(H_S+H_E)t}V(t)e^{-i(H_S+H_E)t}. \tag{2.99}$$

From the quantum equation, we get

$$\frac{d}{dt}\mathscr{P}\widetilde{\rho}(t) = -i\mathscr{P}[\widetilde{V}(t),\widetilde{\rho}(t)], \tag{2.100}$$

$$\frac{d}{dt}\mathscr{Q}\widetilde{\rho}(t) = -i\mathscr{Q}[\widetilde{V}(t),\widetilde{\rho}(t)]. \tag{2.101}$$

Denote $\mathscr{V}\equiv -i[\widetilde{V}(t),\cdot]$, then from the above two equations, we can derive the Nakajima-Zwanzig equation

$$\begin{aligned} \frac{d}{dt}\mathscr{P}\widetilde{\rho}(t) &= \mathscr{P}\mathscr{V}(t)\mathscr{P}\widetilde{\rho}(t) + \mathscr{P}\mathscr{V}(t)\mathscr{G}(t,0)\mathscr{Q}\widetilde{\rho}(0) \\ &+ \int_0^t du\,\mathscr{P}\mathscr{V}(t)\mathscr{G}(t,u)\mathscr{Q}\mathscr{V}(u)\mathscr{P}\widetilde{\rho}(u), \end{aligned} \tag{2.102}$$

where the propagator is defined as $\mathscr{G}(t,s)=\mathscr{T}\exp\int_s^t d\tau\mathscr{Q}\mathscr{V}(\tau)$. Here $\mathscr{T}$ is time ordering.

Assume $\mathscr{P}\mathscr{V}(t)\mathscr{P}=0$, which is the Friedrichs condition, and $\rho(0)=\rho(0)\otimes\rho_E(0)$, then the above integro-differential equation reduces to

$$\frac{d}{dt}\mathscr{P}\widetilde{\rho}(t) = \int_0^t du\mathscr{K}(t,u)\mathscr{P}\widetilde{\rho}(u), \tag{2.103}$$

with the kernel $\mathscr{K}(t,u)\equiv\mathscr{P}\mathscr{V}(t)\mathscr{G}(t,u)\mathscr{Q}\mathscr{V}(u)$. Generally, this equation is highly non-Markovian due to the memory kernel. To solve the Nakajima-Zwanzig equation, some approximation is needed, one of them is the weak coupling limit, i.e., Markovian case. Make the replacement $V\to\alpha V$, the solution to the second-order is

$$\widetilde{\rho}(t) = \rho(0) - \alpha^2\int_0^t ds\int_0^s du\,\mathrm{tr}_E([\widetilde{V}(s),[\widetilde{V}(s-u),\widetilde{\rho}(s)\otimes\rho_E]]). \tag{2.104}$$

Suppose the interaction can be generally written as $V = \sum_k A_k \otimes B_k$, where $A_k = A_k^\dagger$ ($B_k = B_k^\dagger$) acting on system S (environment E). Also suppose the spectrum of H_S is discrete, with eigenstate $|s\rangle$ and eigenvalue ε_s. Define $A_k(\omega) = \sum_\omega |s\rangle\langle s|A_k|s'\rangle\langle s'|$, with $\omega = \varepsilon_{s'} - \varepsilon_s$, and $\sum_\omega A_k(\omega) = A_k$, $[H_S, A_k^\dagger(\omega)A_l(\omega)] = 0$. After some algebra, the solution for the state of the system is

$$\widetilde{\rho}(t) = \rho(0) + \int_0^t d\tau \sum_\omega \sum_{k,l} \Gamma_{kl}(\omega)[A_l(\omega)\widetilde{\rho}(\tau), A_k^\dagger(\omega)] + \Gamma_{kl}^*(\omega)[A_l(\omega), \widetilde{\rho}(\tau)A_k^\dagger(\omega)], \tag{2.105}$$

with the coefficients $\Gamma_{kl}(\omega)$ given by the one-sided Fourier transformation of the correlation function of the environment as

$$\Gamma_{kl}(\omega) = \int_0^\infty ds e^{i\omega s} \mathrm{tr}([\widetilde{B}_k(s)B_l\rho_E]), \tag{2.106}$$

which is a complex number, then can be written as $\Gamma_{kl}(\omega) = \frac{1}{2}\gamma_{kl}(\omega) + iS_{kl}(\omega)$. Physically, the real part describes the decay of the system state, while the imaginal part describes the shift of the system state. The final formula of the Nakajima-Zwanzig equation arrives at

$$\begin{aligned} \dot{\rho}(t) &= -i[H_S + H_{LS}, \rho(t)] \\ &+ \alpha^2 \sum_\omega \sum_{kl} \gamma_{kl}(\omega)[A_l(\omega)\rho(t)A_k^\dagger(\omega) - \frac{1}{2}\{A_k^\dagger(\omega)A_l(\omega), \rho(t)\}], \end{aligned} \tag{2.107}$$

which is in the non-diagonal form of Lindblad equation, where

$$H_{LS} = \sum_\omega \sum_{kl} S_{kl}(\omega)A_k^\dagger(\omega)A_l(\omega) \tag{2.108}$$

represents the Lamb shift, and $[H_S, H_{LS}] = 0$.

Question 45. *How to tell whether a process is Markovian or not?*

In practice, to determine whether an open-system dynamics is Markovian or not is a difficult problem, hence Markovianity is used as an approximation. For instance, when the time scales τ_s for system and τ_e for bath evolution differs by several orders, e.g., $\tau_s \gg \tau_e$ for RWA in quantum optics, $\tau_e \gg \tau_s$ for quantum Brownian motion, and Born–Oppenheimer approximation for atoms, the correlation time between system and bath can be set to be infinitesimally small. Another way is from the feedback and information exchange aspect. Markovian means there is no feedback from the bath to the system, thus the information only flows from the system to the bath, i.e., the entropy of the system gets bigger.

There are ways to quantify the *degree of Markovianity* using relative entropy, Fisher information, the breakdown of the semigroup property, etc. However, there is no unique method to characterize the non-Markovian memory effects, since there

cannot be a total order on the set of quantum channels. Any measure can only be a partial order, similar with the case of entanglement. For instance, there are two *sufficient* criteria. In Markovian process, the *relative entropy difference* defined as $\Delta S(t) = S[\rho(0)||\rho(\tau)] - S[\rho(t)||\rho(t+\tau)]$ is necessarily positive, and the *fidelity* $F[\rho(t),\rho(t+\tau)] = \{\mathrm{tr}(\sqrt{\sqrt{\rho(t)}\rho(t+\tau)\sqrt{\rho(t)}})\}^2$ never decrease. So the non-Markovianity is marked by $\Delta S(t) < 0$ or $\dot{F}[\rho(t),\rho(t+\tau)] < 0$.

2.4.4 Beyond

As we know, a mixed state can still be understood in terms of pure states. Similarly, open-system dynamics can also be done in this way, and these are often known as *Stochastic Schrödinger equation*. Next, we present the theoretical framework for the nonlinear Schrödinger equation.

Instead of constructing the environment, the reduced density matrix of the system is constructed by averaging the system wave function which satisfies the stochastic Schrödinger equation, i.e., $\rho_S = \overline{|\psi_i\rangle\langle\psi_i|}$, the average is over an ensemble of system state $|\psi_i\rangle$. Actually, the environment is assumed to be a set of Bargmann coherent state $|z\rangle$, which is not normalized, the state of the whole system is written as

$$|\Psi_t\rangle = \int \frac{d^2z}{\pi} e^{-|z|^2} |\psi(t,z^*)\rangle|z\rangle, \tag{2.109}$$

where the system states $|\psi(t,z^*)\rangle = \langle z|\Psi_t\rangle$, By tracing out the environment, the system density matrix is

$$\rho_S(t,z) = \int \frac{d^2z}{\pi} e^{-|z|^2} |\psi(t,z^*)\rangle\langle\psi(t,z^*)|, \tag{2.110}$$

which can be viewed as a classical average.

Assume the initial state is $|\Psi_0\rangle = |\psi_0\rangle|0_e\rangle$, the environment is modeled as collection of harmonic oscillators $H_E = \sum_e \omega_e a_e^\dagger a_e$, and the interaction is $H_I = LB^\dagger + BL^\dagger$, with L as system operator, $B = \sum_e g_e a_e$ as bath operator. Then, the equation for the system state can be derived as

$$\dot{\psi}(t,z^*) = -iH_S\psi(t,z^*) + Lz_t^*\psi(t,z^*) - L^\dagger \int_0^t ds\alpha(t-s)\frac{\delta\psi(t,z^*)}{\delta z_s^*}, \tag{2.111}$$

where $z_t^* = -i\sum_e g_e^* z_e^* e^{i\omega_e t}$, $\alpha(t-s) = \int_0^\infty d\omega J(\omega) e^{-i\omega(t-s)}$, $J(\omega)$ is the *spectral density* of the environment, and we drop the Dirac notation. Then define

$$\frac{\delta\psi(t,z^*)}{\delta z_s^*} = \mathscr{O}(t,s,z^*)\psi(t,z^*), \tag{2.112}$$

the equation for the system reduces to

$$\dot{\psi}(t,z^*) = (-iH_S + Lz_t^* - L^\dagger\overline{\mathscr{O}}(t,z^*))\psi(t,z^*), \tag{2.113}$$

with $\overline{\mathscr{O}}(t,z^*) = \int_0^t ds\alpha(t-s)\mathscr{O}(t,s,z^*)$.

By the ensemble average, the evolution for the system density matrix is

$$\dot{\rho}_S(t,z) = -i[H_S,\rho_S(t,z)] + [L,\rho_S(t,z)\overline{\mathscr{O}}^\dagger(t,z^*)] + [\overline{\mathscr{O}}(t,z^*)\rho_S(t,z),L^\dagger]. \quad (2.114)$$

Under Markovian limit, $\overline{\mathscr{O}} \propto L$, the equation reduces to the Lindblad equation.

However, there is one drawback of the above equations, namely, it is not trace-preserving. By some modification, the equation can be reformed as trace-preserving formula, which we do not discuss in detail here.

Let us finally discuss an example of the above formalism. The Gross-Pitaevskii equation

$$i\dot{\psi}(x,t) = \left(-\frac{\nabla^2}{2m} + V + g|\psi(x,t)|^2 - i\gamma(x)/2\right)\psi(x,t), \quad (2.115)$$

describes BEC state $\psi(x,t)$, $\int |\psi(x,t)|^2 dx = N(t)$. This equation can be derived from master equation with $\psi(x,t) = \langle\hat{\Psi}\rangle \simeq \hat{\Psi} \simeq \hat{\Psi}^\dagger$. The master equation

$$i\dot{\rho} = [H,\rho] + i\mathscr{L}\rho, \quad (2.116)$$

with Hamiltonian

$$H = \int dx^3 \hat{\Psi}^\dagger(x)\left(-\frac{\nabla^2}{2m} + V\right)\hat{\Psi} + \frac{g}{2}\int dx^3 \hat{\Psi}^\dagger(x)\hat{\Psi}^\dagger(x)\hat{\Psi}(x)\hat{\Psi}(x) \quad (2.117)$$

and

$$\mathscr{L}\rho = -\int dx^3 \frac{\gamma(x)}{2}\left(\hat{\Psi}^\dagger(x)\hat{\Psi}(x)\rho + \rho\hat{\Psi}^\dagger(x)\hat{\Psi}(x) - 2\hat{\Psi}(x)\rho\hat{\Psi}^\dagger(x)\right) \quad (2.118)$$

with $[\hat{\Psi}(x),\hat{\Psi}^\dagger(x')] = \delta^3(x-x')$. Time evolution of $\hat{\Psi}(x)$ is given in terms of ρ as

$$\langle\dot{\hat{\Psi}}(x)\rangle = \mathrm{tr}(\dot{\rho}\hat{\Psi}(x)), \quad (2.119)$$

which reduces to

$$i\langle\dot{\hat{\Psi}}(x)\rangle = \left\langle\left[-\frac{\nabla^2}{2m} + V + g\hat{\Psi}^\dagger(x)\hat{\Psi}(x)\right]\hat{\Psi}(x)\right\rangle - i\frac{\gamma(x)}{2}\langle\hat{\Psi}(x)\rangle, \quad (2.120)$$

which is the Gross-Pitaevskii equation.

2.5 Matrix product states

Matrix product states (MPS) are very powerful representations of quantum states. Here we study this formalism in details. For N qudit, the most general state is

$$|\Psi\rangle = \sum_{i_1,\ldots,i_N}^{d} C(i_1,\ldots,i_N)|i_1,\ldots,i_N\rangle. \quad (2.121)$$

We assume the local dimensions are all d for simplicity.

Question 46. *Can the state be viewed as a tensor?*

A tensor is a higher-rank version of a matrix, which is rank two. The state above can be viewed as a rank-N tensor, which yields all the magic of MPS.

Now introduce a matrix $\mathfrak{C}$ with dimension $d \times d^{N-1}$ and elements $\mathfrak{C}_{i_1,(i_2,\ldots,i_N)} = C(i_i,\ldots,i_N)$. Then apply singular-value decomposition (SVD) on it we have $\mathfrak{C} = USV^\dagger$ and $\mathfrak{C}_{i_1,(i_2,\ldots,i_N)} = \sum_{a_1}^{r_1} U_{i_1,a_1} S_{a_1,a_1} V^\dagger_{a_1,(i_2,\ldots,i_N)}$. $r_1 \leq d$ is the rank of matrix $\mathfrak{C}$. Denote $S_{a_1,a_1} V^\dagger_{a_1,(i_2,\ldots,i_N)} = C(a_1,i_2,\ldots,i_N)$, and a *row* vector A^{i_1} with element $A^{i_1}_{a_1} = U_{i_1,a_1}$, then $\mathfrak{C}_{i_1,(i_2,\ldots,i_N)} = \sum_{a_1}^{r_1} A^{i_1}_{a_1} C(a_1,i_2,\ldots,i_N)$.

We put A^{i_1} on the most left. The coefficients $C(a_1,i_2,\ldots,i_N)$ can form a new matrix $\mathfrak{C}'$, and then we apply SVD again and find $\mathfrak{C}_{i_1,(i_2,\ldots,i_N)} = \sum_{a_1}^{r_1}\sum_{a_2}^{r_2} A^{i_1}_{a_1} A^{i_2}_{a_1,a_2}$ $C(a_2,i_3,\ldots,i_N)$. $r_2 \leq r_1 d$ is the rank of the new matrix $\mathfrak{C}'$, elements $A^{i_2}_{a_1,a_2}$ forms a $r_1 \times r_2$ matrix. At the end we will have

$$C(i_1,i_2,\ldots,i_N) = \sum_{a_1,\ldots,a_N}^{r_1,\ldots,r_N} A^{i_1}_{a_1} A^{i_2}_{a_1,a_2} \cdots A^{i_{N-1}}_{a_{N-2},a_{N-1}} A^{i_N}_{a_N-1}. \tag{2.122}$$

Finally we obtain

$$|\Psi_{\text{left}}\rangle = \sum_{i_1,\ldots,i_N}^{d} A^{i_1} A^{i_2} \cdots A^{i_N} |i_1,\ldots,i_N\rangle. \tag{2.123}$$

The above is called the *left-canonical form*, since we start the SVD from the left site. We also use *right-canonical form*, which starts the SVD from site N and put A^{i_N} on the most left, the state takes the form

$$|\Psi\rangle = \sum_{i_1,\ldots,i_N} \langle A^{i_N}|A^{i_{N-1}} \cdots A^{i_2}|A^{i_1}\rangle |i_1 \ldots i_N\rangle \tag{2.124}$$

for the open boundary condition (OBC) case.

These A matrices act on the so-called correlation space, also known as virtual space, ancillary space etc, and the correlation space dimension χ is also known as the bond or virtual dimension. Tracing out the system results in a sequence of quantum channels $\mathscr{E}_n$ on the correlator such that

$$\mathscr{E}_n(\rho) = \sum_{i_n} A^{i_n} \rho A^{i_n\dagger}, \tag{2.125}$$

and $\sum_{i_n} A^{i_n\dagger} A^{i_n} = \mathbb{1}$ for each $n = 1,\ldots,N$.

The boundary condition is specified by the set of column vectors $\{|A^{i_1}\rangle\}$ and the set of row vectors $\{\langle A^{i_N}|\}$. For the first site, $\sum_{i_1} A^{i_1\dagger} A^{i_1} = 1$, each A^{i_1} is a column vector but not-normalized, while its norm is a singular value. For the last site, $\sum_{i_N} A^{i_N\dagger} A^{i_N} = \mathbb{1}$, each A^{i_N} is a row vector and normalized, and they come from each

column of a unitary operator that appears in the first step of SVD to derive MPS. For the OBC case, the form of MPS is usually simplified as

$$|\Psi\rangle = \sum_{i_1,\ldots,i_N} \langle R|A^{i_N}\cdots A^{i_1}|L\rangle |i_1\ldots i_N\rangle, \tag{2.126}$$

which may not be normalized due to the probability of the final projection $\langle R|$. However, the normalization condition can be easily handled, so it does not cause problem. For the PBC case, the MPS takes the following form

$$|\Psi\rangle = \sum_{i_1,\ldots,i_N} \mathrm{tr}(A^{i_N}\cdots A^{i_1})|i_1\ldots i_N\rangle. \tag{2.127}$$

This state can be prepared by using $|\omega\rangle = \sum_i |ii\rangle$ as both the initial and final states of the correlator. The bond dimension is actually χ^2, but the A matrices only act on half of the space, so the effective bond dimension is still χ. Also the PBC case can be viewed as a special case of OBC when each vector $|A^{i_1}\rangle$ is equivalent to $|A^{i_N}\rangle$ and $\{|A^{i_1}\rangle\}$ forms a basis of the correlation space.

2.5.1 MPS circuit

To prepare a MPS (2.124) by a quantum circuit, the dilation for each of the channels $\mathscr{E}_n$ (2.125) is employed. The first channel $\mathscr{E}_1$ is defined by the set of Kraus operators $\{|A^{i_1}\rangle\}$, and the last channel $\mathscr{E}_N$ is defined by $\{\langle A^{i_N}|\}$. The channel $\mathscr{E}_1$ maps from dimension $\chi_0 = 1$ to dimension χ_1, while the channel $\mathscr{E}_N$ maps from dimension χ_{N-1} to dimension $\chi_N = 1$, while each other channel $\mathscr{E}_n$ in between maps the bond dimension from χ_{n-1} to χ_n. From the SVD process there exists relations between each χ and d, e.g., $\chi_{N-1} \le d$. Implementing each $\mathscr{E}_n$ requires the dilation of channels that alter dimension. For a rank-r CPTP channel from dimension n to m, one input ancilla with dimension $\lceil \frac{rm}{n} \rceil$ is needed. Note that the input system and ancilla do not correspond to the output system and ancilla, respectively, due to the change of dimension. Now a channel $\mathscr{E}_n$ can be realized by a unitary U_n with dimension $d\chi_n$, and from $A^{i_n} = \langle i_n|U_n|0\rangle$, $\{A^{i_n}\}$ occupy the first block-column of U_n. The circuit takes the form in Fig. 2.3. For the last unitary U_N, special cares are needed. If $\chi_{N-1} = d$, then no ancilla is needed, which means the correlator itself becomes the last physical spin, and then it is traced out after a unitary rotation U_N such that $\langle A^{i_N}| = \langle i_N|U_N$, which appears in the first step of SVD for the right-canonical form. If $\chi_{N-1} < d$ then an ancilla is needed and $\langle A^{i_N}| = \langle i_N|U_N|0\rangle$ for $|0\rangle$ as the initial state of this ancilla. The whole state preparation process is as follows. First, apply a sequence of unitary gates from U_1 till U_{N-1}

$$U_{N-1}\cdots U_1|0\rangle_v|0\rangle_1\cdots|0\rangle_{N-1} = \sum_{i_1,\ldots,i_{N-1}} A^{i_{N-1}}\cdots A^{i_2}|A^{i_1}\rangle|i_1\ldots i_{N-1}\rangle, \tag{2.128}$$

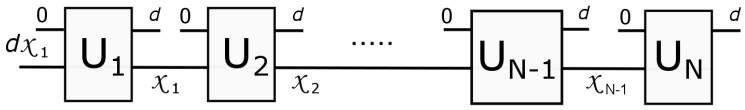

Figure 2.3: Quantum circuit to prepare a general MPS with altering bond dimensions. For each unitary U_n as the dilation of a channel $\mathscr{E}_n$, an input ancilla with initial state $|0\rangle$ is needed, the dimension of which varies from site to site. The final dimension of each spin is always d.

where $|0\rangle_v$ is the initial state of the virtual correlator. If $\chi_{N-1} = d$, apply U_N first and trace out the correlator, the state becomes

$$\begin{aligned}&\sum_{i_N} \langle i_N|U_N U_{N-1}\cdots U_1|0\rangle_v|0\rangle_1\cdots|0\rangle_{N-1}|i_N\rangle \\ &= \sum_{i_N}\sum_{i_1,\ldots,i_{N-1}} \langle i_N|U_N A^{i_{N-1}}\cdots A^{i_2}|A^{i_1}\rangle|i_1\ldots i_{N-1}\rangle|i_N\rangle,\end{aligned} \tag{2.129}$$

which is the MPS (2.124). If $\chi_{N-1} < d$, append the last ancilla with $|0\rangle$ such that $W := U_N|0\rangle$ and $\langle A^{i_N}| = \langle i_N|W$. Applying U_N and tracing out the final system (both correlator and ancilla) yields the MPS (2.124).

Question 47. *Can the bond dimensions be a constant?*

The change of bond dimension complicates the MPS circuit, so usually these matrices are enlarged to have the same bond dimension as the largest one, and indeed, in practice many states can be described by MPS with constant bond dimensions. Therefore, it is assumed that all the A matrices have dimension χ, and each quantum channel becomes dimension-preserving. For the quantum circuit, the first dilation U_1 maps from dimension $d\chi$ to d-dimensional spin and χ-dimensional correlator, and the channels in the middle are simple to deal with, while the last one deserves some attention. The set $\{\langle A^{i_N}|\}$ still forms a channel, but now it may hold $d \leq \chi$, while injectivity requires $d \geq \chi^2$. This means for both injective and also $\chi^2 \geq d \geq \chi$ cases the method described in subsection 2.5.1 can be used. For the case $d < \chi$, the channel cannot be TP since each vector $\langle A^{i_N}|$ is extended to a larger vector. This means partial projection on the correlator is required, which leads to probabilistic events. However, one can employ the method in subsection 2.5.2 to avoid this.

2.5.2 Avoid the final projection

Consider the generation of a MPS in the form (2.126) with a constant bond dimension. It seems a projection has to be done on the correlator.

Question 48. *Is the final projection on the correlator inevitable?*

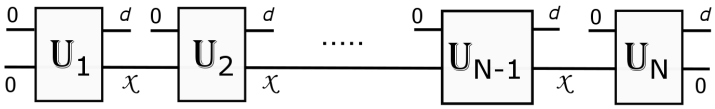

Figure 2.4: Quantum circuit to prepare a general MPS with constant bond dimension χ and an automatically decoupled correlator $|0\rangle$ at the end. Each unitary $\mathbb{U}_n$ is the dilation of $\mathbb{V}_n$, the embedding of V_n'. The initial state of the correlator can be chosen to be $|0\rangle$ by absorbing a unitary gate, which converts $|0\rangle$ to $|L'\rangle$, into the first gate $\mathbb{U}_1$.

It is not. With the isometry $V_n := \sum_{i_n} |i_n\rangle A^{i_n}$ for each site, a MPS with OBC (2.126) can be written as

$$|\Psi\rangle = \langle R|V_N \cdots V_1|L\rangle. \tag{2.130}$$

With $\langle R|V_N = (\mathbb{1}_d \otimes \langle R|)V_N$ and from SVD

$$(\mathbb{1}_d \otimes \langle R|)V_N = V_N' M_N, \tag{2.131}$$

for (i) $d < \chi$, M_N of size $d \times \chi$, unitary V_N' of size $d \times d$, and (ii) $d \geq \chi$, M_N of size $\chi \times \chi$, isometry V_N' of size $d \times \chi$. Now $M_N V_{N-1}$ is $(\mathbb{1}_d \otimes M_N)V_{N-1}$, and perform SVD for the rest sites, and for the last one define $|L'\rangle = M_1|L\rangle$, so

$$|\Psi\rangle = V_N' \cdots V_1'|L'\rangle. \tag{2.132}$$

From a rank consideration, the size of V_{N-k}' is $d\min(\chi, d^k) \times \min(\chi, d^{k+1})$, and the size of M_k is always at most $\chi \times \chi$. Now each V' can be embedded into an isometry $\mathbb{V}$ of size $d\chi \times \chi$, although the embedding is not unique. This means a quantum circuit to realize the sequence of $\mathbb{V}_k$ can be used to prepare the MPS: start from the state $|L'\rangle$, and perform the dilation $\mathbb{U}_k$ for each $\mathbb{V}_k$. To show that the correlator can automatically decouple at the end, there are three cases to consider:

1. For $d^2 \geq \chi > d$, the size of V_{N-1}' is $d^2 \times \chi$, while the size of its embedding $\mathbb{V}_{N-1}$ is $d\chi \times \chi$. This embedding can be done by appending $\chi - d$ rows of zeros to each of the $d \times \chi$ matrices in V_{N-1}', and this means after the action of V_{N-1}', the χ-level correlator will only have amplitude on d levels. The embedding $\mathbb{V}_N$ can be obtained by first appending $\chi - d$ columns of normalized vectors, and then inserting $\chi - 1$ rows of zeros after each row in V_N', and this means that the state of the correlator will be annihilated by V_N', i.e., mapped to dimension one, and the correlator is converted to the last spin by V_N'.

2. For $\chi \leq d$, the size of V_{N-1}' is $d\chi \times \chi$, and its embedding is the same with itself; and size of V_N' is $d \times \chi$, and its embedding $\mathbb{V}_N$ can be obtained by inserting $\chi - 1$ rows of zeros after each row in V_N'. In this case, after V_{N-1}' all levels of the correlator are occupied, yet V_N' will still annihilate the correlator.

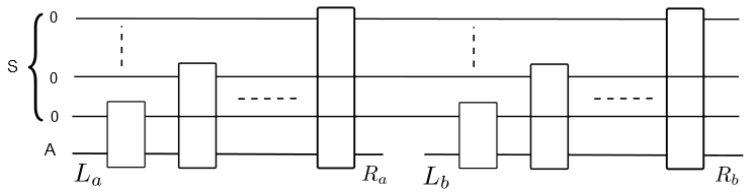

Figure 2.5: Sequential composition property of MPS circuit.

3. For $\chi > d^2$, the size of V'_{N-1} is $d^2 \times d^2$, and its embedding can be obtained by first appending $\chi - d^2$ columns of normalized vectors, and then appending $\chi - d$ rows of zeros to each of the $d \times \chi$ matrices in V'_{N-1}. Still in this case after V'_{N-1} only d levels of the correlator are occupied, which are further annihilated by $\mathbb{V}_N$.

The quantum circuit can be shown as that in Fig. 2.4.

2.5.3 Composition

Question 49. *What are the operations on a set of matrix-product states?*

It is not hard to see that a sum (i.e., linear combination) of MPS leads to a state whose MPS form is not so clear. A reason for this is that the MPS form effectively convert states to operators acting on the correlator, and the natural operation on operators is not sum but product.

In the MPS circuit the starting state of system is usually $|\mathbf{0}\rangle \equiv |0\cdots 0\rangle$. If the input $|\mathbf{0}\rangle$ is substituted by another MPS, the output is still a MPS, but with a larger bond dimension, as shown in Fig. 2.5. Let's denote a MPS by $|\Xi_a\rangle$ and the sequence of unitary operators in it as $\mathscr{U}^{(a)}$, and $|\Xi_a\rangle := \langle R_a|\mathscr{U}^{(a)}|L_a\rangle|\mathbf{0}\rangle$ with bond dimension χ_a, and similarly for another MPS by $|\Xi_b\rangle$. The composition of the two circuits leads to the state

$$|\Xi_{ab}\rangle = \langle R_b|\mathscr{U}^{(b)}|L_b\rangle|\Xi_a\rangle = \langle R_b|\langle R_a|\mathscr{U}^{(ab)}|L_b\rangle|L_a\rangle|\mathbf{0}\rangle, \tag{2.133}$$

with $\mathscr{U}^{(ab)} := \mathscr{U}^{(a)} \diamond \mathscr{U}^{(b)}$ for composition $\diamond$ defined as follows. For $\mathscr{U}^{(a)} := \prod_i U_i^{(a)}$, $\mathscr{U}^{(b)} := \prod_i U_i^{(b)}$, Let $\tilde{U}_i^{(a)} = U_i^{(a)} \otimes \mathbb{1}^{(b)}$, $\tilde{U}_i^{(b)} = U_i^{(b)} \otimes \mathbb{1}^{(a)}$, then $\mathscr{U}^{(ab)} = \prod_i U_i^{(ab)}$ for $U_i^{(ab)} := \tilde{U}_i^{(a)}\tilde{U}_i^{(b)}$. The state $|\Xi_{ab}\rangle$ has bond dimension $\chi_{ab} = \chi_a\chi_b$, and the boundary states of the correlator are $|L_b\rangle|L_a\rangle$ and $\langle R_b|\langle R_a|$.

This property also holds when the technique to avoid the final projection from section 2.5.2 is employed. In the setting of LQTM, this means the action of several groups of sequential unitary operations on the same set of qubits on the tape still prepares a MPS.

In addition, the tensor product of $|\Xi_a\rangle$ and $|\Xi_b\rangle$ also yields a new MPS $|\Xi_{a\otimes b}\rangle$ with bond dimension $\chi_{a\otimes b}=\chi_a\chi_b$ and

$$|\Xi_{a\otimes b}\rangle = \langle R_b|\langle R_a|\mathscr{U}^{(a\otimes b)}|L_b\rangle|L_a\rangle|\mathbf{00}\rangle, \tag{2.134}$$

with $\mathscr{U}^{(a\otimes b)}=\prod_i U_i^{(a\otimes b)}$ for $U_i^{(a\otimes b)}:=U_i^{(a)}\otimes U_i^{(b)}$.

2.5.4 Extract the edge state

For open boundary condition, the states $|L\rangle$ and $|R\rangle$ serve as boundary condition. They are not the physical edge states, i.e. d.o.f. To have an actual edge states, we have to introduce this extra d.o.f. by a slight extension of MPS. Recall that a MPS or PEPS is prepared by projections on entangled pairs. Let $|L\rangle$ be the state of site 1, $|R\rangle$ the state of site N, $|\omega\rangle_n$ the state of sites n and $n+1$. Let the projector $P_n=\sum_{i\alpha\omega}A^i_{\alpha\omega}[n]|i\rangle\langle\alpha\omega|$ act on sites n and $n+1$. Then a MPS is defined as

$$|MPS\rangle = \otimes_{n=1}^{N-1}P_n(|L\rangle\otimes_{n=2}^{N-2}|\omega\rangle_n|R\rangle), \tag{2.135}$$

which is obviously not normalized. By the mapping $P_n\mapsto V_n=\sum_{i\alpha\omega}A^i_{\alpha\omega}[n]|i\rangle|\alpha\rangle\langle\omega|=\sum_i A^i[n]|i\rangle$, $|\omega\rangle_n\mapsto\mathbb{1}_n$, the MPS can also be written as

$$|MPS\rangle = \langle R|V_{N-1}\mathbb{1}_{N-2}\cdots\mathbb{1}_2V_1|L\rangle, \tag{2.136}$$

while the norm comes from the 'projection' $\langle R|$. If we treat $|L\rangle$ as the initial state of an edge, then it is projected out at the end. In order to keep it, we have to avoid the projection. The idea is to replace $|R\rangle$ at site N by a state $|\omega\rangle_N$ with the extra d.o.f. at site $N+1$. The site $N+1$ serves as the space to hold the edge states. Now the state becomes

$$|MPS'\rangle = \mathbb{1}_N V_{N-1}\mathbb{1}_{N-2}\cdots\mathbb{1}_2V_1|L\rangle, \tag{2.137}$$

which contains the actual edge states. Note that in the above we could also replace $|L\rangle$ to obtain the actual edge states at site 0, say.

2.5.5 Area law

Question 50. *Does MPS only describe 1D systems?*

The answer is no. The form of MPS so far does not have relation with spatial dimension yet. The sites in it are just labelled numerically. When we use MPS to describe many-body systems on lattices, there are natural orders of sites. For 1D system, this can be easily done, while for 2D and 3D systems, the ordering is not unique. States for 2D and 3D systems are in general known as tensor-network states, which are also MPS by treating tensors as matrices. In many-body physics, MPS is reserved for the case of 1D gapped system when the bond dimension is a finite constant that does not increase with the system size or other parameters.

For gapped systems, ground states obey the so-called area law of entanglement: the entanglement E between two connected parts is proportional to their interface area A, up to some constants. For a pure state $|\psi\rangle$ on two parts, a and b, the entanglement between them can be measured by the entropy $E = -\mathrm{tr}\rho\log\rho$ for ρ as the reduced state of a or b. We know that E is upper bounded by $\log D$ for D as the dimension of ρ, which is exponential of the volume of ρ. So it seems E shall relate to the volume instead of area. This is indeed true for generic states, however, for gapped ground states described by MPS with constant bond dimensions, this is not true. Instead, gapped ground states have the key feature that correlation functions of local observable are exponentially decaying, which indicates that sites that are far away are not correlated or entangled. This has been proven rigorously and these states are known as short-range entangled. As a result, the two parts a and b are only entangled near their interfaces and we have the area law

$$E = -\mathrm{tr}\rho\log\rho \propto A, \text{ for } \log\dim\rho \propto A. \tag{2.138}$$

We will see later that the additional constants account for additional structure of states: the topology, which is described as long-range entanglement, and others that describe non-short-range entanglement

Chapter 3

Advanced Formalism

The former chapter is about Hilbert space, but we have not dealt with the usual 'space' yet: the space of positions of objects. In a many-body system, the position labels for each subsystem are still classical; i.e., numbers. The question is:

Question 51. *How to make 'positions' quantum?*

If this can be done, we can then imagine quantum coherence 'flows' among these subsystems. The formal treatment is the so-called 'second quantization', or called 're-quantization', which treats the quantum system itself as a classical one and then quantize it again. It serves to explain the origin of infinite-dimensional space and quantum field theory.

In this chapter we explain second quantization and its consequences, such as fermion/boson, relativity, and quantum field theory. We also discuss subtheories including semi-classical and semi-quantum ones since fully-quantum models are usually difficult to solve. We discriminate *semi-classical*, *semi-quantum*, *quantum*, and *classical* forms. Here 'quantum' means the system needs a Hilbert space, 'classical' means the system needs a conventional phase space, 'semi-classical' means it is mainly classical but with a slight quantum addon, 'semi-quantum' means it is mainly quantum but with a slight classical addon.

3.1 Phase space and quantization

In classical mechanics, a particle has a position $q \in \mathbb{R}^3$, and the Newtonian dynamics describes the trajectory $q(t)$ for time $t \in \mathbb{R}^+$. The momentum p is associated with the velocity $\dot{q}$ and is defined as $p = \frac{\partial L}{\partial \dot{q}}$, with Lagrangian $L(q, \dot{q}, t)$. Denote $V \equiv \mathbb{R}^3$, then $p \in V_q^*$, and $(q, p) \in T^*V$. The *phase space* is the cotangent bundle T^*V for the particle dynamics specified by parameters (q, p). The symplectic form is $\omega = dq \wedge dp$. The Hamiltonian is defined as $H = p\dot{q} - L$, from which, the standard

Hamiltonian dynamics follows

$$\frac{\partial H}{\partial p} = \dot{q}, \;\; \frac{\partial H}{\partial q} = -\dot{p}. \tag{3.1}$$

The Poisson bracket for two functions f, g takes the form

$$\{f,g\} = \frac{\partial f}{\partial q}\frac{\partial g}{\partial p} - \frac{\partial g}{\partial q}\frac{\partial f}{\partial p}, \tag{3.2}$$

which also satisfies the Jacobi identity

$$\{f,\{g,h\}\} + \{h,\{f,g\}\} + \{g,\{h,f\}\} = 0, \tag{3.3}$$

for $f,g,h \in C(T^*V)$, and $C(\cdot)$ denotes the set of functions on the phase space. The Hamiltonian dynamics can also be expressed as

$$\{q,H\} = \dot{q}, \;\; \{p,H\} = \dot{p}. \tag{3.4}$$

The Hamiltonian H induces a Hamiltonian flow $h_t : T^*V \to T^*V$, and a vector field $\upsilon_H := \frac{\partial H}{\partial p}\frac{\partial}{\partial q} - \frac{\partial H}{\partial q}\frac{\partial}{\partial p}$, and the Hamiltonian dynamics is equivalent to

$$\dot{\eta} = \upsilon_H \eta, \tag{3.5}$$

with the combined column vector $\eta := (q,p)$, which looks very much like a quantum state.

Question 52. *Is phase space classical or quantum?*

First, phase space is not a space of 'phase', also it is not strictly defined as a space in the sense of mathematics. In physics, phase space just means the set of states for a system, phase means state, status, or configuration, not the phase of a complex variable. So phase space is defined irrelevant to being classical or quantum.

3.1.1 Hamiltonization

The power of phase space is that it describes a system by *conjugate variables*, like position and momentum. The dimension of a phase space is always even. It turns out Hilbert space is a phase space (with an additional structure compared with the usual case). The quantum equation can be written in the form of Hamilton equations! This is known as geometric quantum mechanics (GQM), which is an approach to unify both classical mechanics and quantum mechanics in the phase-space approach. In GQM, the real and imaginary parts of quantum state play conjugate roles and they can be viewed as observable. Also GQM provides a connection between 1st and 2nd quantization, since a quantum system can be viewed as a classical system and then be quantized again.

We start from finite-dimensional unitary evolution of pure states. Quantum states live in projective Hilbert space $\mathscr{H}$ since they are normalized vectors with any global phase physically trivial. Distance between any two states $|\psi\rangle$ and $|\phi\rangle$ is based on the overlap function $\langle\phi|\psi\rangle$. GQM shows that the space $\mathscr{H}$ is a Kähler manifold, which is specified by a symmetric Riemannian form, $\mathrm{Re}(\langle\phi|\psi\rangle)$, and a skew-symmetric symplectic form, $\mathrm{Im}(\langle\phi|\psi\rangle)$. The symplectic form implies that a Hamiltonian dynamics exists and the space $\mathscr{H}$ can be viewed as a phase space. For an orthonormal basis $\{|i\rangle\}$, a state $|\psi\rangle \in \mathscr{H}$ is mapped to a set of coefficients $\psi_i := \langle i|\psi\rangle$. Name the real part $q_i := \mathrm{Re}(\psi_i)$ as "position" and imaginary part $p_i := \mathrm{Im}(\psi_i)$ as "momentum" [1] such that the normalization condition becomes

$$\langle\psi|\psi\rangle = \sum_i |\psi_i|^2 = -i\sum_i \pi_i\psi_i = \sum_i (p_i^2 + q_i^2) = 1, \tag{3.6}$$

for $\pi_i := i\psi_i^*$. As a result, the state equation $i|\dot{\psi}\rangle = \hat{H}|\psi\rangle$ can be equivalently written as Hamilton's equations

$$\frac{\partial H}{\partial q_i} = -\dot{p}_i, \ \frac{\partial H}{\partial p_i} = \dot{q}_i, \ \forall i, \tag{3.7}$$

with classical Hamiltonian (or energy)

$$H = \langle\psi|\hat{H}|\psi\rangle = \sum_{i,j=1}^{d} \psi_i^* H_{ij}\psi_j. \tag{3.8}$$

The symplectic form is $\omega = \sum_{i=1}^{d} d\psi_i \wedge d\pi_i$, and the Hamiltonian dynamics system is (Σ, ω, H). Here Σ denoting the Kähler manifold which is actually the Hilbert space $\mathscr{H}$ itself. We also put a hat on $\hat{H}$ to make clear it is an operator, not the same as H. The classical Hamiltonian H is quadratic of the dynamical variables, there is no higher-order terms, which may appear for nonlinear modifications. Note the above equations are equivalent to

$$\frac{\partial H}{\partial \psi_i} = -\dot{\pi}_i, \ \frac{\partial H}{\partial \pi_i} = \dot{\psi}_i, \ \forall i. \tag{3.9}$$

Also the Hamilton's equations hold for time-dependent Hamiltonian $\hat{H}(t)$. This shows that the unitary dynamics of a quantum state $|\psi\rangle$ is equivalent to the Hamiltonian dynamics of a set of $d := \dim\mathscr{H}$ coupled "particles" (q_i, p_i) in phase space with the normalization condition (3.6).

We define the *hamiltonization* map

$$\text{Hamiltonization} : \mathscr{H} \to \Sigma : \hat{H} \mapsto H, |\psi\rangle \mapsto (\psi_i, \pi_i), \tag{3.10}$$

which is a homomorphism instead of an isomorphism since given a Hamiltonian dynamics different quantum Hamiltonian operators can be deduced. This is similar with the fact that given a classical dynamics there can be different quantum versions.

[1] Note one can also use ψ_i and $\pi_i := i\psi_i^*$ instead. Position can be denoted by either q or x.

If the Hilbert space is infinite dimensional, e.g., $L_2(\mathbb{R}^3)$, the classical dynamical variables become scale field $\psi(x)$ and $i\psi^*(x)$, and the Hamilton's equations take the form for a scale field. Note that there is a significant difference between the field $\psi(x)$ here and a scale field $\phi(x)$ in classical wave mechanics, where the 'momentum' is the time-derivative of the field.

Quantum observable corresponds to special kind of vector field, such as the Kählerian function, and the Poisson bracket between two vector fields is equivalent to the expectation value of commutator

$$\{F,G\} = -i\langle\psi|[\hat{F},\hat{G}]|\psi\rangle, \tag{3.11}$$

with observable $\hat{F},\hat{G}$ and the corresponding vector fields F,G.

The GQM above provides a hidden 'classical' picture of quantum dynamics in terms of Hamiltonian dynamics of coupled classical particles. However, there also exist many other bases hence other collection of hidden particles dynamics, which are equivalent to each other via unitary basis transformations. This is due to the extra Riemannian form for QM, which is absent for classical case, and related to the non-commutativity (or complementarity) of quantum operators. This also implies that a quantum dynamics may arise from a set of Hamilton dynamics such that the Riemannian form is respected.

3.1.2 Wigner function

Question 53. *How to define Wigner functions?*

Phase space formula based on Wigner function can be viewed as a variation of geometric quantum mechanics. Instead of using expectation values, it uses a particular Wigner function which in a sense is also an average over states. Note that Wigner functions are not unique.

We denote $W(\hat{A})$ as a Wigner map that maps to the Wigner function. For infinite-dimensional case, the Wigner function of any observable $\hat{A}$ is

$$W(\hat{A}) = a(x,p) = \int ds e^{ips/\hbar}\langle x-s/2|\hat{A}|x+s/2\rangle, \tag{3.12}$$

and its inverse map is the so-called Weyl transformation

$$\langle x|\hat{A}|y\rangle = \int dp e^{ip(x-y)/\hbar} a((x+y)/2,p). \tag{3.13}$$

It uses the basis $\{|x\rangle\}$ and $\{|p\rangle\}$ from observable $\hat{x}$ and $\hat{p}$.

For finite-dimensional case, Wigner functions can also be defined. The analog of $\hat{x}$ and $\hat{p}$ are the Heisenberg-Weyl operators X^i and Z^j (Chapter 2). A quantum state ρ is expanded as a set of coefficients $w_{ij} = \mathrm{tr}(\rho X^i Z^j)$.

For density operator, the usual Wigner function is

$$w(x,p) = \int ds e^{ips/\hbar} \langle x - s/2|\rho|x + s/2\rangle. \tag{3.14}$$

It satisfies

$$\int dx \int dp w(x,p) = \mathrm{tr}\rho, \tag{3.15}$$

$$\int dx w(x,p) = \langle p|\rho|p\rangle, \tag{3.16}$$

$$\int dp w(x,p) = \langle x|\rho|x\rangle, \tag{3.17}$$

$$\int dx \int dp w(x,p) a(x,p) = \mathrm{tr}\rho\hat{A}. \tag{3.18}$$

In particular, the moments are

$$\langle x^m p^n\rangle = \int dx \int dp \hat{x}^m \hat{p}^n w(x,p). \tag{3.19}$$

The important fact is that all information of state is contained by the set of moments.

Question 54. *It turns out Wigner function can be negative. What does this mean?*

It seems $w(x,p)$ acts like probability, yet this is not ture. Wigner function $w(x,p)$ can take negative value. Wigner function negativity is viewed as a quantum feature compared with classical physics. However, this is not correct. Wigner function negativity is only an apparent fact, and it can be modified to a positive-definite function. There are Glauber P representation and Husimi Q representation. The crucial fact is that they cannot be viewed as standard probability distribution, since either they can be negative, or they do not describe independent events.

Further, we know that the distribution $\psi(x)$ satisfies Hamilton equation, how about the Wigner function? It turns out Wigner function satisfies the Moyal equation

$$\dot{w}(x,p,t) = -\{\{w(x,p,t), H(x,p)\}\} \tag{3.20}$$

for Hamilton $H(x,p)$ as the Wigner version of $\hat{H}$ and Moyal bracket

$$\{\{f,g\}\} := (f \star g - g \star f) = \{f,g\} + O(\hbar^2) \tag{3.21}$$

and $\star$ is known as Moyal product. This can be viewed as a variation of Hamilton equation. Also note that Moyal's approach is related to the $\hbar$-deformation quantization since the limit $\hbar \to 0$ leads to the classical case.

3.1.3 Quantization

Question 55. *What is quantization?*

Quantization is an operation which maps functions on a space to operators acting on the corresponding Hilbert space $\mathscr{H}$. For $\mathbb{R}^3$, $\mathscr{H} = L_2(\mathbb{R}^3)$, the set of square-integrable functions on it (ignoring the normalization condition). The dynamical variables (q,p) are mapped to operators $(\hat{q},\hat{p} := -i\frac{\partial}{\partial q})$, which are both hermitian observable, and the Hamiltonian H is mapped to quantum Hamiltonian operator $\hat{H}$. The Schrödinger equation takes the form

$$i|\dot{\psi}\rangle = \hat{H}|\psi\rangle, \text{ or } i\dot{\psi}(x) = \hat{H}\psi(x), \tag{3.22}$$

with $|\psi\rangle = \int dx\psi(x)|x\rangle$, and $|x\rangle, |\psi\rangle \in \mathscr{H}$, $\psi(x) \in \mathscr{H}^*$, with $\mathscr{H}^*$ as the dual of $\mathscr{H}$. Here $\hat{H}$ is diagonal in the basis $\{|x\rangle\}$.

With quantization and hamiltonization, we can further introduce the second quantization and the quantum (or operator-valued) phase space. The Ehrenfest theorem reveals that a classical description of quantum dynamics can be obtained by taking the expectation value of operators on a quantum state. From this point of view, the dynamics (3.7) is classical in the sense that the Hamiltonian is the expectation values of the quantum operators, which is consistent with the Ehrenfest theorem. It turns out we can generalize the form above to the quantum case, or termed as an operator-valued case wherein the expectation is absent. That is, we can further quantize the Hamiltonian dynamics on Σ to a new quantum dynamics.

The quantization map takes the form

$$\text{Quantization} : \Sigma \to L_2(\Sigma) : H \mapsto \mathbb{H}, (\psi_i, \pi_i) \mapsto (\hat{\psi}_i, \hat{\pi}_i), \tag{3.23}$$

and the new quantum dynamics is

$$i|\dot{\chi}(\psi)\rangle = \mathbb{H}|\chi(\psi)\rangle, \text{ or } i\dot{\chi}(\psi) = \mathbb{H}\,\chi(\psi), \tag{3.24}$$

with $\chi(\psi) \in L_2(\Sigma)$, and $|\chi(\psi)\rangle = \int d\psi\chi(\psi)|\psi\rangle$, and

$$\mathbb{H} = \hat{\psi}^\dagger \hat{H}\hat{\psi} = \sum_{i,j=1}^{d} \hat{\psi}_i^\dagger H_{ij}\hat{\psi}_j, \tag{3.25}$$

with $\hat{\psi} = \sum_{i=1}^{d} \hat{\psi}_i|i\rangle$, and $\mathbb{H}, \hat{\psi}_i : L_2(\Sigma) \to L_2(\Sigma)$. Note H_{ij} could be an operator so the order in $\mathbb{H}$ cannot be changed generally. We may call $\hat{\psi}$ as field operator. The form $d\psi$ represents a measure on space $L_2(\Sigma)$. As is well known, there is no Lebesgue measure on infinite-dimensional Hilbert space; however, there can be a Borel measure, which is indeed the Wiener measure. Also, employing Fock space $\mathscr{F}$ and particle-number basis or coherent-state basis (studied below) instead of space $L_2(\Sigma)$, a measure and integral can be well defined.

The formalism above provides a proper foundation for second quantization. We can view $\hat{\psi}_i$ as the analog of ladder operator $\hat{a}$ of a position particle and then $\sum_i \hat{\psi}_i^\dagger \hat{\psi}_i$ is the particle (or excitation) number operator. In order to describe the boson or fermion, the Fock space $\mathscr{F}$ is often introduced, which is isomorphic to $L_2(\Sigma)$, since their dimensions are the same. As the result, we can study the dynamics in $\mathscr{F}$ instead. In Fock space, there exists the particle number basis $\{|n\rangle\}$, and state $|\chi\rangle$ (omitting the ψ index) can be expanded in this basis as

$$|\chi\rangle = \sum_{n=0}^{\infty} \chi_n |n\rangle, \ \ |\chi\rangle \in \mathscr{F}. \tag{3.26}$$

Question 56. *Schrödinger vs. Heisenberg, how equivalent?*

Next we study the dynamics of field operators $\hat{\psi}_i$. We find that the Schrödinger equation and Heisenberg equation are equivalent, and furthermore, the equivalence is closely related to the boson/fermion statistics. Starting from Heisenberg equation

$$i\dot{\hat{\psi}}_i = [\hat{\psi}_i, \mathbb{H}], \ \ \& \ \ i\dot{\hat{\psi}} = [\hat{\psi}, \mathbb{H}], \tag{3.27}$$

we find if the operator $\hat{\psi}_i$ is bosonic or fermionic

$$\text{Boson: } [\hat{\psi}_i, \hat{\psi}_j] = 0, [\hat{\psi}_i^\dagger, \hat{\psi}_j^\dagger] = 0, [\hat{\psi}_i, \hat{\psi}_j^\dagger] = \delta_{ij}, \tag{3.28a}$$

$$\text{Fermion: } \{\hat{\psi}_i, \hat{\psi}_j\} = 0, \{\hat{\psi}_i^\dagger, \hat{\psi}_j^\dagger\} = 0, \{\hat{\psi}_i, \hat{\psi}_j^\dagger\} = \delta_{ij}, \tag{3.28b}$$

then

$$i\dot{\hat{\psi}}_i = [\hat{\psi}_i, \mathbb{H}] = \sum_{j=1}^{d} H_{ij}\hat{\psi}_j, \ \ \& \ \ i\dot{\hat{\psi}} = [\hat{\psi}, \mathbb{H}] = \hat{H}\hat{\psi}. \tag{3.29}$$

The Kronecker delta δ_{ij} is changed to Dirac delta function $\delta(i-j)$ for infinite-dimensional case. The equation (3.29) above is the operator-valued version of $i\dot{\psi}_i = \sum_{j=1}^{d} H_{ij}\psi_j$ and $i|\dot{\psi}\rangle = \hat{H}|\psi\rangle$.

Conversely, if we assume Eq. (3.29) first and then we can derive the boson/fermion statistics (3.28). That is, the derivation of statistics does not refer to spin or special relativity. Also, the equivalence between Schrödinger equation and Heisenberg equation above is more fundamental than the common one, which is about the expectation value of observable in two different 'pictures'.

Furthermore, there is also an operator-valued version of the Hamiltonian dynamics (in Eq. (3.9)). We find that

$$\frac{\partial \mathbb{H}}{\partial \hat{\psi}_i} = -\dot{\hat{\pi}}_i \ , \ \frac{\partial \mathbb{H}}{\partial \hat{\pi}_i} = \dot{\hat{\psi}}_i. \tag{3.30}$$

Then the following equation holds

$$i\dot{\hat{\psi}} = [\hat{\psi}, \mathbb{H}] = \hat{H}\hat{\psi} = \frac{\partial \mathbb{H}}{\partial \hat{\psi}^\dagger}. \tag{3.31}$$

This equation looks simple, while it may have deep physical foundations. The first equality is Heisenberg equation, the second one is Schrödinger equation, and the third one is Hamilton equation. The symplectic form is $\omega = d\hat{\psi} \wedge d\hat{\pi} = \sum_i d\hat{\psi}_i \wedge d\hat{\pi}_i$, and the triplet $(\mathscr{F}, \omega, \mathbb{H})$ is the *quantum (or operator-valued) Hamiltonian dynamics system*.

For hermitian operator $\hat{O}$ acting on Hilbert space $\mathscr{H}$, define the observable acting on $\mathscr{F}$ as

$$\mathbb{O} := \hat{\psi}^\dagger \hat{O} \hat{\psi}. \tag{3.32}$$

A quantum Poisson structure can also be defined as in the case of classical mechanics. For operators $\mathbb{F}$ and $\mathbb{G}$ acting on $\mathscr{F}$, the quantum Poisson bracket is defined as

$$\{\mathbb{F}, \mathbb{G}\}_Q := \frac{\partial \mathbb{F}}{\partial \hat{\psi}} \frac{\partial \mathbb{G}}{\partial \hat{\pi}} - \frac{\partial \mathbb{G}}{\partial \hat{\psi}} \frac{\partial \mathbb{F}}{\partial \hat{\pi}}. \tag{3.33}$$

It is straightforward to check that $[\mathbb{O}, \mathbb{H}] = i\{\mathbb{O}, \mathbb{H}\}_Q$, and more generally,

$$[\mathbb{F}, \mathbb{G}] = i\{\mathbb{F}, \mathbb{G}\}_Q. \tag{3.34}$$

We see that the commutator $[,]$ is equivalent to quantum Poisson bracket $\{,\}_Q$, which means that the commutator plays the roles of quantum Poisson bracket. This is different from the common view that the expectation value of commutator corresponds to the classical Poisson bracket, also see Eq. (3.11).

The quantum Poisson bracket also satisfies the Jacobi identity

$$\{\mathbb{F}, \{\mathbb{G}, \mathbb{E}\}\}_Q + \{\mathbb{E}, \{\mathbb{F}, \mathbb{G}\}\}_Q + \{\mathbb{G}, \{\mathbb{E}, \mathbb{F}\}\}_Q = 0, \tag{3.35}$$

which is equivalent to the Jacobi identity of the corresponding operators acting on Hilbert space; i.e. $[F, [G, E]] + [E, [F, G]] + [G, [E, F]] = 0$.

Next we consider observable dynamics in Heisenberg picture. Define time-dependent observable $\hat{O}_t := e^{i\hat{H}t} \hat{O} e^{-i\hat{H}t}$, and then $\mathbb{O} := \hat{\psi}_t^\dagger \hat{O} \hat{\psi}_t = \hat{\psi}_0^\dagger \hat{O}_t \hat{\psi}_0$. We find

$$i\dot{\mathbb{O}} = [\mathbb{O}, \mathbb{H}] = \hat{\psi}_t^\dagger [\hat{O}, \hat{H}] \hat{\psi}_t = \hat{\psi}_0^\dagger [\hat{O}_t, \hat{H}] \hat{\psi}_0, \tag{3.36}$$

from which we derive

$$\dot{\hat{O}}_t = -i[\hat{O}_t, \hat{H}], \tag{3.37}$$

which is the usual Heisenberg equation, assuming $\dot{\hat{O}} = 0$. Also, the Hamiltonian-Poissonian dynamics of field operator can be summarized as

$$\dot{\hat{\psi}} = \frac{\partial \mathbb{H}}{\partial \hat{\pi}} = -i\hat{H}\hat{\psi} = -i[\hat{\psi}, \mathbb{H}] = \{\hat{\psi}, \mathbb{H}\}_Q, \tag{3.38a}$$

$$\dot{\hat{\pi}} = -\frac{\partial \mathbb{H}}{\partial \hat{\psi}} = i\hat{\pi}\hat{H} = -i[\hat{\pi}, \mathbb{H}] = \{\hat{\pi}, \mathbb{H}\}_Q. \tag{3.38b}$$

Question 57. *Is there an origin of quantum harmonic oscillators?*

We see that a field operator can be bosonic or fermionic. If we start from the space $\mathbb{C}$ and a trivial model h, then the quantized model is just $H = \hat{\psi}^\dagger \hat{\psi}$, which actually looks like a harmonic oscillator when $h = \omega$, the size of energy. When we start from $\mathbb{C}^d$ but with $d \to \infty$ and a trivial model, then the quantized model is $H = \int dx \hat{\psi}^\dagger(x)\hat{\psi}(x)$, i.e., an infinite collection of free harmonic oscillators. Nontrivial models will lead to interactions among them, e.g., jump terms $\hat{\psi}^\dagger(x)\hat{\psi}(y)$ may exist.

For a single harmonic oscillator, the field operator $\hat{\psi}$ is the usual operator $\hat{a}$ that we are familiar with. Now we can see a difference between fermion and boson. For the model $H = \hat{a}^\dagger \hat{a}$, for boson we define $\hat{x} = \hat{a} + \hat{a}^\dagger$, $\hat{p} = i(\hat{a} - \hat{a}^\dagger)$, which are the position and momentum operators, then

$$H = \hat{x}^2 + \hat{p}^2 - 1/2, \tag{3.39}$$

since $[\hat{x}, \hat{p}] = i$. For fermion we define $\chi_1 = \hat{a} + \hat{a}^\dagger$, $\chi_2 = i(\hat{a} - \hat{a}^\dagger)$, which are so-called Majorana fermions, then

$$H = -i\chi_1\chi_2 + 1/2, \tag{3.40}$$

since χ_1^2 and χ_2^2 are constants.

3.2 Relativistic subtheory

It turns out special relativity can be put in the framework of quantum theory. It describes the coherence between vacuum and matter, i.e., the space itself needs to be treated in a quantum way. When matter moves fast in vacuum, the coherence between them will become evident: the evidence is that the speed of matter cannot exceed the speed of light in vacuum, which is a constant, c.

3.2.1 Special Relativity

The discovery of relativity was attributed to Poincaré, Lorentz, and Einstein. Einstein postulated the two principles:

- The Principle of Relativity: The laws by which the states of physical systems undergo change are not affected, whether these changes of state be referred to the one or the other of two systems in uniform translatory motion relative to each other.
- The Principle of Invariant Light Speed: Light in vacuum propagates with the speed c (a fixed constant, independent of direction) in at least one system of inertial coordinates, regardless of the state of motion of the light source.

Question 58. *How to derive Lorentz transformation?*

Suppose

$$\begin{cases} x' &= ax - bvt, \\ t' &= ct - dx/v, \end{cases} \tag{3.41}$$

a, b, c, d are parameters to be derived. The following three conditions are required:

- (i) The inverse transformation satisfies similar relations, except that the velocity v changes to $-v$. As the result,

$$\begin{cases} x &= ax' + bvt, \\ t &= ct' + dx'/v, \end{cases} \tag{3.42}$$

  From this, we find $a = c$, $ac - bd = 1$.

- (ii) The case for light $x = ct$ is equivalent to $x' = ct'$. Then we find $bv^2 = dc^2$.

- (iii) The velocity in x' frame is zero $v' = dx'/dt' = 0$. We get $a = b$.

From above conditions, we find $a = b = c = \gamma \equiv 1/\sqrt{1-\beta^2}$, $d = \gamma\beta^2$, $\beta \equiv v/c$. Then, the Lorentz transform Λ is

$$\begin{pmatrix} x' \\ ct' \end{pmatrix} = \Lambda \begin{pmatrix} x \\ ct \end{pmatrix}, \Lambda = \gamma \begin{pmatrix} 1 & -\beta \\ -\beta & 1 \end{pmatrix}. \tag{3.43}$$

When $\beta = 0$, the matrix Λ reduces to identity. The transformation forms the Lorentz group. The vector (x, ct) is in the covariant form, which is suitable for field dynamics since both x and t are treated on the equal footing as external parameters.

Lorentz group is the generalized orthogonal group $O(3,1)$. It is not compact and is not connected. It has four connected components which are not simply connected, but rather doubly connected. The connected component of identity is often called the restricted Lorentz group $SO^+(3,1)$, which preserves the direction of time and contains proper rotation (with determinant 1), and is six dimensional. It contains $SO(3)$ as a subgroup for rotations. It is isomorphic to the projective special linear group PSL(2,$\mathbb{C}$), which is, in turn, isomorphic to the Möbius group, the symmetry group of conformal geometry on the Riemann sphere. The isomorphism is easy to see as a Möbius map $f(z) = \frac{az+b}{cz+d}$ uniquely maps to an invertible matrix $[a, b; c, d]$, which can be decomposed into rotations and boosts. This leads to the spacetime qubit.

Now, we study the kinetic observable energy E and momentum $\vec{p}$ in covariant form. Suppose the transform is

$$\begin{pmatrix} cp' \\ E' \end{pmatrix} = \gamma \begin{pmatrix} 1 & -\beta \\ -\beta & 1 \end{pmatrix} \begin{pmatrix} cp \\ E \end{pmatrix}. \tag{3.44}$$

In x' frame, it is required $p' = 0$, $E' = mc^2$, then

$$E = \gamma mc^2, \vec{p} = \gamma m\vec{v}. \tag{3.45}$$

The Einstein relation follows

$$E^2 - p^2c^2 = m^2c^4. \tag{3.46}$$

Question 59. *How to interpret Einstein relation?*

In equation (3.45), the energy is positive, yet from equation (3.46), the energy can take two values $E_\pm = \pm\sqrt{p^2c^2 + m^2c^4}$. Which equation is more fundamental? The question is quite deep since it questions the meaning of energy and mass. The standard way is to interpret the negative energy as for *anti-particle* with opposite charge. Then the new question arises: what is charge? In modern particle physics, a neutral particle can also have anti-particle. Why in the Lorentz transform there is no charge d.o.f? Further, if there exists anti-particle and anti-matter, why the symmetry between matter and anti-matter is broken? In gravity, does anti-particle attract particle? At present, we cannot answer the questions above. Nowadays, the standard interpretation is that: there exists anti-matter, its mass is positive, yet its energy is opposite to particle, also anti-particle travels backwards in time. Thus, a particle with (cp, E) corresponds to the anti-particle with $(-cp, -E)$, and the Lorentz transform is the same for particle and anti-particle.

Question 60. *How to treat relativistic effect in a quantum way?*

The special relativity of a single particle reveals there is a fundamental limitation of its "freedom", that is to say, there exists a kind of internal coherence which depends on the velocity of the particle. The effect of special relativity can be treated quantally, it turns out the spacetime state can be expressed as one two-level system, i.e. qubit. Define the spacetime qubit as

$$\hat{s} = \begin{pmatrix} ct + z & x - iy \\ x + iy & ct - z \end{pmatrix} \tag{3.47}$$

The norm (determinant) of $\hat{s}$ is the proper distance $\|\hat{s}\| = c^2t^2 - r^2$. The Lorentz transform is a "determinant-preserving" completely-positive map. The qubit can be labeled on the Bloch ball, then light is a pure spacetime qubit forming the sphere, and massive particles live in the ball, the center of the ball is for particle at rest (the origin). The purity is $\mathrm{tr}\hat{s}^2/(\mathrm{tr}\hat{s})^2 = (1+n^2)/2$ for $n = r/ct$ as the Bloch vector.

A mass-qubit can also be defined as

$$\hat{m} = \begin{pmatrix} E - p_z c & p_x c - i p_y c \\ p_x c + i p_y c & E + p_z c \end{pmatrix} \tag{3.48}$$

The norm of $\hat{m}$ is the rest energy $\|\hat{m}\| = m^2c^4$.

Usually, relativistic quantum mechanics is treated as the extension of standard quantum mechanics. However, by the concept of spacetime qubit, the effect of relativity will be represented by a two-dimensional Hilbert space. There is a basis formed by $|\circ\rangle$ (anti-particle) and $|\bullet\rangle$ (particle).

3.2.2 Relativistic equations

Schrödinger equation is proposed for non-relativistic motions, but it sets the fundamental form of quantum equation

$$i|\dot{\psi}\rangle = H|\psi\rangle. \tag{3.49}$$

It turns out relativistic free motions, with Einstein's relation as its energy, all take this form but with different Hamiltonian and states. Different cases are labelled by the value of spins, for which the Dirac equation lies at the heart of these forms.

A wave propagating in a massive string is one standard example of a scalar field. The field equation is

$$\left(\frac{1}{v^2}\frac{\partial^2}{\partial t^2} - \nabla^2 + \xi\right)\phi(x) = 0, \tag{3.50}$$

where v is the propagation speed of the wave in the medium, ξ is mass-like parameter.

The covariant (relativistic) scalar field which modifies the above case is Klein-Gordon equation

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2 + \frac{m^2c^2}{\hbar^2}\right)\phi(x) = 0, \tag{3.51}$$

which describes spinless scalar particles. It can be viewed as special cases of Maxwell equation and Dirac equation.

Question 61. *How to express Klein-Gordon equation in a quantum form?*

We start from Einstein's relation $E^2 = p^2c^2 + m^2c^4$. Quantize it as $E \to i\hbar\frac{\partial}{\partial t}$, $p \to -i\hbar\nabla$, we get Klein-Gordon equation (3.51). Let $\phi(x) = \psi_1(x) + \psi_2(x)$, then each of the two satisfies Klein-Gordon equation. With the basis $|\bullet, \circ\rangle$, we define the Klein-Gordon state $|\Psi\rangle = \begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix}$, the up branch is for particle, the down for anti-particle. Suppose

$$\begin{cases} i\dot{\psi}_1 &= a\psi_1 + b\psi_2, \\ i\dot{\psi}_2 &= c\psi_2 + d\psi_1, \end{cases} \tag{3.52}$$

a, b, c, d are parameters to be derived. The above coupled equation is equivalent to Schrödinger equation with Hamiltonian $H = \begin{pmatrix} a & b \\ d & c\end{pmatrix}$. The four parameters need to satisfy two conditions: (1) Klein-Gordon equation; (2) eigenvalues of H are $E_\pm = \pm\sqrt{p^2c^2 + m^2c^4} = \left(a + c \pm \sqrt{(a-c)^2 + 4bd}\right)/2$. We find $c^2 - a^2 = (a+c)(d-b)$.

- $a = c \Rightarrow b = d \Rightarrow a = 0$, $b = \sqrt{p^2c^2 + m^2c^4} \Rightarrow$

$$H = \begin{pmatrix} 0 & \sqrt{p^2c^2 + m^2c^4} \\ \sqrt{p^2c^2 + m^2c^4} & 0\end{pmatrix}. \tag{3.53}$$

Diagonalize as $H_d = \sqrt{p^2c^2 + m^2c^4}I$. This is "relativistic Schrödinger form".

- $a=-c \Rightarrow E_{\pm}=\pm\sqrt{a^2+bd}$;
  - (i) $a=mc^2+p^2/2m$, $b=-d=p^2/2m \Rightarrow$

$$H=\begin{pmatrix} mc^2+p^2/2m & p^2/2m \\ -p^2/2m & -mc^2-p^2/2m \end{pmatrix}. \tag{3.54}$$

    This is called the "Hamiltonian form", yet it is not hermitian. Diagonalized form is called "Feshbch-Villars representation". $\dot{\phi}=-imc^2(\psi_1-\psi_2)$. This form has a drawback that mass m cannot be reduced to zero.
  - (ii) $a=mc^2$, $b=d=p^2/2m \Rightarrow$

$$H=\begin{pmatrix} mc^2 & pc \\ pc & -mc^2 \end{pmatrix}. \tag{3.55}$$

    $\dot{\phi}=-imc^2(\psi_1-\psi_2)-ipc(\psi_1+\psi_2)$. This can be named as "Dirac form" since it can be derived from Dirac equation.

The reason that there could be different forms of Hamiltonian is both values of ϕ and $\dot{\phi}$ are required. The above forms yield different $\dot{\phi}$, the Dirac form is preferred. Thus, the Klein-Gordon equation is written in Dirac form as

$$i\hbar|\dot{\Psi}\rangle=\begin{pmatrix} mc^2 & pc \\ pc & -mc^2 \end{pmatrix}|\Psi\rangle. \tag{3.56}$$

One consequence is that if a particle is accelerated, it will eventually becomes a superposed state $|+\rangle$ of particle and anti-particle, i.e. a flash, just like light whose antiparticle is itself.

Question 62. *How to derive Klein-Gordon equation from Dirac equation?*

Dirac equation describes the motion of a massive charged spin-half particle, e.g. electron, it takes the form

$$i\hbar|\dot{\Psi}\rangle=(\vec{\alpha}\cdot\vec{p}c+\beta mc^2)|\Psi\rangle, \tag{3.57}$$

for $|\Psi\rangle$ in the basis of $|x\rangle$, $|\bullet,\circ\rangle$, and spin $|\uparrow,\downarrow\rangle$, with Hamiltonian

$$H_D=\begin{pmatrix} mc^2 & 0 & cp_z & c(p_x-ip_y) \\ 0 & mc^2 & c(p_x+ip_y) & -cp_z \\ cp_z & c(p_x-ip_y) & -mc^2 & 0 \\ c(p_x+ip_y) & -cp_z & 0 & -mc^2 \end{pmatrix}, \tag{3.58}$$

and Dirac matrices

$$\beta\equiv\alpha^0\equiv\gamma^0=Z\otimes\mathbb{1}=\begin{pmatrix} \mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix}, \tag{3.59}$$

$$\alpha^1=X\otimes X=\begin{pmatrix} 0 & X \\ X & 0 \end{pmatrix}, \alpha^2=X\otimes Y=\begin{pmatrix} 0 & Y \\ Y & 0 \end{pmatrix}, \alpha^3=X\otimes Z=\begin{pmatrix} 0 & Z \\ Z & 0 \end{pmatrix}. \tag{3.60}$$

The covariant Dirac matrices are defined as

$$\gamma^\mu = \gamma^0\alpha^\mu, \gamma^1 = \begin{pmatrix} 0 & X \\ -X & 0 \end{pmatrix}, \gamma^2 = \begin{pmatrix} 0 & Y \\ -Y & 0 \end{pmatrix}, \gamma^3 = \begin{pmatrix} 0 & Z \\ -Z & 0 \end{pmatrix}. \tag{3.61}$$

There exists one additional matrix $\gamma^5 = X \otimes \mathbb{1}$.

In Dirac equation, there is a 'spin-orbit' coupling, i.e., the momentum and spin are coupled dynamically. As the result, independent observation of each of them is not proper. A quantity representing the coupling property of them is called "helicity". The spin operator is defined as $\vec{S} = \frac{1}{2}\mathbb{1} \otimes \sigma$. Helicity is defined as

$$h := \vec{S} \cdot \vec{p}/|p|. \tag{3.62}$$

Klein-Gordon equation can be derived by the following trick: define $W = (p_x X + p_y Y + p_z Z)/|p|$, then $W^2 = \mathbb{1}$. H_D can be reduced as $H_D = (mc^2 Z + |p|cX) \otimes \mathbb{1}$, the identity represents the information of spin which is ignored, and the rest part is just the Klein-Gordon Hamiltonian. In short, Klein-Gordon equation is a spinless or 1D Dirac equation.

Now we study the spin-1 case, which is nothing but the traditional Maxwell equation for photon

$$\nabla \cdot \vec{E} = \frac{\rho}{\varepsilon}, \tag{3.63}$$

$$\nabla \times \vec{B} - \frac{1}{c^2}\dot{\vec{E}} = \mu\vec{J}, \tag{3.64}$$

$$\nabla \cdot \vec{B} = 0, \tag{3.65}$$

$$\nabla \times \vec{E} + \dot{\vec{B}} = 0. \tag{3.66}$$

Introduce 4-vector forms $J^\mu = (\rho, \vec{J})$, $J_\mu = (\rho, -\vec{J})$, $A^\mu = (\phi, \vec{A})$, $A_\mu = (\phi, -\vec{A})$, and we know $\vec{B} = \nabla \times \vec{A}$, $\vec{E} = -\nabla\phi - \dot{\vec{A}}$. The field tensor is $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ with

$$F_{\mu\nu} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{pmatrix}. \tag{3.67}$$

The norm of it is $\|F_{\mu\nu}\| = (\vec{B} \cdot \vec{E})^2$, and $F_{\mu\nu}F^{\mu\nu} = 2(B^2 - E^2)$. The eigenvalues of $iF_{\mu\nu}$ are $\pm|E|$, $\pm|B|$. The Lagrangian density takes the form

$$\mathscr{L}_{EB} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + J^\mu A_\mu. \tag{3.68}$$

The Maxwell equation becomes

$$\partial_\nu F^{\mu\nu} = J^\mu, \tag{3.69}$$

$$\varepsilon_{\mu\nu\rho\gamma}\partial^\rho F^{\mu\nu} = 0, \tag{3.70}$$

$$\partial_\mu J^\mu = 0. \tag{3.71}$$

The last equation is the continuity equation.

Question 63. *How to express Maxwell equation in the quantum form?*

Without sources terms, the Maxwell Hamiltonian is

$$\hat{H}_{\rm EB} = i \begin{pmatrix} 0 & -p_x & p_z \\ p_x & 0 & -p_y \\ -p_z & p_y & 0 \end{pmatrix}, \tag{3.72}$$

and the quantum state $|\psi\rangle$ is three-component with each as $E_l + iB_l$, $l = x, y, z$. (check this!) This form is consistent with the fact that Maxwell equation can be written as a Hamiltonian system, with (E_l, B_l) as the dynamical variable. That is, (E_l, B_l) forms the coordinates of a phase space (relating to GQM). Accordingly, when the operator-valued quantum phase space form is defined, $\vec{E}$ and $\vec{B}$ will be quantized.

Different from electrons, photon is massless and its antiparticle is itself. However, as the case of electron, there is the d.o.f of helicity. The Hamiltonian above can be extended to include the information of helicity as

$$\hat{H}'_{\rm EB} = \begin{pmatrix} \hat{H}_{\rm EB} & \mathbf{0} \\ \mathbf{0} & -\hat{H}_{\rm EB} \end{pmatrix}, \tag{3.73}$$

$\mathbf{0}$ denoting block zero matrix. The quantum state is extended to $|\psi\rangle = \begin{pmatrix} \vec{E} + i\vec{B} \\ \vec{E} - i\vec{B} \end{pmatrix}$. The eigenvalues of $\hat{H}'_{\rm EB}$ are $0, p, -p$ each of degeneracy two, for helicity $h = \pm 1$.

Question 64. *Can Dirac equations be generalized to arbitrary spins?*

The answer is yes, but there are many ways. This problem is interesting since it might have a relation with Einstein's field equation of gravity, or, graviton which is spin-2.

Recall that Dirac equation takes the form $(\gamma^\mu \partial_\mu - m)\psi = 0$. The Bargmann-Wigner method generalizes the spinor ψ to a tensor Ψ, which can be viewed as product states of spinors. The tensor $\Psi_{n_1 n_2 \cdots n_{2s}}$ has 4^{2s} component for each n_i having four values. Further, the tensor Ψ is defined to be symmetric under permutation of index n_i, and then the number of independent component reduces to $2(2s+1)$. For instance, it is 2 for Klein-Gordon, 4 for Dirac, and 6 for Maxwell, agree with the forms before. It is 8 for spin-3/2 case, and 10 for spin-2 case! Bargmann-Wigner equation can be presented for each index

$$(\gamma^\mu \partial_\mu - m)_{nn'} \Psi_{n' \ldots} = 0, \forall n. \tag{3.74}$$

It includes many equations as special cases. For spin-1, it leads to the Proca equation which is a massive Maxwell equation, which can also be written as Kemmer equation.

Question 65. *How to express Einstein gravity equation in the quantum form?*

So far this is still an open problem. Eienstein field equation is

$$G_{ab} = \kappa T_{ab}, \tag{3.75}$$

with $\kappa = -\frac{8\pi G}{c^4}$ and

$$G_{ab} = R_{ab} + (\Lambda - R/2)g_{ab}, \tag{3.76}$$

for Ricci tensor R_{ab} and scalar $R = g^{ab}R_{ab}$. The energy-momentum tensor T_{ab} is the source (analog to charge and current) generating the gravitational field, which is equivalent to the spacetime geometry G_{ab} by the equivalence principle.

According to quantum field theory, graviton is a spin-2 boson. In Einstein equation there are 10 independent components in its tensor, and this agrees with the number of independent fields in Bargmann-Wigner equation. The quantum state $|\psi\rangle$ contains components from the Christoffel symbol Γ. The Hamiltonian contains two blocks for the two helicity graviton, each block contains 20 elements. The 10 independent variable in T_{ab} forms the nonlinear term.

There exists a linearized (or weak) form, called Bronshtein equation, which can be treated as spin-2 Bargmann-Wigner equation. Compared with Maxwell equation for electric and magnetic fields

$$\dot{E}_i = \varepsilon_{ikl}\partial_{x_k}B_l,\ \dot{B}_i = -\varepsilon_{ikl}\partial_{x_k}E_l,\ \partial_{x_i}E_i = 0,\ \partial_{x_i}B_i = 0, \tag{3.77}$$

the Bronshtein equation for electric-like and magnetic-like fields are

$$\dot{E}_{ij} = \varepsilon_{ikl}\partial_{x_k}B_{lj},\ \dot{B}_{ij} = -\varepsilon_{ikl}\partial_{x_k}E_{lj},\ \partial_{x_i}E_{ij} = 0,\ \partial_{x_i}B_{ij} = 0, \tag{3.78}$$

for E_{ij} and B_{ij} as symmetric traceless tensors defined by curvature tensor R.

3.3 Semi-classical subtheory

Before we discuss semi-classical subtheory, let's make a note on semi-quantum ones.

Question 66. *Are there semi-quantum subtheories?*

It turns out lots of quantum systems are actually semi-quantum, i.e., not fully but mainly quantum. The Schrödinger equation and Pauli equation are semi-quantum, for instance. The Pauli equation is a non-relativistic limit of Dirac equation. With external fields, even Dirac equation is semi-quantum since the field is treated classically, the precise interaction between electron and photons is not explicit. The Schrödinger equation is the non-relativistic limit of Klein-Gordon equation, or spinless Dirac equation. In Schrödinger equation, external potentials V can be included usually, which is classical since the potentials are not treated in quantum ways. This applies to situations when the external system is "huge" such that the back action on itself is tiny. The Born-Oppenheimer approximation is an example by treating nuclei

classically. The Lindblad equation is semi-quantum since the bath is treated classically. Many quantum field theories are also semi-quantum since, in addition to the free theories, semi-quantum additional terms are added without a quantum account of external d.o.f. For instance, in scalar field theory, terms like $\cos\Phi$ or Φ^4 for a field Φ, which is analog to external potentials in Schrödinger equation, is semi-quantum since the cause of it, which are those external d.o.f, are not included in the dynamics. In general, a fully quantum model between two quantum systems H_1 and H_2 are

$$H = H_1 + H_2 + H_I \tag{3.79}$$

for an interaction term H_I. When one of the systems are treated classically so that its dynamics is trivial, say H_2, the model will reduce to

$$H = H_1 + V \tag{3.80}$$

for an effective potential V. Semi-quantum, usually just called quantum, models are extremely useful since it reduces the fully quantum problems, which are harder, to much more tractable problems for various situations.

Now we discuss semi-classical subtheories, which are of distinct natures from semi-quantum ones. They are developed by 'classical' theorists who prefer to think quantum theory in classical ways. This can be done but not always. If you think quantum theory as operators or matrices, while classical theory as numbers, then matrices do not always behave like numbers. But to identify 'the unique quantum features' is difficult and cannot be done in principle. Below we survey some of the well-studied semi-classical subtheories.

3.3.1 Bohmian mechanics

Question 67. *What is the power of Bohmian mechanics?*

Bohmian mechanics (BM) is motivated by the quantum-classical distinction. It shows a 'classical' picture of the double-slit interference, hence provides 'more' than the standard QM. However, this picture is redundant. It cannot explain the origin of coherence. BM is usually used in hydrodynamic numerical simulation to view the trajectories of a set of identical particles, and this is used to provide an intuitive understanding of the underlying quantum dynamics.

BM is usually viewed as a hidden-variable theory (HVT), which is a post-quantum theory, but this is not correct. HVT requires that the probability can be derived from a deterministic dynamics, which cannot be done and shall not be pursued. However, in BM this is not the case. In BM there are two basic equations, one is the quantum equation, and the other is the equation of motion (for position). In BM a particle has a position, whose dynamics is affected by a so-called "quantum potential", which is determined by the quantum equation. That is to say, the position of a particle is determined by the wavefunction, while a HVT requires the opposite.

BM is a *semi-classical* theory. It can be derived from the Ehrenfest theorem. Ehrenfest theorem draws the connection between classical mechanics and quantum mechanics, namely, classical dynamics is the average of quantum dynamics. From Heisenberg equation of arbitrary observable A and take the average on a state $|\psi\rangle$, it holds

$$\frac{d}{dt}\langle\psi|A|\psi\rangle = \frac{1}{i\hbar}\langle\psi|[A,H]|\psi\rangle + \langle\psi|\frac{\partial A}{\partial t}|\psi\rangle. \tag{3.81}$$

For the Hamiltonian $H = p^2/2m - V(x)$, this reduces to

$$m\frac{d}{dt}\langle\psi|x|\psi\rangle = \langle\psi|p|\psi\rangle, \tag{3.82}$$

$$\frac{d}{dt}\langle\psi|p|\psi\rangle = \langle\psi|\frac{\partial V(x)}{\partial x}|\psi\rangle. \tag{3.83}$$

This is the classical version of the underlying quantum dynamics according to the phase space approach.

For a particle with position $x(t) := \langle\psi|x|\psi\rangle$, the Bohmian guiding equation is

$$m\dot{x}(t) = \mathrm{Im}\left(\frac{\nabla\psi}{\psi}\right)(x,t), \tag{3.84}$$

and this is from Ehrenfest theorem by noticing that $p = i\hbar\nabla$, and $\frac{\nabla\psi}{\psi} = \frac{\psi^*\nabla\psi}{\psi^*\psi}$.

We can see that the position in BM is not the usual position in classical mechanics, instead it is the average position of the quantum observable $\hat{x}$. However, from a different viewpoint, we can say the usual position in classical mechanics actually *is* the average position of $\hat{x}$.

As wavefunction is complex, so we can define $\psi = Re^{iS}$ with the real part $R^2 = \psi^*\psi$. This leads to a quantum hydrodynamics

$$\frac{\partial R}{\partial t} = -\frac{1}{2m}(R\nabla^2 S + 2\nabla R\cdot\nabla S), \tag{3.85}$$

$$\frac{\partial S}{\partial t} = -\frac{|\nabla S|^2}{2m} - V - Q \tag{3.86}$$

for the so-called *quantum potential*

$$Q := -\frac{\hbar^2}{2m}\frac{\nabla^2 R}{R}. \tag{3.87}$$

To include spin, we use two-component spinor field and inner product is defined as

$$\langle\psi|\phi\rangle = \psi_1^*\phi_1 + \psi_2^*\phi_2. \tag{3.88}$$

The guiding equation can be simply extended, and the quantum equation is the Pauli equation. However, this is not relativistic yet. The relativistic version of Pauli equation is the Dirac equation. So we can start from Dirac equation and apply Ehrenfest theorem to derive a Bohmian mechanics. In addition, BM has not been well developed for finite-dimensional Hilbert space and open-system dynamics.

3.3.2 Path integral

Path integral is a technique that merely uses the following simple relation

$$\sum_i |i\rangle\langle i| = \mathbb{1}, \tag{3.89}$$

which is known as the completeness relation given a basis $\{|i\rangle\}$ of a Hilbert space $\mathscr{H}$. However, this completeness relation is not trivial

- it encodes some geometry or topology feature of the space $\mathscr{H}$.
- the basis $\{|i\rangle\}$ is not unique, and can be overcomplete, or non-orthogonal etc.
- it introduces a pseudo dimension along the set of completeness relation.

We briefly introduce it for the model

$$H = \frac{p^2}{2m} + V. \tag{3.90}$$

The unitary evolution operator is

$$U(t_b, t_a) = e^{-i(t_b - t_a)H}. \tag{3.91}$$

The amplitude from a state $|x_a\rangle$ to $|x_b\rangle$ is

$$K(x_b t_b, x_a t_a) := \langle x_b | U(t_b, t_a) | x_a \rangle. \tag{3.92}$$

Now the evolution $U(t_b, t_a)$ can be sliced into many segments, say $N+1$, and denote $\varepsilon = t_n - t_{n-1} = (t_b - t_a)/(N+1)$. With the completeness relation at each time point

$$\int dx_n |x_n\rangle\langle x_n| = \mathbb{1},\ n = 1, \ldots, N, \tag{3.93}$$

the amplitude becomes

$$K(x_b t_b, x_a t_a) = \prod_n \int dx_n \prod_n K(x_n t_n, x_{n-1} t_{n-1}) \tag{3.94}$$

for

$$K(x_n t_n, x_{n-1} t_{n-1}) = \langle x_n | e^{-i\varepsilon H(t_n)} | x_{n-1} \rangle \tag{3.95}$$

and finally

$$K(x_b t_b, x_a t_a) = \prod_n \int dx_n \prod_n \int dp_n e^{i\mathscr{A}} \tag{3.96}$$

for

$$\mathscr{A} = \sum_{n=1}^{N+1} p_n (x_n - x_{n-1}) - \varepsilon H(p_n, x_n, t_n) = \int_{t_a}^{t_b} dt (p(t)\dot{x}(t) - H(p(t), x(t), t)) \tag{3.97}$$

known as "action." If $H = 0$, then $K(x_b t_b, x_a t_a) = \delta(x_b - x_a)$. Note here we need to use the completeness relation at different times. Sometime we do not need to use infinite many completeness relation.

Question 68. *Why path integral is viewed as a quantization method?*

By inserting the complete relation and taking inner product, the operators are converted into classical values, similar with the hamiltonization procedure, while the former deals with the evolution, and the later deals with the Hamiltonian (and also observable). Conversely, we can start from a classical system and then make 'path' integral or sum to get to the quantum case. Usually, we say classical mechanics is only one trajectory of the whole quantum dynamics. However, the path integral is very hard to define and this is still an open problem in mathematics.

Despite the difficulties of path integral, it has broad applications such as in quantum magnetism, quantum field theory, and statistical mechanics. For the space of a single spin, the path integral can reveal the geometry difference between integer and half-integer spins. In statistics, path integral can convert a D dimensional quantum system into a D+1 dimensional classical system based on its partition function.

3.3.3 Stabilizer formalism

In quantum computing, a popular class of states are known as stabilizer states. The standard setting is for several qubits, while it can also be generalized to qudits or continuous-variable cases. Here we focus on the case for qubits.

Recall that a qubit state can be specified by the observation values of Pauli σ^x and σ^z. In Heisenberg picture, we can keep track of the evolution of σ^x and σ^z, which is equivalent to study the evolution of the state. This can be generalized to the setting of stabilizer states, for which the observable are commuting stabilizers, which are products of Pauli operators.

When $S|\psi\rangle = |\psi\rangle$, we say S 'stabilize' the state $|\psi\rangle$. It is actually a symmetry. A stabilizer state is defined such that it is specified by a set of commuting stabilizers. For n qubits, the stabilizers form an abelian group, known as stabilizer group, $\mathscr{S}$.

Question 69. *How many generators there are for $\mathscr{S}$?*

It turns out the answer is n. This can be shown by defining the projector $P_0 = \prod_{j=1}^{n}(\mathbb{1}+S_j)/2^n$ for n generators S_j. It is clear that P_0 projects onto the stabilizer state, $|\psi\rangle$. Now we define more projectors as $P_x = \prod_{j=1}^{n}(\mathbb{1}+(-1)^{x_j}S_j)/2^n$ for $x = (x_1, x_2, ..., x_n)$ as a binary vector. From the commutation relation of Pauli matrices it is easy to see $P_x = g_x P_0 g_x$ for a g_x as a tensor product of Pauli matrices. Each P_x projects onto a unique state $|\psi_x\rangle$. Now as these projectors are orthogonal and complete, the set of states $\{|\psi_x\rangle\}$ form a basis of the n-qubit Hilbert space. This shows that indeed a n-qubit stabilizer state is uniquely defined by n stabilizer generators. For instance, the Bell state $|00\rangle + |11\rangle$ is stabilized by XX and ZZ, and other Bell states are stabilized similarly.

Furthermore, when a certain number of stabilizers are missing, we expect that the stabilizer state is replaced by a set of them, i.e., a subspace, usually known as a

'stabilizer code'. It is easy to see the stabilizer code encodes k qubits when there are $n-k$ stabilizers.

In general, a n-qubit state needs $O(2^n)$ parameters, yet a stabilizer state only needs the knowledge of n stabilizers. A simple way to specify a stabilizer state is to first encode each stabilizer by several bits, and in total by poly(n) bits. There is no need to construct the stabilizers themselves if not required. So in a sense, a stabilizer state is just a classical bit strings.

We can then consider evolution of stabilizer states. It turns out stabilizer states map to each other by the so-called Clifford circuits. A Clifford circuit is sequence of gates from the Clifford group

$$\mathscr{C}^{(2)} := \{U | UPU^\dagger \in \mathscr{C}^{(1)}, \forall P \in \mathscr{C}^{(1)}\}, \tag{3.98}$$

and $\mathscr{C}^{(1)}$ is the n-qubit Pauli group $\mathscr{P}_n := \mathscr{P}^{\otimes n}$, and $\mathscr{P} = \langle i\mathbb{1}, X, Z\rangle$ is the well-known Pauli group. Clifford gates include the phase gate, the Hadamard gate, the controlled-not gate, i.e., CNOT, and the controlled-phase gate, i.e., CZ. We see that Clifford gates map Pauli operators to Pauli operators, so map stabilizer states to each other. For a classical simulation, it is easy to see that Clifford circuits can be easily simulated. It is of course not universal for quantum computing. In addition, another nontrivial result is that Clifford circuits are not even universal for classical computing.

3.4 Quantum field theory

Question 70. *What is a quantum field?*

A quantum field is an operator acting on a Fock space, which is an infinite-dimensional Hilbert space. Quantum fields generalize our model of harmonic oscillator. A quantum field can be a quantized classical field, or a quantized quantum state. A classical field is a pile of spatially well-ordered particles $\{p_i, q_i\}$, $i = 1, 2, \ldots, n$. Instead of q_i, $\phi(x)$ is employed to label the position function of the field, $\pi(x)$ is employed to label the momentum function. The standard examples of classical fields are scalar field of harmonic oscillator, vector field of electromagnetic (EB) photonic "ether". Quantum fields can be used to describe quasiparticle, fluctuations, correlation, gapped and gapless properties of large systems. A (quantum) field operator creates a field from the vacuum with a certain amplitude, just like the creation operator creates a phonon or photon from the vacuum for a harmonic oscillator.

3.4.1 Bosonic and fermionic fields

When we quantize a quantum system, we see that it can lead to either bosonic or fermionic quantum fields. Which type shall it be? It turns out this is a nontrivial issue, and addressed by the spin-statistics theorem.

Question 71. *How to prove the spin-statistics theorem?*

The spin-statistics theorem states that half-integer spins are fermion, and integer spins are boson. This theorem can be proved from primary rules: e.g., locality, causality, and stability of the vacuum. A heuristic idea is that exchange of two particles is equivalent to the rotation of one particle by 2π. We know that the 2π rotation of half-spins will lead to a minus sign, while a positive sign for integer-spins; the theorem follows.

How about the spin-zero case? Originally, the spin-zero field is treated as boson. However, we realize that when spin is zero, i.e., either there is no spin or the spin is effectively decoupled from the dynamics, the spin-statistics theorem has to be altered. There are both spinless (i.e., spin-zero) fermion and boson, as demonstrated by the two forms of the Klein-Gordon equation.

The solution of Klein-Gordon equation is plane wave. In quantized form, the bosonic form can be written as

$$\hat{\phi}(x) = \int \frac{d^3k}{\sqrt{2\omega_k}} \left(\hat{a}_k e^{-i\vec{p}\cdot\vec{r}} + \hat{b}_k^\dagger e^{i\vec{p}\cdot\vec{r}} \right), \tag{3.99}$$

with $\omega_k = \sqrt{c^2p^2 + m^2c^4}/\hbar$, and the Klein-Gordon particle ($\hat{a}_k$) and anti-particle ($\hat{b}_k$) are boson,

$$[\hat{a}_k, \hat{a}_{k'}^\dagger] = [\hat{b}_k, \hat{b}_{k'}^\dagger] = (2\pi)^3 \delta^3(\vec{k} - \vec{k}'), \tag{3.100}$$

$$[\hat{a}_k, \hat{b}_{k'}^\dagger] = [\hat{b}_k, \hat{a}_{k'}^\dagger] = 0, \tag{3.101}$$

$$[\hat{\phi}(x), \hat{\phi}(y)] = [\hat{\pi}(x), \hat{\pi}(y)] = 0, \tag{3.102}$$

$$[\hat{\phi}(x), \hat{\pi}(y)] = i\delta^3(x-y). \tag{3.103}$$

The Hamiltonian becomes

$$H_{KG} = \int d^3k \omega_k \left(\hat{a}_k^\dagger \hat{a}_k + \hat{b}_k^\dagger \hat{b}_k + 1 \right). \tag{3.104}$$

We know that Klein-Gordon equation also describes spinless fermion, so its fermionic solution takes the form

$$\hat{\Psi}_{KG} = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_k}} \left(\hat{a}_k |\bullet\rangle e^{-i\vec{p}\cdot\vec{r}} + \hat{b}_k^\dagger |\circ\rangle e^{i\vec{p}\cdot\vec{r}} \right), \tag{3.105}$$

which is the 1D Dirac form, for fermionic $\hat{a}_k$ and $\hat{b}_k$. The Hamiltonian becomes

$$H = \int dk \omega_k (\hat{a}_k^\dagger \hat{a}_k + \hat{b}_k^\dagger \hat{b}_k - 1). \tag{3.106}$$

Dirac field is defined as

$$\hat{\Psi}_D = \int \frac{d^3k}{\sqrt{2\omega_k}} \left(\hat{c}_{1k} |\bullet\uparrow\rangle e^{-i\vec{p}\cdot\vec{r}} + \hat{c}_{2k} |\bullet\downarrow\rangle e^{-i\vec{p}\cdot\vec{r}} + \hat{d}_{1k}^\dagger |\circ\uparrow\rangle e^{i\vec{p}\cdot\vec{r}} + \hat{d}_{2k}^\dagger |\circ\downarrow\rangle e^{i\vec{p}\cdot\vec{r}} \right). \tag{3.107}$$

The Dirac Hamiltonian becomes

$$H_D = \int d^3k\omega_k(\hat{c}^\dagger_{1k}\hat{c}_{1k} + \hat{c}^\dagger_{2k}\hat{c}_{2k} - \hat{d}_{1k}\hat{d}^\dagger_{1k} - \hat{d}_{2k}\hat{d}^\dagger_{2k}). \tag{3.108}$$

which includes the fermionic Klein-Gordon form as a spinless case.

The quantized form of EB field

$$H_{EB} = \frac{1}{2}\int_V \varepsilon dV(E^2 + B^2) = \int d^3k \sum_{s=-1,1} \omega_k \left(\hat{a}^\dagger_{\vec{k}s}\hat{a}_{\vec{k}s} + \frac{1}{2}\right), \tag{3.109}$$

is well-known in quantum optics which describes photons with ladder operators $\hat{a}_{\vec{k}s}$ and conjugate.

We see above that all those free fields are collections of non-interacting harmonic oscillators with different spin s and energy ω_k. Scalar fields are common in particle physics, fermions such as electrons are common both in particle physics and condensed-matter physics, while bosons like photons are well studied in quantum optics.

Define the photon number operator $N_{ks} = \hat{a}^\dagger_{\vec{k}s}\hat{a}_{\vec{k}s}$, then $[N_{ks}, N_{k's'}] = 0$, which just manifests that different oscillators commute. The field Ψ_{EB} can be expanded in the basis of energy $\{|n\rangle\}$ which is Fock mode basis, and spin basis $\{|\uparrow,\downarrow\rangle\}$. Denote the basis as $|n\rangle = |n,k,s\rangle$, then

$$N_{ks}|n,k,s\rangle = n_{ks}|n,k,s\rangle, \tag{3.110}$$

$$\hat{a}^\dagger_{ks}|n,k,s\rangle = \sqrt{n_{ks}+1}|n+1,k,s\rangle, \tag{3.111}$$

$$\hat{a}_{ks}|n,k,s\rangle = \sqrt{n_{ks}}|n-1,k,s\rangle. \tag{3.112}$$

The state $|n,k,s\rangle$ can be viewed as the excitation from vacuum $|0\rangle$ as $|n,k,s\rangle = \frac{1}{\sqrt{n!}}(a^\dagger_{ks})^n|0\rangle$.

A quantized EB field can be in various state. For instance, a pulse from a laser is usually described as a coherent state

$$|\alpha\rangle = e^{-\frac{|\alpha|^2}{2}} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}}|n\rangle, \tag{3.113}$$

which is the eigenstate of the annihilation operator $\hat{a}|\alpha\rangle = \alpha|\alpha\rangle$, $|\alpha|^2 = \bar{n}$ is the average number of photons in the coherent state. We ignore the spin and modes for simplicity. The photon number distribution is $P(n) = |\langle n|\alpha\rangle|^2 = e^{-\bar{n}}\frac{\bar{n}^n}{n!}$, which is Poissonian. As the result, the coherent state sets at the boundary between quantum and classical, which has *minimal uncertainty*

$$\Delta X \Delta P = \hbar/2. \tag{3.114}$$

Besides, squeezed states also have minimal uncertainty by changing the values of the two while keeping their product. The state $|\alpha\rangle$ can also be viewed as being displaced

from the vacuum state by the displacement operator $D(\alpha)$ as $|\alpha\rangle = D(\alpha)|0\rangle$, and $D(\alpha) = e^{\alpha\hat{a}^\dagger - \alpha^*\hat{a}}$. Two different coherent states are quasi-orthogonal

$$\langle\beta|\alpha\rangle = e^{-(|\alpha|^2+|\beta|^2-2\alpha\beta^*)/2} \neq \delta(\alpha-\beta). \tag{3.115}$$

However, the coherent states can still form a complete set of basis, which is the Sudarshan-Glauber P representation, since

$$\frac{1}{\pi}\int |\alpha\rangle\langle\alpha| d^2\alpha = \mathbb{1}. \tag{3.116}$$

What is the average number of photons when the EB field is at temperature $\beta = 1/k_BT$? When temperature is not zero, the Hamiltonian needs to be modified to include the chemical potential μ. The state of EB field at equilibrium is $\rho = e^{-\beta H}$, then the average number of photon is

$$n_{ks} = \frac{\mathrm{tr}(\rho N_{ks})}{\mathrm{tr}(\rho)} = \frac{\sum_{n_{ks}=0}^{\infty} n_{ks} e^{-\beta n_{ks}(\omega_{ks}-\mu_{ks})}}{\sum_{n_{ks}=0}^{\infty} e^{-\beta n_{ks}(\omega_{ks}-\mu_{ks})}} \tag{3.117}$$

$$= \frac{1}{e^{\beta(\omega_{ks}-\mu_{ks})}-1}, \tag{3.118}$$

which is the Bose-Einstein distribution. The partition function is the product of the partition function for each mode (k,s), $Z = \Pi_{ks} Z_{ks}$.

Free field theory is trivial since we only obtain harmonic oscillators! In practice, there will be additional terms that induce nontrivial physics. For instance, term like $\int d^3x \int d^3x' \hat{\psi}^\dagger(x,t)\hat{\psi}^\dagger(x',t)V(x,x')\hat{\psi}(x,t)\hat{\psi}(x',t)$ describes interaction between fields. But note this term is semi-quantum instead of (fully) quantum since the interaction potential $V(x,x')$ does not account for entanglement between the interacting fields.

Question 72. *Is photon a real particle?*

A photon travels at the speed of light in vacuum, it does not have rest mass. We can imagine there is one particle called "ether" which is always oscillating harmonically, photon is just the package of energy of the oscillating ether. As the result, photon is not a particle, instead it is the standing wave of the propagation of the energy of the ether. The ether will never have one definite position and does not have dissipation in vacuum, yet in medium the travel speed of light will go down which means there is dissipation for the oscillation of ether. The study of ether dynamics will lead to new physics beyond standard quantum physics and explain behavior of vacuum. At present, the method based on photon instead of ether directly is convenient for applications.

Question 73. *What does '2nd quantization' really mean?*

Historically, 2nd quantization means the quantization for field, e.g., EB field, leading to quantum field theory (QFT), while 1st quantization means the quantization for particles. We now know that EB state and Dirac state are already quantum states, so the quantization is indeed a 2nd quantization since it quantizes a quantum system. The quantization of classical field is not a 2nd quantization, though. Classical fields are approximation of many-body classical systems, and quantization of it is a 1st quantization.

3.4.2 Topological fields

We see that a quantum field is spatially extensive, i.e., it describes a collection of particles. It turns out there are exotic models that only depends on the topology of these systems, and this is the subject of topological QFT.

Question 74. *What is topological in topological QFT?*

The topology is classical, namely, it refers to the topology on a classical manifold M, e.g., the spacetime or a torus. However, in topological QFT quantum observable is topological, i.e., they are topological invariants. These observable include order parameter, correlation function, partition function, etc.

There are basically two classes of TQFT:

1. Schwarz-type: explicit metric-independent. e.g., BF theory, Chern–Simons theory.

2. Witten-type: apparent metric-dependent. The action has an exact symmetry and relates to cohomology.

TQFT has crucial application for topological states of matter, i.e., topological order. Chern–Simons theory, with U(1) gauge group, can describe quantum Hall states, BF theory can describe topological insulator and quantum dimer models. TQFT can describe the low-energy sectors of a model, including ground states and a few excitations. The feature of a TQFT is best understood from its excitations: anyons.

Question 75. *Why there are anyons?*

Anyons are not elementary particles, since which can only be boson or fermion. Anyons are quasi-particles, or excitations, that can only exist depending on the underlying dimension. For point-like anyons, they only exist in 2D space. For string-like anyons, they only exist in 3D space. Due to the constraint of space, 'exchange' of anyons are not permutation, instead they are 'braiding' operations, forming the braid groups. Braiding operations are determined from the F tensor and R tensor of an anyonic system, which are again found from the modular operators S and T. The modular operator S is unitary, hence an anyonic system is often known as 'unitary

modular tensor category'. It determines the multiplicity N_{ab}^c in the anyon fusion rule by the Verlinde formula

$$N_{ab}^c = \sum_x S_{ax} S_{bx} S_{\bar{c}x} S_{1x}^{-1}, \tag{3.119}$$

for $S_{1x} = d_x/D$, for the so-called total dimension $D := \sqrt{\sum_a d_a^2}$, for d_a as the quantum dimension of anyon type a. For a collection of anyons $\{f_i\}$, the fusion rule is the Verlinde algebra

$$f_a \times f_b = \sum_c N_{ab}^c f_c, \tag{3.120}$$

which is a commutative associative $*$-algebra, $N_{ba}^c = N_{ab}^c$, $f_{\bar{a}} = f_a^\dagger$. The fusion operations act on the Hilbert spaces of these anyons. Abelian anyons have quantum dimension one, otherwise they are non-Abelian. The important application of anyons is topological quantum computing.

Chern-Simons (CS) theory can describe both non-Abelian and Abelian anyons, especially for quantum Hall states. CS model is a gauge model, i.e., it has local (gauge) symmetry (or redundancy). Usually, the gauge symmetry is a compact connected Lie group G, such as $U(1)$, $U(N)$, $O(N)$, and $Sp(2n)$. Given an odd-dimensional manifold M, such as spacetime, the connection A can be defined

$$A := A_\mu^a T^a dx^\mu,\ x^\mu \in M, \tag{3.121}$$

for T^a generators of G. It is a generalized notion of 'potential' in electromagnetism. Also a 'field tensor', or curvature F is

$$F := dA + A \wedge A \tag{3.122}$$

for exterior product $\wedge$ and differential operator d. In terms of components, $A \wedge A = \varepsilon^{\mu\nu} A_\mu A_\nu$ for anti-symmetric tensor $\varepsilon^{\mu\nu}$. The 2+1 CS action at 'level'-$k \in \mathbb{N}$ takes the form

$$S_k = \frac{k}{4\pi} \int_M \mathrm{tr}(A \wedge dA + \frac{2}{3} A \wedge A \wedge A). \tag{3.123}$$

It is so-defined since it is gauge invariant under G, up to $2\pi n$. The constant n is proportional to the winding number w of $U \in G$ in the gauge group

$$w := \frac{1}{24\pi^2} \int d^3x \varepsilon^{\mu\nu\rho} \mathrm{tr}(\Pi_\mu \Pi_\nu \Pi_\rho) \tag{3.124}$$

for $\Pi_\mu := U^\dagger \partial_\mu U$. This is actually the Wess–Zumino (WZ) term for a conformal field theory. When CS theory is defined on a manifold with boundary, then the gauge group operation will lead to the WZ term on the boundary, and this leads to the correspondence between CS theory in the bulk and WZW theory on the boundary, for the same group G (see later section).

Theory	Spectrum	Observable	Dimension	Order	Qubits
TFT	gapped	Wilson loop, braiding	2,3	TOP	fusion space
CFT	gapless	vertex operator	1,2,3	any	conformal tower
QFT	any	any	1,2,3	any	degeneracy

Table 3.1: Brief information of field theories.

The observable in CS theory is Wilson loops and their correlation functions which are gauge invariant. A Wilson loop on a loop γ takes the form

$$W(\gamma) := \mathrm{tr}\mathscr{P}e^{i\int_\gamma A}. \tag{3.125}$$

They are nothing but the logical operators for topological quantum computing. It is known that $SU(2)_k$ CS theory is universal for $k = 3$ and $k > 4$. The $k = 4$ case is not universal, in particular, since it describes Ising anyons.

3.4.3 Conformal fields

Question 76. *Why 'conformal' transformation?*

Conformal transformation (CT) acts on space and time which preserves 'angles'. It includes Lorentz transformation. CT is a usual map on a manifold widely used in geometry. In physics people realize long time ago that lots of equations or systems are conformal invariant, more general than scale invariant, and this applies to critical (gapless) systems at phase transition boundaries or points. CT has great impact in condensed matter physics and string theory, but not very much in other areas. The reason is that it first acts on space and time, but lots of problems do not deal with space. The other reason is that it deals with gapless systems which are very special.

Given space and time and the fields defined on it, we have to distinguish 'local' and 'global' transformation, just like local (gauge) and global symmetry. Usually local symmetry is larger than global ones, and local ones are often considered as redundancy while global ones are the real symmetry.

The global CT (gCT) is simpler than the local CT (lCT) since gCT forms a group while lCT does not. Note that as a group, gCT $\mathscr{C}$ is not compact, e.g., Lorentz group is not compact. For $\mathbb{R}^{3,1}$ with three space dimensions and one time dimension, the Lorentz group $\mathscr{L}$ is a generalized orthogonal group $O(3,1)$. With translation, it forms the Pincare group $\mathscr{P}$. We have

$$O(3) \subset \mathscr{L} \subset \mathscr{P} \subset \mathscr{C}. \tag{3.126}$$

As a group, gCT $\mathscr{C}$ contains translation P, Lorentz map L (rotation R and boost B), dilation (scaling) D, special conformal transformation (SCT) K. The SCT is

actually a composition of inversion I, translation, and again inversion. With the qubit form of spacetime, their actions are

$$L(\rho) = L\rho L,\ P(\rho) = \rho + \sigma,\ D(\rho) = \Delta\rho,\ I(\rho) = \rho^{-1}, \tag{3.127}$$

note generically ρ is invertible (eigenvalues $ct \pm r$).

If you are interested in the generators of them, they take the form

$$p_\mu = -i\partial_\mu,\ d = -ix^\mu\partial_\mu,\ l_{\mu\nu} = i(x_\mu\partial_\nu - x_\nu\partial_\mu),\ k_\mu = -i(2x_\mu x^\nu\partial_\nu - x^2\partial_\mu). \tag{3.128}$$

They satisfy

$$[p_\mu, p_\nu] = 0,\ [k_\mu, k_\nu] = 0,\ [d, p_\mu] = p_\mu,\ [d, k_\mu] = k_\mu,\ [k_\mu, p_\nu] = \eta_{\mu\nu}d - il_{\mu\nu}. \tag{3.129}$$

For CT there is also a difference between $d > 2$ and $d \leq 2$. The $d = 1$ case is not interesting for now. For $d = 2$, $x^2 = x_1^2 + x_2^2$ for $\mathbb{R}^2$ for two-dimensional classical fields, $x^2 = t^2 - x_3^2$ for $\mathbb{R}^{1,1}$ for one-dimensional quantum fields. 2D gCT is the set of Möbius map on Rienmann sphere $\mathbb{C} \cup \infty$, which contains holomorphic functions $f(z) = \alpha z + \beta$ ($z = it + x$) as a subset. For $d > 2$, gCT is isomorphic to a generalized orthogonal group.

2D lCT is well studied. Contrary to gCT, the lCT does not form an infinite-dimensional group. For 2D cases, the classical version of lCT is the Witt algebra, and the quantum version, as central extensions with central charges, is the Virasoro algebra. The generators of lCT is

$$L_n = -z^{n+1}\partial_z. \tag{3.130}$$

This is from an infinitesimal shift of z. To see the connection with holomorphic functions for gCT, we need to take the exponent of $\sum_{n=-\infty}^{\infty} L_n v_n$. From non-singularity at $z = 0$, we see $v_n = 0$ for $n \leq -2$. From its inverse, i.e. non-singularity at $z = \infty$, we see $v_n = 0$ for $n \geq 2$. So there are only components $n = 0, \pm 1$ for gCT, which is the Möbius group $\mathscr{M}$ isomorphic to PSL(2,$\mathbb{C}$) and the restricted Lorentz group $O^+(3,1)$, which is six-dimensional. So

$$\mathscr{M} \cong PSL(2,\mathbb{C}) \cong O^+(3,1). \tag{3.131}$$

This is easy to see as a Möbius map

$$f(z) = \frac{az + b}{cz + d} \tag{3.132}$$

can be mapped to a matrix $[a, b; c, d]$ with nonzero determinant. A Lorentz map can be written as an invertible matrix. Their representations are different, however. Namely, a Möbius map acts on complex numbers z, while an invertible matrix or Lorentz map acts on the full spacetime qubit which preserves determinant (proper distance).

From now on we will focus on the 2D lCT. The famous Virasoro algebra is

$$[L_n, L_m] = (n-m)L_{n+m} + \frac{c}{12}(n^3 - n)\delta_{n+m,0} \tag{3.133}$$

for central charge c. This is the central extension of Witt algebra which is for $c = 0$. The Zamolodchikov c-theorem states that the central charge decreases monotonically under the renormalization group flow for 2D field systems.

Question 77. *How field theory relates to Virasoro algebra?*

The connection with field theory is to treat a 1+1 quantum field system as a representation of Virasoro algebra. The set of eigenstates generated by L_n from a vacuum $|\Omega\rangle$ is called 'conformal tower'. The highest-weight state corresponds to a 'primary' field ϕ, i.e.,

$$\phi|\Omega\rangle = |\phi\rangle, L_n|\phi\rangle = 0, n > 0, \tag{3.134}$$

and others

$$|\phi, \vec{n}\rangle := \prod_{n_j>0} L_{-n_j}|\phi\rangle \tag{3.135}$$

obtained from L_n for $n < 0$ are 'descendants'. The set $\{|\phi, \vec{n}\rangle\}$ spans the Hilbert space of the model. A primary field has a scaling dimension Δ from the dilation (scaling) transform D. If the field has spin s, then the holomorphic conformal dimension h and its anti part are

$$h = (\Delta + s)/2,\ \bar{h} = (\Delta - s)/2. \tag{3.136}$$

The highest-weight is nothing but the conformal dimension h such that

$$L_0|\phi\rangle = h|\phi\rangle. \tag{3.137}$$

Primary fields (or chiral fields) are invariant with respect to lCT

$$L_n\phi(z) = -z^{n+1}\partial_z\phi(z) - (n+1)hz^n\phi(z). \tag{3.138}$$

A primary field $\phi(z, \bar{z})$ ($\bar{z} = z^*$) transforms as

$$\phi(z, \bar{z}) \mapsto \left(\frac{dw}{dz}\right)^{-h} \left(\frac{d\bar{w}}{d\bar{z}}\right)^{-\bar{h}} \phi(z, \bar{z}). \tag{3.139}$$

The scaling transform D can be done by renormalization flow.

How many primary fields there are? It turns out this is a deep question, and physicists decide to make it finite. Such models are called 'minimal' and easy to study. Representations of Virasoro algebra are called Verma modules. Minimal models are defined as systems with finite number of Verma modules, and labelled by central charge and conformal dimension. Well known examples include the critical Ising model with $c = 1/2$, and two nontrivial primary fields, Majorana fermion ψ and

Ising anyon σ, the tricritical Ising (Ising with vacant sites) with $c = 7/10$, and five nontrivial primary fields, and the Z_3 Potts (qutrit) model with $c = 4/5$, and five nontrivial primary fields. The $c = 1/2$ Ising is interpreted as a single Majorana fermion, and there are a family of models, e.g., WZW models, that can be viewed as n Majorana fermions with $c = n/2$. This include the famous XY model with $c = 1$, as a critical region in the XXZ model.

Correlation functions of exponents of primary fields $\langle \prod_n e^{i\alpha_n \phi_n} \rangle$ are the main physical observable of the theory. However, correlation functions are especially difficult to compute, which needs operator product expansions (OPE) and vertex operator algebra (VOA), and also non-factorizable 'conformal blocks'. A CFT is called 'rational' when the z part and $\bar{z}$ part factorizes and only one sector matters. Usually, we only study rational and minimal CFT. General strategy to make a CFT well defined is called 'conformal bootstrap'.

Question 78. *What does fusion rules mean?*

The most primary algebra on primary fields is the fusion rules with fusion denoted by '$\times$'. A fusion between ϕ_i and ϕ_j in general takes the form

$$\phi_i \times \phi_j = \sum_k N_{ij}^k \phi_k. \tag{3.140}$$

Is it an isomorphism between direct sum and product of Hilbert spaces? Do the fields have to brought at the same location? In CFT, fusion means the generation of new fields from product of fields using operator-product expansion (OPE), so the fields do not have to be brought together. The symbol $\times$ shall not be viewed as tensor product $\otimes$ of Hilbert spaces, instead, it means there exist several fields in the state; while the symbol $+$ can be treated as $\oplus$ of Hilbert spaces. So fusion means that the Hilbert space of fields on the left of fusion equation can be written as direct sum of Hilbert space of fields on the right of fusion equation. The isomorphism is by a unitary process, U. We can define the fusion process as U.

For instance, for toric code we can use a segment of Wilson loop, C, to create a pair of charges, e. This process is unitary. To fusion them, $e \times e = 1$ means it will go back to a ground state. The fusion is done by C itself since it is invertible.

For non-abelian anyons, the fusion only means that there are additional dimensions for the Hilbert space, which is the 'quantum dimension' of anyons. For abelian anyons the quantum dimension is one, and this means that they can be created or annihilated deterministically from the vacuum. However, non-abelian anyons cannot be created or annihilated deterministically; in order to make them unitary, additional space is required, which is the quantum dimension.

The braiding of anyons in TQFT is an additional structure different from the fusion rules. In CFT, there is no braiding. The quantum dimension supports the 'fusion space', and braiding of anyons generate unitary operations on it.

We have not explained OPE in details, which is central in CFT. The OPE, also called OP algebra, can be proved rigorously. Here we show it for vertex operators, $V(\phi,z)$ for field ϕ. The OPE takes the form

$$V(\phi,z)V(\psi,w) = V(V(\phi,z-w)\psi,w), \tag{3.141}$$

while the RHS can be expanded as a sum of vertex operators for fields, weighted by power-law factors. This is the underlying physics for the fusion rule. In addition, the OPE is well established in CFT, and whether it can be extended to all QFT is still an on-going research.

We now study CFT with symmetry which brings more structures to a model. A Wess–Zumino–Witten (WZW) model is 2D CFT with a Lie-group symmetry, and the symmetry algebra is an affine Lie algebra, which is also a Kac–Moody algebra. An affine Lie algebra is an infinite-dimensional Lie algebra that is constructed in a canonical fashion out of a finite-dimensional simple Lie algebra.

Question 79. *What is the physical observable for gapless system?*

It is current. The current operators satisfy Kac–Moody algebra. For a Lie group G with generators t^a and $[t^a,t^b] = if^{abc}t^c$, $\mathrm{tr}(t^at^b) = \delta_{ab}/2$, and the coordinate of 2D system $z = \tau + ix$, the current operators $J^a(z)$ satisfy the OPE form

$$J^a(z)J^b(z') = \frac{k}{8\pi^2(z-z')^2} + if^{abc}\frac{J^c(z')}{2\pi(z-z')} + \cdots. \tag{3.142}$$

The constant k is the 'level' of the model, G_k. With Laurent expansion,

$$J^a(z) = \sum_n z^{-n-1}J^a_n, n \in \mathbb{Z}, \tag{3.143}$$

we have the Kac–Moody algebra

$$[J^a_n, J^b_m] = if^{abc}J^c_{n+m} + kn\delta_{ab}\delta_{n+m,0}. \tag{3.144}$$

The zero-th component J^a_0 satisfy the usual Lie algebra. The current operators are chiral and commute $[J^a_n, \bar{J}^b_m] = 0$. The current operators are defined in terms of free fermion operators $L_{\alpha,n}$ and $R_{\alpha,n}$ for n as species by

$$J^a(z) = \sum_n R^\dagger_{\alpha,n}(z)t^a_{\alpha\beta}R_{\beta,n}(z), \tag{3.145}$$

and L for z^*. The current operators serve as lowering and raising operators to generate states of the space.

The number k is related to the central charge c by

$$c = \frac{kD}{k+v_c} \tag{3.146}$$

for D as the dimension of the adjoint irrep of group G and v_c as the eigenvalue of Casimir operator in adjoint rep. For $SU(N)$, $c = \frac{k(N^2-1)}{k+N}$. The generator of Virasoro algebra is

$$L_n \propto \sum_m : J_m^a J_{n-m}^a : \tag{3.147}$$

where : : means normal-ordering.

The current operators can also be expressed with the G-valued field g as

$$J(z) = -\frac{k}{2\pi} g\, \partial g^{-1},\ J(\bar{z}) = \frac{k}{2\pi} g\, \bar{\partial} g^{-1}, \tag{3.148}$$

and the field g is the Wess-Zumino field in the WZW action $S = \Gamma_0(g) + \Gamma(g)$, for 'nonlinear sigma' term

$$\Gamma_0(g) = \frac{-k}{16\pi} \int d^2x\, \mathrm{tr}(\partial_\mu g^{-1} \partial_\mu g) \tag{3.149}$$

and topological term

$$\Gamma(g) = \frac{ik}{24\pi} \int d\xi \int d^2x\, \varepsilon^{\alpha\beta\gamma} \mathrm{tr}(\Pi_\alpha \Pi_\beta \Pi_\gamma) \tag{3.150}$$

for $\Pi_i = g^{-1}\partial_i g$. The integral in $\Gamma(g)$ is one-dimensional higher than the model, taking the model as the boundary, while it is a total derivative so does not depend on the bulk. This relates to the Stokes formula. For $SU(N)$, the matrix field g is unitary in the fundmental irrep. The WZW action can describe fermionic system, as such, it is also known as 'non-abelian bosonization' method. Furthermore, it is also equivalent to abelian bosonization by decomposing g in terms of several bosonic fields, such as in sine-Gordon field theory.

In WZW model, in addition to g there are also other primary fields, such as current operators. For $SU(N)_k$ WZW models, primary fields can be viewed as irreps of $SU(N)$. The $SU(2)_k$ WZW models, which are the boundary of $SU(2)_k$ CS models, for $k > 3$ and $k \neq 4$ are proved to be universal for quantum computing. The $k = 4$ does not work since it reduces to Ising anyon case.

$SU(N)$ WZW models can describe the critical points in $SU(N)$ spin chains, such as the dimer-VBS transition. The $SU(2)_1$ has $c = 1$, two Majorana fermions, describes XXZ model. The $SU(2)_2$ has $c = 3/2$, three Majorana fermions, and it can describe spin-1/2 ladder which has three Majorana fermions as a triplet and another Majorana fermion as a decoupled singlet. For $SU(3)$ VBS model with adjoint irreps, the central charge is $c = 16/5$.

Fusion rules of WZW models are well developed based on representations of affine Lie algebra. There is a remarkable duality between rank and level: e.g., the model of $SU(N)_k$ is dual to $SU(k)_N$. This can be seen from the symmetry of the Yound tableau. We will not discuss this in details.

CFT can serve as the edge of TQFT. We have seen that there is the correspondence between CS theory in the bulk and WZW theory on the boundary, for the same group G. Such a bulk-boundary correspondence or duality is an example of 'holographic principle', the exact content of which is still elusive. It roughly states that the boundary encodes all the information of the bulk, and vice versa. However, people find that this correspondence is not one-to-one.

We discuss the correspondence for Abelian anyon models, e.g., integer quantum Hall states, using $U(1)^N$-CS theory for N species of anyons. Without the level k, the model is

$$\mathscr{L} = \frac{\varepsilon^{\mu\nu\rho}}{4\pi} \sum_{I,J=1}^{N} a^I_\mu K_{IJ} \partial_\nu a^J_\rho. \tag{3.151}$$

The model is defined by a^I and a matrix K, which is symmetric and has real-integer elements. When the microscopic particle is boson, all K_{II} are even integers, while if there is an odd K_{II}, there will be fermions in the system. K encodes the statistics of anyons.

When the model has a boundary, we know that the boundary is a CFT. Here, there are N chiral boson fields ϕ^I for each a^I, which are compactified $\phi^I \equiv \phi^I + 2\pi$, since the physical observable are 'vertex operators' $V(\phi^I) = e^{i\phi^I}$. The edge is one-dimensional so the fields are chiral: left mover z and right mover $\bar{z}$. The vertex operators satisfy VOA, and the fields satisfy Kac-Moody algebra

$$[\phi^I_x, \partial_y \phi^J_y] = \pm 2\pi K^{-1}_{IJ} i\delta(x-y) \tag{3.152}$$

and $\pm$ for chirality. The edge can be described by a sine-Gordon field theory with terms like $\cos\phi^I$, which can induce a gap when relevant.

Chapter 4

Quantum Computation

Question 80. *Is computation a physical process?*

Computation is a process to transform information, which is basically, entropy. Information or entropy comes with energy since it has to be carried by physical objects. Therefore, of course, computation is physical.

Quantum computation has become an important brunch of quantum physics. It appears as applied physics, however, it often deals with problems harder than physical ones, such as efficiency, complexity, and as such, it has greatly improved quantum physics. By treating the evolution of the universe as a computation process, lots of physical rules and problems can be refreshed from the point of view of computation.

4.1 Universal computing models

Question 81. *What can be computed, and how?*

This question is so easy to answer. But, if you take it seriously, it could be hard. The notion 'compute' is similar but more general than 'calculate', and this is why a computer is more powerful than a calculator. About fifty years ago, a computer is nothing but a person who would use a device to calculate something. Nowadays a computer is a device that can process all kinds of *digital signals*. It is digital since all information is encoded by strings of 0s and 1s, which are encoded by values of voltage or geometric shapes on a disc. Problems that can or cannot be computed, with finite (i.e., efficient) amount of time and storage space or not, are the subject of computational complexity theory. When a certain problems turn to be hard, there usually exist some underlying mathematical or physical reasons for that. I believe, the mathematical ones can also be attributed to physical ones, since the computation process is constrained by physical laws.

When information is encoded as strings of bits, computation is merely the change of it, namely, flip of bits. How about qubits or pbits? Well, you can view them as collection of bits: a qubit is a coherent set, while a pbit is an incoherent (i.e., mixing)

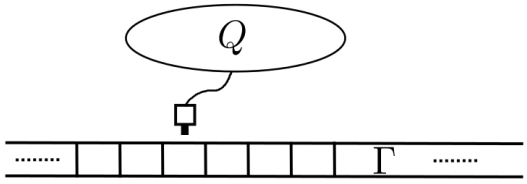

Figure 4.1: The model of Turing machines. There is a *shift function* acting on the head, the head together with the wire shows there is interaction (*computing function*) between one cell of the tape Γ and the processor Q.

set. Quantum or probabilistic computation is just a coherent or incoherent collection of computations with bits, usually called 'classical' computation. We will denote a bit by 'cbit', and use 'bit' as a general name for either of the three: cit, pbit, or qubit.

We already know how these models work. But if you have a chance to reinvent them, how could you do? Given your sense of how nature works, and given a collection of bits, what simple rules to use to manipulate them? A computation can be divided into many steps. It might be hard to change them all at the same time, so only a constant number of bits can be changed each time. How the value of a bit can be changed? There can be different ways:

1. it can be changed depending on the values of other bits.

2. it can be changed each at a time by an external control.

3. several bits can be changed by an external control.

These methods are known in order as: cellular automaton, Turing machine, and circuit model. We will only analyse the last two, as the first one is not well understood yet for the quantum case.

For a Turing machine, the bits do not interact with each other. Instead, each bit interacts with an external control. You may imagine the bits are distributed far apart and it is hard for them to interact directly. A bit can be anything that has a discrete number of states. For instance, the communication of two stations by a satellite is such a situation. For the circuit model, the bits can interact directly with the help of an external control. Usually at most two or three bits can interact directly. This is the model that is being used in our modern computers. These features apply to all the three (c,p,q) models.

4.1.1 Turing machine

A Turing machine, classical, probabilistic, or quantum, has a processor (also known as control), denoted by the symbol Q, state of which is often called *internal state*, and a register (tape) Γ of a string of non-interacting bits, which usually contains the input and output, and a head, which can read, write, move left or right by at most one

step, and a transition function δ for each step, which forms the program to solve a certain problem. Usually the processor Q is specified to have an initial internal state $q_0 \in Q$ and a set of halting states $F \subseteq Q$ so that the machine halts when the internal state reaches a halting state.

Question 82. *Can there be classical elements in a quantum Turing machine?*

A quantum Turing machine does not have to be fully quantum, just like a quantum model does not have to be fully quantum; instead, it could be semi-quantum or semi-classical. The head position shall be classical since it induces the local interactions between the tape and the processor.

Let Γ_n be the space of the nth bit, and it is $\{0,1\}$ for a cbit, $\Omega = \{(p, 1-p) : p \in [0,1]\}$ for a pbit, $\mathscr{P}(\mathbb{C}^2)$ for a qubit. The transition function contains a computing map

$$\delta : Q \times \Gamma_n \to Q \times \Gamma_n \tag{4.1}$$

and a head position shift function

$$\mathbb{Z} \to \mathbb{Z} : p_\ell \mapsto p_{\ell+1} = p_\ell \pm 1, 0. \tag{4.2}$$

Note here in each step the head has a definite location! There will not be superposition or probability distribution of the head position. This is a reasonable setting since in each step we need to make sure only one bit is acted upon. Otherwise, the uncertainty of the head position will induce nonlocal interaction between the tape and the control. A state of the whole machine is often known as a *configuration*, including the state of tape, head position, processor (and some others). A computation on TM can be viewed as a sequence of configurations, and a conversion between any successive two configurations can be described by a permutation, stochastic, or unitary matrix, which are for CTM, PTM, or QTM, respectively.

Our definition is different from a global description of the transition function

$$\delta : Q \backslash F \times \Gamma \times Q \times \Gamma \times \{L, R, N\} \to \mathscr{D}, \tag{4.3}$$

for

$$\mathscr{D}_{\text{CTM}} = \{0,1\},\ \mathscr{D}_{\text{PTM}} = [0,1],\ \mathscr{D}_{\text{QTM}} = \mathbb{C}. \tag{4.4}$$

Here L (left), R (right), and N (no movement) specifies the motion of the head. This form does not reveal the locality explicitly.

Although simply defined, the TM has nontrivial computational and physical features. Besides the locality, the other feature is that the head needs to move in both directions to achieve universality. Namely, each bit can be acted upon many times. A one-way head is called unilateral, and a unilateral CTM is called a transducer, which is known to be non-universal.

For a CTM, the tape is formed by a string of bits, and each step can be described by a permutation, Π. A PTM is usually understood as a randomized CTM, and the

randomness can be realized by random variables, which can be encoded by a string of pbits on a so-called random tape, and the computation by a PTM is a randomized permutation, which can be described by a stochastic matrix. Also each step of a PTM is a stochastic matrix

$$S = \sum_{\lambda} \mathfrak{p}_{\lambda} \Pi_{\lambda} \tag{4.5}$$

for a certain permutations Π_λ with probability $\mathfrak{p}_\lambda$, which is represented on the random tape. The product of a sequence of stochastic matrices can be expressed as

$$\prod_i S_i = \sum_{\lambda_1,\lambda_2,\ldots} \mathfrak{p}_{\lambda_1}\mathfrak{p}_{\lambda_2}\cdots\left(\prod_i \Pi_{\lambda_i}\right), \tag{4.6}$$

and each sequence in the parenthese above represents a CTM with corresponding probability. That is, a CTM realizes a particular trajectory of a PTM. The output of a PTM is the final state $\gamma \in \Gamma$ on the tape with probability

$$\mathscr{P}(\gamma) = \sum_{q\in Q} \mathscr{P}(q,\gamma), \tag{4.7}$$

where the sum is over internal state q for the same γ. Note there is no sum over $p \in P$ as the head position is fixed.

Replacing permutation by unitary matrix, a CTM generalizes to a QTM. The quantum control (processor) Q contains a set of qubits, which could interact with each other or not, or a higher-dimensional system. It is also beneficial, although not necessary, to have a classical control *CC*, which is formed by a set of classical states $\{c\}$ that corresponds to the computation part, and has a starting state c_0 for the quantum starting state, and some halting states $\{c_f\}$ for the quantum halting states of the processor. The function of *CC* is to signal the process of the machine such that the machine halts when the classical control is at a halting state, without actually measuring the quantum state of the machine. Furthermore, with entanglement and teleportation, which is unique for quantum system, a two-way head QTM can be simulated by a unilateral QTM (see later).

At last, how about replacing a stochastic matrix by something quantum? This refers to the so-called quantum stochastic process, which is nothing but a CPTP map. As we know, a CPTP map can be dilated to a unitary evolution, also unique for quantum system. Tracing out the ancilla part will result in a CPTP map. As a result, the quantum stochastic TM is equivalent to QTM described above. However, it is not known whether PTM is equivalent to CTM since there is no classical version of dilation.

4.1.2 Circuit model

For Turing machines, the control somehow seems mysterious: its structure is not so clear. However, the physical requirement is less restrictive than that of the tape.

Information is encoded on the tape and the control merely assists the information processing on the tape.

In circuit model, the external control plays a trivial role: it does not participate the computation. A circuit contains some number of *gates* acting on a certain number of bits, with final measurements. A circuit is a function that maps the input initial state to the output final state. A circuit shall be *digital*: its gates shall be only of forms from a finite set. Such a set is said to be *universal* such that any gates or functions can be constructed to an arbitrary accuracy. Classical circuits, known as Boolean circuits, contains Boolean logic gates. A common basis is the set {AND, OR, NOT}, which is non-invertible. A universal invertible gate is the Toffoli gate, i.e. the controlled-controlled-not gate

$$CCX = |11\rangle\langle 11| \otimes X + \text{trivial}, \; X \equiv NOT. \tag{4.8}$$

Classical universality means that arbitrary permutation can be realized.

Question 83. *What does universality mean for quantum computing?*

Quantum circuits, on the other hand, need a bit more: it needs the Hadamard gate

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \tag{4.9}$$

which looks simple but is really powerful since the set $\{CCX, H\}$ is universal for quantum computing. The Hadamard gate can generate superposition hence interference, and also entanglement between qubits. In addition, another universal set of gates is $\{CZ, T, H\}$. The so-called T gate causes us a lot of trouble, as we will see later on in much details. Quantum universality means that arbitrary unitary operation can be realized to arbitrary accuracy. It cannot be realized exactly since a discrete set can only generate a finite set given a finite product of them. It is this definition of universality that represents the digital feature of quantum computing.

How about the stochastic case? Due to the convex decomposition of any stochastic matrix to a sum of permutations, a stochastic circuit can be reduced to a set of random Boolean circuits. Well, instead, you might be interested to ask: what is a universal set of stochastic gates? This problem is hard to deal with since there is no group structure for this set, compared with the permutation or unitary cases. Also it appears not necessary to find such a set since stochastic process is no easier to control than a permutation or unitary evolution.

4.1.3 Turing machine vs. circuit model

Question 84. *What does it mean by 'A simulates B'?*

Simulation can be used to prove the equivalence between two models. As the 'equivalence' can be of different senses, there are many kinds of simulation methods.

As the word 'scientist' means differently for different scientists, the word 'simulation' also means differently for different simulations. We will use simulation to prove the equivalence between Turing machine model and the circuit model. Then we will analyze other kinds of simulations.

The simulation of a TM M by another TM U is a task such that

$$\mathrm{U}([\mathrm{M}],[x]) = [\mathrm{M}(x)],\ \forall x, \tag{4.10}$$

here $[\cdot]$ represents encoding, e.g., $[\mathrm{M}]$ is the bit-string description of M. The simulation is efficient if there is only a polynomial overhead of cost for all input x. Furthermore, as $[\mathrm{M}]$ is only being read during the simulation, $[\mathrm{M}]$ does not have to be the input of U, hence in fact

$$\mathrm{U}([x]) = [\mathrm{M}(x)],\ \forall x, \tag{4.11}$$

and there exists a program

$$\mathrm{P}([\mathrm{M}],[x]) = [\mathrm{U}],\ \forall x, \tag{4.12}$$

such that P specifies the process of U to simulate M on arbitrary x. Each x is an input of P since the simulation is to simulate the action of M on x, and both P and U are generically x-independent. This is a *strong* simulation, which employs an input-independent simulator U to simulate the effect of the simulatee (simulated object) M on arbitrary input x without building the simulatee itself, and a classical algorithm P is employed to design the simulator given the simulatee.

Now we show the equivalence between the circuit model and the Turing machine. Given a circuit, the input bits can be represented on the tape of TM, and each gate in it can be simulated by a finite number of steps of a TM. Given a Turing machine with a fixed size of tape, the size of the tape and the state of control decide the number of input bits in a circuit. Each step in TM can be realized by a finite number of gates, hence in total a TM can be simulated efficiently.

There also exists the so-called universal TM that can simulate other TMs. This merely says that within the framework of TM, it is universal to compute any computable functions. It is equivalent to say within the framework of circuit model that there is a universal set of gates. Given a TM, its configurations can be encoded on the tape of the universal TM. By reading them its steps can be simulated efficiently.

This carries over to the quantum case. Given a QTM, its tape and control states can be encoded as qubits for a circuit. Each step in the QTM is unitary and can be simulated by a finite number of gates. Given a circuit, each gate in it can be simulated. The gate can only be from $\{CZ, H, T\}$ as Toffoli gate can be reduced to them. The H and T gates can be directly applied to the qubits. The CZ, or equivalently CX, gate can be realized as

$$CZ_{12}|\psi\rangle_{12}|0\rangle_a = S_{1a}CZ_{2a}S_{1a}|\psi\rangle_{12}|0\rangle_a, \tag{4.13}$$

which is easy to prove for S as swap gate. It carries the information of qubit 1 to the ancilla first, then do the CZ gate and then carries the information back. The

gates are not sequential in the sense that the ancilla interacts with qubit 1 twice. The wisdom that is hard to see is that any entangling gate cannot be realized sequentially by interacting with an ancilla. (show it!) This is the key reason that a QTM has to have a two-way head, otherwise it cannot simulate entangling gates. With the method, a circuit can be simulated efficiently on a QTM, which is, actually a universal QTM. The gate in a QTM may not be directly from a universal gate set, but it can be simulated by product of them. A QTM such that each step is a gate from a universal gate set is a universal QTM.

Furthermore, QTM can be simplified a bit to a unilateral one with the presence of entanglement and teleportation, which needs projective measurement and classical feedback. The non-unilateral feature is due to the gate CZ (or CX). To convert the procedure (4.13) to a 'sequential' one, we can teleport qubit 1 to another qubit $1'$ before the 2nd gate S_{1a}. The qubit $1'$ and $2'$ is in the maximally entangled state $|\omega\rangle = \frac{1}{\sqrt{2}}\sum_i |ii\rangle$. Bell measurement on 1 and $2'$ will teleport qubit 1 to qubit $1'$. This is not sequential yet. We need further to express the state $|\omega\rangle$ as a matrix-product state such that the correlator decouples. Let the two qubits be α and β, and a qubit ancilla be a, we find

$$|\omega\rangle = \langle 0|_a BA|0\rangle_a,\ A = |0\rangle_\alpha A^0 + |1\rangle_\alpha A^1,\ B = |0\rangle_\beta B^0 + |1\rangle_\beta B^1, \tag{4.14}$$

with the tensors defined as

$$A^0 = \mathbb{1}/\sqrt{2},\ A^1 = \sigma^x/\sqrt{2},\ B^0 = P_0,\ B^1 = \sigma^+, \tag{4.15}$$

for $P_0 = (\mathbb{1} + \sigma^z)/2$, $\sigma^+ = (\sigma^x + i\sigma^y)/2$, and Pauli matrices σ^x, σ^y, σ^z, and $|0\rangle = (1,0)^t$, $|1\rangle = (0,1)^t$. The pair of matrices A^0 and A^1, B^0 and B^1 each form a quantum channel. The quantum circuit to prepare $|\omega\rangle$ is also easy to find

$$|\omega\rangle = \langle 0|_a U_{\beta a} U_{\alpha a}|000\rangle_{\beta\alpha a}, \tag{4.16}$$

for $U_{\beta a} = S_{\beta a}$ as a swap gate $S_{\beta a}$ realizing B^0 and B^1, $U_{\alpha a} = CX_{\alpha a}H_\alpha$ with the controlled-not (α as control) and Hadamard gate realizing A^0 and A^1. The qubit ancilla a automatically decouples simply because it is swapped with the qubit β.

As a result, the whole process now is sequential with the additional Bell measurement. This means that, a QTM can be converted to a unilateral one with additional Bell measurements at the end of computation, and an enlargement of the control system by product of the bond dimensions of the Bell state $|\omega\rangle$.

4.1.4 Simulation

Simulation deals with closeness between objects. There is no unique way to define closeness and how they can get close. So we shall define simulation carefully and classify them. Here we analyze classification of simulations for quantum operators and probability vectors.

Question 85. *How to classify simulations?*

For quantum operators, the problem of quantum simulation can be characterized with notions from topology of bounded linear operators on Hilbert space. There exist different kinds of convergence in different topologies of the set of bounded linear operators on Hilbert space. Most commonly, there are uniform (or called norm), strong operator, and weak operator topologies.

Correspondingly, there are three kinds of quantum simulations:

- Weak quantum simulation: approximate $T \in \mathscr{L}(\mathscr{H})$ by $\tilde{T}$ within distance $\varepsilon > 0$ such that $|\langle\psi|\tilde{T} - T|\psi\rangle| \leq \varepsilon, \forall\, |\psi\rangle \in \mathscr{H}$.
- Strong quantum simulation: approximate the action of $T \in \mathscr{L}(\mathscr{H})$ on state $|\psi\rangle \in \mathscr{H}$ by $\tilde{T}$ within vector 2-norm distance $\varepsilon > 0$ for the worst case such that $\|T - \tilde{T}\| := \sup_{|\psi\rangle} \|(T - \tilde{T})|\psi\rangle\| \leq \varepsilon$.
- Uniform quantum simulation: approximate $T \in \mathscr{L}(\mathscr{H})$ by $\tilde{T}$ within distance $\varepsilon > 0$ quantified by a certain operator norm. An example is to use channel-state duality and prepare the dual state of T.

The problem of strong quantum simulation of one unitary operator U, e.g. $U = e^{-iHt}$ if it is generated by a time-independent Hamiltonian H, is to approximate it by another unitary $\tilde{U}$ satisfying the spectral norm distance condition $\|U - \tilde{U}\| \leq \varepsilon$. The approximation can be achieved by, e.g., either constructing an approximate Hamiltonian $\tilde{H}$ using easy-to-implement interactions or a direct approximation $\tilde{U}$ using elementary quantum gates.

The problem of quantum state generation is to generate a state $|\psi\rangle$ within distance ε so that $\||\psi\rangle - |\tilde{\psi}\rangle\| \leq \varepsilon$. Now suppose $|\psi\rangle = U|0\rangle$, and $|\tilde{\psi}\rangle = \tilde{U}|0\rangle$ for some unitary operators U and $\tilde{U}$, and then the accuracy condition becomes $\|U|0\rangle - \tilde{U}|0\rangle\| \leq \varepsilon$, which can be ensured if we can simulate U by $\tilde{U}$ strongly; i.e. $\|U - \tilde{U}\| \leq \varepsilon$.

Different quantum simulations have natural physical interpretations. The scenario for uniform quantum simulation is that, given an unknown process, one would like to simulate or approximate the process itself after knowing enough information of the process. One closely-related, yet not the same, task is the quantum process tomography, for which one needs to construct the process matrix of the process itself. For strong quantum simulation, one has to make sure that the output state from a simulator should be close enough to the ideal output state for any input state. This only requires the simulator has the similar effects on all input states. The requirement of weak quantum simulation is merely to ensure that the simulation provides similar observable effects for a given quantum state and observable, without referring to quantum process tomography or state tomography.

For probability vectors in probability space, there are also various ways for simulation depending on convergence of variables. For instance, the convergence in distribution is

$$\lim_{n\to\infty} f_n(x) = f(x), \forall x \in \mathbb{R} \tag{4.17}$$

which is a weak convergence, for f, f_n as distribution function $f(x) = p(X \leq x)$. The stronger one, convergence in probability is

$$\lim_{n\to\infty} p(|X_n - X| \geq \varepsilon) = 0, \forall \varepsilon > 0 \tag{4.18}$$

and even stronger one is the pointwise (or sure) convergence

$$\lim_{n\to\infty} X_n(\omega) = X(\omega), \forall \omega \in \Omega \tag{4.19}$$

for the sample space Ω. There are also many other kinds to play with.

Question 86. *How to simulate quantum objects with classical means?*

Can an operator be simulated classically? This is an important subject, as it will relate to the separation between computation powers of quantum computers and classical computers.

When a given quantum object can be represented efficiently with classical means, then it leads to an efficient simulation. But here we need to be careful of the 'representation'. For instance, we could represent a state ρ_0 as a bit 0, and a state ρ_1 as a bit 1, but then there is no way to represent the nonzero overlap between the states. A full representation is also not clever: a n-qubit state needs 2^n number of parameters. Hence some simulation methods in between need to be employed.

The proper scheme is to simulate the observable effects of quantum processes. When the observable is represented as probability vectors, there are two popular ones:

- strong classical simulation: compute the probability;
- weak classical simulation: sample from the probability.

Note here 'strong' and 'weak' are only heuristic terms. It is known that there is a separation between the two methods; namely, a quantum task may be efficiently simulated by a sampling, but cannot by a strong simulation. Sampling is easier since you only need to generate a sample efficiently according to the probability given as an oracle. There is even a separation between quantum sampling and classical sampling: the Boson Sampling problem can be efficiently solved by quantum computers but not classical computers.

A class of quantum process that can be efficiently weakly simulated classically is the Clifford circuits (via stabilizer formalism). Clifford circuits are not powerful since it is not even universal for classical computing. A Clifford circuit can be understood as a game among Pauli matrices: X, Y, and Z, which can be represented by just two bits. The state of a qubit is labelled by X and Z (Y as their product), so by two bits, and n qubits will need $2n$ bits. Clifford circuit maps stabilizer states to each other. A stabilizer state is defined in terms of a set of stabilizers, which is product

of Pauli matrices. So by keeping track of the stabilizers, the simulation is efficient. This is the content of the Gottesman-Knill theorem. Note that the stabilizers, as operators, are not simulated. For strong simulation, it has to construct the stabilizers themselves.

In phase space, a simulation based on Wigner function can be weak, i.e., it simulates observable effects. The method is to sample values of operator according to Wigner function and then make average, and if Wigner function is negative, then it can use the absolute value of it for the sampling and use the negative value in the average. The simulation cost is the size of the sample. For strong simulation based on Wigner function, it requires the simulation of its dynamics: Moyal equation.

Besides the above, there is also a distinction between analog and digital simulations. Analog simulation originates from analog computers which do not represent everything as bits. As such, there are inherent analog errors that cannot be well controlled. However, analog simulation proves to be useful for a certain tasks. You only want to simulate a limited task, such as: is there a phase transition?, and you do not have to know what the transition point is. This is why analog quantum simulators are pursued by physicists. The analog simulators might be non-precise but accurate, non-universal but reliable, and verifiable, too.

4.2 Quantum gate operations

Question 87. *How to perform a quantum gate, which is a unitary operator?*

To realize a gate U, we can try whatever methods we have, such as non-unitary operators, operators with larger dimensions, measurements, etc, as long as the net effect is a unitary gate U.

For a spin, a unitary operation on it is a rotation. This can be done by an external field. But you may wonder: has the state of the field been changed after the rotation? The field is 'big' so any change shall be tiny. The fact is that the field is an external control so that it would not entangle with the spin. If there are entanglement, the evolution of the spin cannot be unitary. The field is 'semi-classical' in the sense that it could interact with a quantum particle while still has classical states. When the field is quantized, we have to study how photons interact with the spin.

Some computing models are defined according to the physical realization of gate operations. Examples are teleportation and braiding.

4.2.1 Teleportation

Suppose the qubits we have are not very good: their coherence time is short. This means you cannot apply many gates on a qubit. How can we use such noisy qubits for computation? One idea is to use teleportation: it can transfer the state of a qubit completely from one carrier to another carrier. That is, before the decoherence of the

carrier affects the qubit state, we transfer it to another carrier which is just refreshed. However, there is a caveat: the teleportation requires an entangling gate, which shall be performed in a short time.

Question 88. *So what is teleportation?*

It turns out teleportation is a sequence of two swap operations, which is also called 'one-bit teleportation'. It seems there is no magic! However, the magic is the 'gauge' symmetry of teleportation so that we could teleport gates. By representing the one-bit teleportation as a tensor, as in matrix-product states, then the gauge symmetry means that a $U(1)$ rotation on the physical leg is equivalent to a $U(1)$ rotation on one (not two) virtual leg.

An equivalent way to describe this is via cluster state. First each qubit is initially at state $|+\rangle$, and the two-qubit controlled-phase gate $CZ = [\mathbb{1}, 0; 0, Z]$ acts on each neighboring pairs. This prepares a cluster state. To simulate an arbitray qubit gate, we need five sites for a linear information flow. The following relation for information propagation (identity gate) holds

$$HP_s|\psi\rangle = \langle s|CZ|\psi\rangle|+\rangle \tag{4.20}$$

for input known state $|\psi\rangle$, projector $P_s = |s\rangle\langle s|$ in Z basis, $s = 0, 1$. Also projection in X basis leads to

$$HZ^s|\psi\rangle = \langle s|(H\otimes\mathbb{1})CZ|\psi\rangle|+\rangle. \tag{4.21}$$

Two successive projection leads to

$$\begin{aligned}&\langle s|\langle t|(H\otimes H\otimes\mathbb{1})CZ\cdot CZ|\psi\rangle|+\rangle|+\rangle\\ &= HZ^tHZ^s|\psi\rangle = X^tZ^s|\psi\rangle\end{aligned} \tag{4.22}$$

You may find that HZ^s are the tensors in MPS form of cluster state. To execute a general qubit gate, we need measurements in rotated bases. We have

$$|\text{out}\rangle = HZ(\alpha_4)Z^{s_4}HZ(\alpha_3)Z^{s_3}HZ(\alpha_2)Z^{s_2}HZ^{s_1}|\text{in}\rangle \tag{4.23}$$

for $Z(\alpha) := e^{-i\alpha Z} = [e^{-i\alpha}, 0; 0, e^{i\alpha}]$ etc.

We could use only two qubits if we allow a circular information flow: that is, from qubit a to b, then back to a, then to b and so on. Let's make the connection with the coherence time of a qubit. Suppose the time to apply gate CZ is τ, projective measurement is τ_1, and refresh a qubit is τ_2, then for the procedure: 1) prepare state $|\psi\rangle|+\rangle$, 2) apply CZ; 3) measure qubit a; 4) refresh qubit a; 5) apply CZ; 6) measure qubit b; the total time is $2\tau + 2\tau_1 + \tau_2$, which shall be shorter than the coherence time of qubit a and b.

To teleport the entangling gate CZ or CX, we can use four qubits a, b, c, d, and a, b form a pair for an actual qubit, c, d form another qubit. We first prepare a four-site cluster state but with a, d at arbitray initial state $|\alpha\rangle$ and $|\beta\rangle$, respectively. Then measure them in X-basis with outcomes s and t induces the output $|\alpha\rangle|\alpha\oplus\beta\rangle$ with byproduct $(Z^tZ^s)\otimes(X^tZ^s)$, which is the CX gate.

4.2.2 Anyon braiding

Anyon braiding realizes gates in topological ways, and also realizes them in a subspace of the total so-called fusion space. Encoding qubits via anyons is a promising scheme since anyons have topological protection. We will focus on non-abelian anyon braiding since abelian cases are not universal. We term non-abelian anyon as 'nanyon' for simplicity.

Nanyons can fuse to different states belonging to the fusion space, which will be the logical space of qubits. Nanyons cannot exist in vacuum, instead, they are excitations (or defects) from topological system. The states for nanyons are ideally degenerate, and braiding will lead to topological holonomy, which form unitary gates. Crucial features of the fusion space $\mathscr{F}$ are that it does not have a natural tensor product structure

$$\mathscr{F} \neq \otimes_i \mathscr{H}_i, \tag{4.24}$$

and the braiding gates usually do not look familiar, e.g., the gates in a universal gate set.

Question 89. *How powerful is Ising anyon?*

Here we discuss Ising anyon, which is not universal but the gates are easy to perform. A famous example of Ising anyon is the Majorana zero mode, supported in vortex core of $p+ip$ superconductor or fractional quantum hall liquid. Ising anyon σ is defined so it can 'eat' a fermion ψ

$$\sigma \times \sigma = 1 + \psi,\ \sigma \times \psi = \sigma,\ \psi \times \psi = 1. \tag{4.25}$$

Due to the first relation, a qubit can be encoded by two anyons. The two states are distinct for their fermion charge $i\sigma_1\sigma_2 = \pm 1$. If we encode a qubit by two anyons, then $Z_L = i\sigma_1\sigma_2$, $X_L = \sigma_1$. Braiding of i and j leads to gate

$$B_{ij} = e^{-\frac{\pi}{4}\sigma_i\sigma_j} \tag{4.26}$$

So the braiding B_{12} is $e^{i\frac{\pi}{4}Z_L} = S$ for phase gate S, square root of Z_L. However, it is difficult to generate superposition of 1 and ψ, i.e., difficult to do the Hadamard gate, usually more anyons are used. It is easy to see four anyons is a good choice, and we can use the subspace stabilized by $-\sigma_1\sigma_2\sigma_3\sigma_4$, which is the space for even number of fermions: zero or two. The minus sign is important. $2n$ anyons spans dimension 2^n, while we can use $4n$ anyons to encode n qubits. The two logical gates are

$$Z_L = i\sigma_1\sigma_2 = i\sigma_3\sigma_4,\ X_L = i\sigma_1\sigma_4 = i\sigma_2\sigma_3. \tag{4.27}$$

So the braiding B_{12} is the phase gate S, the braiding B_{23} is $e^{i\frac{\pi}{4}X_L} = SHS$ for phase gate S and Hadamard H. However, there is no entangling gate from braiding via this encoding, the so-called 'sparse' encoding.

Entangling gate is possible when n qubits encoded by $2n+2$ anyons in the 'dense' encoding. Say, for 2 qubits we need 6 anyons. But there is no Hadamard gate and phase gate for this encoding. So we can switch between the dense and sparse encoding by measurement in order to obtain all the gates.

The T gate is complicated: it can be done with Dehn twist since the conformal spin of Ising anyon is $1/16$. (show it!) We would not describe the scheme here.

4.2.3 Other methods

Physicists have invented many ways to perform gates. Let's look at some more of them below

- Holonomic (Geometric); Adiabatic; Dissipative; Quantum walk

We already know non-abelian geometric phases are also called holonomy. Anyon braiding relies on adiabatic topological holonomy. We could simply use holonomy without the adiabatic and topological features, and also without the braiding operations. This is the geometric or holonomic QC, which can be realized by qubits encoded in small systems. The gate operations in holonomic QC is induced by U that acts on Hamiltonian H, instead of state. The external drive U will introduce time-dependence to the model H, and we know this in general can generate holonomy. Different external drives which form different cycles are used to simulate different gates. However, the external drives are no easier than the Hamiltonian evolution itself.

In the Hamiltonian setting, the continuous-time quantum walk is also shown to be universal by simulation of gates. There is no exponential overhead for interacting quantum walkers. The Hamiltonian is time-independent, so it does not rely on geometric phase and its evolution is not adiabatic.

In addition, there are also models for which the universality is only proved via state generation. No universal gate simulation has been proved. This includes the adiabatic QC and dissipative QC. The model is to use a Hamiltonian or a quantum channel, both can have a notion of 'gap', such that it can prepare the desired state as a ground or fixed state. However, it seems the desired state is only prepared with a certain probability. Whether there is a gate-simulation proof for the universality of these models is an open problem.

4.3 Universal fault-tolerant QC

4.3.1 Quantum codes

Question 90. *What is a quantum code?*

A quantum code is simple: it is just a Hilbert space. In order for it to be nontrivial, it has to be a subspace of a bigger one. For instance, the singlet from two spin-half is a quantum code, which only encodes one state. To make it more useful, a code has to be able to detect or correct errors, which is the main motivation for using a big space to encode a smaller one.

A quantum code is defined by an encoding isometry $V : \mathscr{H}_l \to \mathscr{H}$ and $\dim\mathscr{H}_l = \dim\mathscr{C}$, for projector $\mathscr{P} = VV^\dagger$ on the code space $\mathscr{C} \subset \mathscr{H}$ of the physical system, denoted by Q. The space $\mathscr{H}$ could have a nontrivial structure. For instance, there could be a tensor-product structure $\mathscr{H} = \otimes_{n=1}^{N}\mathscr{H}_n$ for a system with N subsystems, each denoted as Q_n. A notion of locality could be defined for the set of Q_n, which might be relatively 'easy' to access in practice.

Quantum error correction refers to the correction of a set of error operators $\{E_i\}$ on $\mathscr{H}$ so that errors cannot induce or disturb any nontrivial logical operations on $\mathscr{C}$. Given $\{E_i\}$, the necessary and sufficient condition for its correctability is the well-known Knill-Laflamme condition

$$\mathscr{P}E_i^\dagger E_j\mathscr{P} = a_{ij}\mathscr{P} \propto \mathscr{P},\ a_{ij} \in \mathbb{C}. \tag{4.28}$$

Meanwhile, the condition for detection, which is weaker than correction, takes the form

$$\mathscr{P}E_i\mathscr{P} = e_i\mathscr{P} \propto \mathscr{P},\ e_i \in \mathbb{C}. \tag{4.29}$$

Each error operator E_i could be local or nonlocal.

Here the linearity of quantum theory jumps in: when $\{E_i\}$ is a local basis of $\mathscr{B}(\mathscr{H}_n)$ (the subscript n is omitted for error operators E_i) for any n, then for any $F_j = \sum_i f_{ij}E_i \in \mathscr{B}(\mathscr{H}_n)$, condition (4.29) implies

$$\mathscr{P}F_j\mathscr{P} \propto \mathscr{P}, \tag{4.30}$$

which means arbitrary local error can be detected, and condition (4.28) implies

$$\mathscr{P}F_j^\dagger F_l\mathscr{P} \propto \mathscr{P}, \tag{4.31}$$

which means arbitrary local error can be corrected. For instance, if the subsystem is a two-level system, i.e., a qubit, the local error operators to consider are the Pauli operators $\sigma^x, \sigma^y, \sigma^z$ and identity.

It does not care about what are the physical noises in practice. Instead, it is mathematical or *digital*: it reduces all noises into linear combination of the error generators, such as Pauli operators. This is based on the linearity of quantum theory. However, it indeed has an assumption: it requires that there are measurement instruments that can perform measurement that identify the error generators, usually this is projective measurement.

Furthermore, there are codes that can correct multiple local errors or nonlocal errors. We say t errors can be corrected when arbitrary errors on t subsystems can be corrected. A 'code distance'

$$d = 2t + 1 \tag{4.32}$$

can be defined if t errors can be corrected, which also means that $2t$ errors can be detected.

The correction procedure, also known as decoding, usually involves measurement. It can be described via the framework of quantum channel, or quantum instrument, which is mixture of CP mappings that together form a TP mapping. Given a_{ij}, the matrix $[a_{ij}]$ can be rotated to be diagonal by a unitary operation $U = [u_{ij}]$, and the error set becomes $\{F_j\}$ for $F_j = \sum_i u_{ij} E_i$. The coefficients a_{ij} in (4.28) will be replaced by $\delta_{ij} d_j$. Then the recovery channel can be specified by a set of Kraus operators $\{R_j\}$ for $R_j \propto \mathscr{P} F_j^\dagger$. The operator $F_j^\dagger$ might be nonlocal. However, for quantum codes with more structures, e.g., with a set of commuting or frustration-free local terms, these terms can be measured locally which will benefit the decoding.

4.3.2 Universal vs. Fault-tolerant gate set

QC with quantum codes is fault-tolerant since errors can be corrected. Now, how to perform quantum gates between quantum codes? You could imagine various kinds of logical gates. Here one important and simple class is the *transversal* logical gate which in general takes the form

$$U = \bigotimes_j U_j. \tag{4.33}$$

It is a tensor product of non-overlapping U_j. Such gates would not spread errors among the blocks labelled by j, which can be a local subsystem. Note that U is defined up to any permutation of code blocks since permutation cannot spread out errors except the locations of them. Independent error correction on blocks i will delete errors.

An error-correction code block, or code block for short, is a part of the whole system for which error correction can be performed. As such, a system Q is partitioned as a union of non-overlapping code blocks $Q = \bigcup_i Q_i$. The natural choice of a code block is a subsystem. However, a code block can consist of several subsystems, which should usually be a connected local part of the whole system. This applies to topological codes that have macroscopic code distance. Transversal logical gates do not spread out errors across code blocks for the same logical qubit. So the error-correction for each logical qubit can ensure fault tolerance.

For many logical qubits each encoded by a different physical system, denoted as $Q[n]$, usually a one-to-one correspondence of code blocks has to be chosen. For instance, for two logical qubits $Q[n] = \bigcup_i Q_i[n]$, $n = 1,2$, a code block $Q_i[1]$ can be chosen to correspond to $Q_i[2]$ for the same label i. In general, there might be a permutation $\pi(i)$. If each system comes with a Hamiltonian $H[n]$, then the total Hamiltonian is the sum of them $H = \sum_n H[n]$ without interaction terms.

Furthermore, when a physical system is used to encode many qubits, the notion of code block and transversality need to be defined carefully. For instance, if a harmonic

oscillator is used to encode a qudit, which is equivalent to a certain number of qubits, the notion of locality is different from that in many-body systems. If the system has a Hamiltonian H, it cannot be written as the sum of Hamiltonians for each qubit. This means that the locality is defined with respect to the whole system, and code blocks are shared by all qubits. For both single-qubit gates and entangling gates, entangling operations on code blocks might be involved. Nevertheless, logical gates should take the form (4.33) for a fixed partition of code blocks labelled by j.

Question 91. *How powerful are transversal gates?*

It turns out there is a crucial limitation of transversal gates: there is no universal transversal logical gate set on an exact code, which is the Eastin-Knill theorem. A unitary operator U is a logical operator iff

$$UP = PUP. \tag{4.34}$$

A unitary operator $U^\dagger$ is a logical operator iff

$$PU = PUP. \tag{4.35}$$

It turns out, if U is logical, then $U^\dagger$ is also logical. This means a state in the codespace cannot be mapped out of it by U, while a state out of the codespace cannot be mapped into it by U, neither.

If the codespace is defined by a Hamiltonian H, then it has $[P,H] = 0$. If the codewords are further degenerate, then $HP = hP$ for a constant h which can be set to zero. A logical operator U can be viewed as an 'emergent' symmetry of H as

$$P[U,H]P = 0, \tag{4.36}$$

which means U preserves H on the codespace P. While in general U is not a symmetry of H. As a result, we see that emergent symmetry plays more important roles than symmetry in the presence of a Hamiltonian for logical operations.

In order to see the generality of Eastin-Knill theorem, below we review its content in details. There are several crucial assumptions. (a) The error set $\{E_i\}$ spans the space of a code block, namely, it requires arbitrary errors on a code block can be detected. If the dimension of the space spanned by $\{E_i\}$ is smaller that that of a code block, the theorem does not apply. (b) Transversality is fixed, namely, all logical transversal gate takes the form (4.33) for a given partition of code blocks. (c) The Hilbert space dimension of the system is finite.

Given a code space $\mathscr{C}$, it first shows that the set of logical gates (4.34) form a Lie group $\mathscr{L}$. Given a transversality and the form (4.33), it shows that the set of transversal logical gates is also a Lie group $\mathscr{G} = \mathscr{L} \bigcap \mathscr{A}$, for $\mathscr{A} = \bigotimes_j U(d_j)$, d_j as the dimension of a code block. Now the connected component of identity $\mathscr{C}$ in $\mathscr{G}$ contains elements of the form $C = \prod_k e^{i\xi_k D_k}$, $\xi_k \in \mathbb{R}$. An operator $e^{i\xi D}$ is a transversal

logical gate for arbitrarily small ξ, and then it shows $DP = PDP$. The operator D can be written as a sum of local terms $D = \sum_j \alpha_j H_j$ due to the structure of Lie algebra, and each H_j acts on the code block j. Given the detection of arbitrary local errors, it holds $PH_jP \propto P$, and then $DP = PDP \propto P$. As the result, $CP \propto P$, which means that C acts as the logical identity gate, and the whole group $\mathscr{C}$ 'collapses' to identity. As the quotient group $\mathscr{Q} = \mathscr{G}/\mathscr{C}$ is a topologically discrete group, the number of logically distinct operators is finite. In other words, the set of transversal logical gates is not universal.

Next we remark on some points. (1) During the execution of each logical gate, there may be leakage out of the codespace, as long as it goes back to the codespace at the end. (2) Each local unitary U_j can be realized in many ways, even not unitarily, as long as the net effect is unitary. For instance, ancilla and measurement can be used. (3) There is no logical ancilla to realize a logical gate U since U itself must be unitary of the form (4.33). (4) A code block can contains several subsystems, and this especially applies to codes with large code distance, such as topological codes. Error correction on subsystems ensures error correction on a code block. For code distance $d = 2t+1$, a code block can be as big as t. This means that each U_j can be an entangling gate on the underlying subsystems. However, the transversality has to be fixed to ensure that all logical gates are of the form (4.33). (5) it does not apply to gates via finite-depth circuits since they do not form a group given a finite upper bound of depth, and the algebra will not factorize into sum of local ones.

Question 92. *What shall replace transversality in order to achieve universality?*

The no-go theorem can be circumvented by relaxing its assumptions. Measurement-based schemes, such as magic-state injection, violate condition (b) since logical gates do not take the form (4.33). Code switching schemes also violate condition (b) since transversality is not fixed. Quantum computing by non-abelian anyons, when realized in a lattice system, violate conditions (b) and (c) since for braiding, implemented by finite or linear-depth circuits, the transversality is not fixed, and the system size needs to tend to infinity.

In the setting of topological codes, finite-depth local circuits, or locality (homology)-preserving gates are more general than transversal gates and apply to topological codes. Here the depth of a circuit is the length of the longest path from any input to any output subsystem. Below we use code super-block to denote a collection of a constant number of code blocks. A finite-depth local (FDL) circuit is defined by a unitary

$$U = \prod_L \bigotimes_J U_{LJ} \tag{4.37}$$

for L as the index of a finite number of layers, and J as the index of code super-blocks that U_{LJ} acts on. Note here U_{LJ} is local, namely, it acts on a constant number of code blocks which do not need to be neighboring. The definition is modulo permutation of code blocks. Given the existence of any permutation, the notion of neighbourhood

becomes trivial. Also the set of FDL circuits is not closed under multiplication since the depth will not stay as a constant. This is a drawback of the definition of FDL circuits.

Braiding of nanyons in general are *not* transversal, instead they can be realized by linear (instead of finite)-depth local circuits. As such, braiding does not preserve locality, so they can spread local errors out into a nonlocal ones. It is known that for a nanyon model, when the braiding is universal, then FDL gate will be trivial, e.g., the Fibonacci model. For Ising anyon, the braiding generates Clifford group, while FDL gates only form the Pauli group since Pauli gates are not realized via braiding, instead they are served by observable and order parameters. This is an intriguing tradeoff, and we see that universal braiding of nanyons does not have much relation with the Clifford hierarchy

$$P_D = \{U : \forall P \in P_1, UPU^\dagger \in P_{D-1}.\} \tag{4.38}$$

P_1 is the Pauli group, $P_0 = \mathbb{C}$. Braiding treats all the gates on the equal footing.

For topological stabilizer codes, it turns out the FDL gates relates to the spatial dimension of the code, which is stated by the Bravyi-König theorem. For D-dimensional code with local stabilizers (including subsystem codes), any FDL gate is contained in P_D in the Clifford hierarchy. Note that it does not claim the set of FDL gates is universal or not. For instance, in two dimension, there are only Pauli gates and Clifford gates, including CX, H, and S. To realize T gate as a FDL circuit, we need three dimension. A 3D color code has transversal T gate, and as a tradeoff, it does not have transversal H gate. To understand the theorem, first note that logical Pauli gates are string-like, which is the reason for the code being 'topological', in other words, its code distance increase with the system size, by at least logarithmic. Now a gate $U \in P_D$ when $UPU^\dagger \in P_{D-1}$, which means the support of P serves as part of the boundary of the support of U. This motivates a method based on partition of the whole system into the union of faces, edges and vertex. Then the support of a gate can be defined using these regions. As stabilizers commute with each other, the support of a gate can move around and sets to be the smallest one. As Pauli gates correspond to 1-dimensional strings, gates in P_D for $D > 1$ will correspond to higher dimensional supports.

4.4 Examples of universal fault-tolerant QC

Here we present several examples of universal fault-tolerant QC. These examples include anyon and nanyon, concatenated codes, and code switch methods. We do not discuss toric code since it requires magic state distillation, which is quite involved to explain.

Different codes hold different sets of transversal logical gates. If different codes can be combined together in a certain way, then their sets of gates may form a universal set.

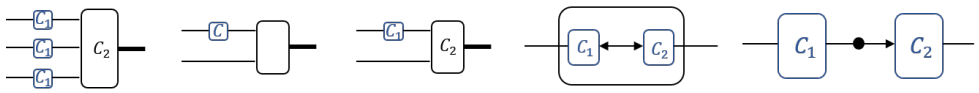

Figure 4.2: Quantum-codes networks. From left to right: concatenation, augmentation of C by non-coding elements, augmentation of C_2 by C_1, code switching, and codes junction (with black dot as interface).

Question 93. *How can codes be combined together?*

Just like resistors, capacitors etc can be combined in various ways, quantum codes can also be combined in many ways to achieve universality. There are at least these ways, see Fig. 4.2:

- Junction: codes can be connected together to form networks of codes.
- Switching: codes can be converted into each other by unitary or non-unitary means.
- Augmentation: a code can be augmented by other codes (or parts) which only play special limited roles in the main code.
- Concatenation: codes are put on different 'levels' such that one code becomes the elementary building blocks of another.

4.4.1 Topological QC via nanyons

There are many nanyon models that are universal, here we discuss the Fibonacci nanyon. Recall that an nanyon model is described by the fusion rule

$$\alpha \times \beta = \sum_\gamma N_{\alpha\beta}^\gamma \gamma \tag{4.39}$$

for α, β, and γ as nanyons. The dimension of an anyon is defined as $\log_d \Omega = n$ for a large number n of anyons and total dimension Ω. The dimension of Fibonacci τ is

$$d_\tau := \phi = \frac{1+\sqrt{5}}{2}, \tag{4.40}$$

which is bigger than 1.5, and bigger than that for Ising anyon $\sqrt{2}$. In the fusion rule, their dimensions satisfy

$$d_\alpha d_\beta = \sum_\gamma N_{\alpha\beta}^\gamma d_\gamma. \tag{4.41}$$

The only nontrivial fusion rule of τ is that

$$\tau \times \tau = \mathbf{1} \oplus \tau \tag{4.42}$$

for **1** as vacuum. There is no boson or fermion in this model. Usually we use 0 to label **1** and 1 for τ and this is helpful to study the fusion of many anyons.

What we use for quantum computing are the braiding of anyons, which are obtained from the so-called F-move and R-move. The R-move decides the effect of a twist on two anyons, and F-move decides the equivalence between different orders of braiding of many anyons, which is a unitary basis transformation. To encode a qubit we need at least 3 anyons since for two anyons, the two basis states, labelled by 0 and 1, cannot be superposed since they have different charges. For 3 anyons, we have $1\times 1\times 1 = 1+0+1$, so there are two states with charge 1. A common way to denote a state is by the fusion tree (history). We need to label the anyons from one to three, then the state (01) means the first two fuse to 0 and then fuse with the third to 1. We use $|0\rangle$ for (01) and $|1\rangle$ for (11). Now we have two braids: σ_{12} and σ_{23}, and also their inverse σ_{21} and σ_{32}. It is found from the F-move and R-move

$$\sigma_{12} = \omega^6 diag(1,\omega^7), \omega = e^{i\pi/5}, \sigma_{23} = F\sigma_{12}F, F = \phi^{-1}Z + \phi^{-1/2}X. \tag{4.43}$$

It is easy to see, since they do not commute, the two braids σ_{12} and σ_{23}, and their inverse, can generate the whole group of $SU(2)$ on a qubit.

With the same encoding, it is known that an entangling gate can also be realized between two qubits. However, there is leakage in addition to approximation due to gate compiling. For two qubits encoded in six (or eight) anyons, the total dimension is 13 while only four states are used for the two qubits. The braidings can induce a tiny leakage out of the code space. Fortunately, the leakage is not intrinsic and can be made arbitrarily small. This shows the universality of Fibonacci nanyon.

Despite topological, there are noises including, but not limited to, the followings:

- thermal excitations that are not used for qubits: they could travel on the system or interact with logical anyons. This can be suppressed by lowering temperature and control of the logical anyons.
- braiding of anyon: may not return to the exact location. This can be compensated by subsequent measurement which pairing up anyons.
- non-adiabatic braiding: braiding time shall be big $t \gg 1/\Delta$. Also it shall be faster than a noisy error, so $t \ll Te^{\Delta T}$, which is the time for a noise to induce a logical gate.

Question 94. *Is braiding perfect, ideally?*

Despite topological protection and universality, and suppose noises can be suppressed, the gates from braiding do not appear beautiful. For instance, there is no short-enough sequence of braids to realize Pauli gates and Clifford gates. The algebraic structures of braiding of Fibonacci nanyon and Clifford hierarchy do not match each other well. In addition, there is no physical system, except fractional quantum hall liquids, that might support Fibonacci nanyon. Also, it is very difficult to braid

nanyons since the process is not transversal, namely, on-site product unitary. So physicists also try to develop topological computing models based on other kinds of nanyons or gates.

4.4.2 3D gauge color codes

A 3D gauge color code is a self-dual CSS code so that it has transversal H and CX gate. The self-duality is required to have a transversal H gate, while the CSS feature itself guarantees transversal CX gate. However, it does not have transversal T gate. Instead of using magic state, the code switch method can be employed to switch between gauge color code and a color code, which has transversal T gate. The switch is done by the measurement of the stabilizers of the color code, and apply corrections, similar with error correction on stabilizer codes.

Color codes are toy models or software type codes. No 'color code phase' of many-body system is well defined yet. It has transversal T gate due to the magic of 'eight'. The code space is the group subspace of a commuting model

$$H = -\sum_{c\in C} B_c^X - \sum_{f\in F} B_f^Z, \tag{4.44}$$

for X stabilizers B_c^X on each cell, and Z stabilizers B_f^Z on each face. Each face (cell) has a number of site as a multiple of four (eight). So not all 3D lattices can be used to define color codes. Electric charge is point-like and created by string operators, while magnetic charge is loop-like and created by membrane operators, similar with 3D toric code. Logical X_L and Z_L are both global

$$X_L = \otimes_n X_n,\ Z_L = \otimes_n Z_n, \tag{4.45}$$

yet their weights can be reduced to be linear with the system size by products with local stabilizers. The total size of the system L is assumed to be odd. It turns out logical CX is transversal

$$CX_L = \otimes_n CX_{n_1 n_2}, \tag{4.46}$$

while logical H_L is not since it does not preserve the stabilizers (Hamiltonian). The T gate is transversal

$$T_L = \otimes_n T_n, \tag{4.47}$$

now let's see why. First, the lattice is required such that faces (cells) have a number of sites as a multiple of four (eight). The codeword $|0\rangle$ can be written as a loop condensation

$$|0\rangle = \sum_{v_8} |v_8\rangle \tag{4.48}$$

for $|v_8\rangle$ as state of product of zeros except ones with a number as a multiple of eight. The codeword $|1\rangle$ is the flip of $|0\rangle$. Now it is easy to see $T_L|1\rangle = e^{i\frac{\pi}{4}l}|1\rangle$, $l = L \mod 8$, and $T_L|0\rangle = |0\rangle$.

4.4.3 Triorthogonal codes

Triorthogonal codes are a type of CSS stabilizer codes that have transversal CCZ gate. The logical H gate can be done by code switch, also called gauge fixing. The set $\{CCZ, H\}$ is universal.

For S and G as stabilizer group and gauge group, a code with S_2, G_2 can be obtained from gauge fixing on S_1, G_1 when $S_1 \subseteq S_2 \subseteq G_2 \subseteq G_1$, Let's call the code 1 as 'source' code, and 2 as 'target' code. We see that the source code has smaller stabilizer group but larger gauge group, while the target code has larger stabilizer group but smaller gauge group. So fixing the gauge part of the source code can induce new stabilizers for the target code.

Triorthogonal codes are defined in a classical way. No physical system is known to be a triorthogonal code. For a $m \times n$ binary generator matrix G, with rows $f_i \in \{0,1\}^n$, it is triorthogonal when $|f_i \cdot f_j| = 0 \mod 2$ $|f_i \cdot f_j \cdot f_k| = 0 \mod 2$. The weight of each row f_i (number of 1s) can be even or odd. Now X stabilizers are obtained by mapping 1 to X for each even-weight row, Z stabilizers are obtained by mapping 1 to Z for each even-weight row in the orthogonal complement $G^\perp$. Logical X_L and Z_L are from odd-weight rows of G.

Let G_0 be the linear span of all even-weight rows of G, then $G_0 \subset G^\perp$. This means each even-weight rows of G maps to a X and a Z stabilizers, while there are additional Z stabilizers for $G^\perp$ G_0, denoted by $\tilde{Z}$. These $\tilde{Z}$ are the obstacle for transversal H gate.

For a triorthogonal encoding of a single qubit, we need a G with a single odd-weight row, $f_\star$. Let $G_1 = f_\star + G_0$. Then the codewords are

$$|a\rangle = \frac{1}{\sqrt{|G_a|}} \sum_{g \in G_a} |g\rangle,\ a = 0, 1. \tag{4.49}$$

Note that neither the size of the system n nor the code distance is fixed. The transversal CCZ is obtained from

$$CCZ^{\otimes n}|a,b,c\rangle = \sum_{g,h,i} (-1)^{|g \cdot h \cdot i|} |g,h,i\rangle = (-1)^{abc}|a,b,c\rangle. \tag{4.50}$$

For logical H gate, we first apply transversal $H^{\otimes n}$, which will disturb $\tilde{Z}$. Then with gauge fixing the correct values of $\tilde{Z}$ can be restored. When the value of $\tilde{Z}$ is -1, Pauli X correction is applied. The gauge fixing is fault tolerant.

4.4.4 Concatenated codes

Concatenation is a useful method to construct better codes, reduce error-threshold, and enlarge the set of logical gates.

Concatenated error-detecting (ED) codes can make error-correcting (EC) codes. ED codes can only detect errors but cannot correct them. One reason for this is

it cannot locate the errors. By concatenation, the error locations can be identified with higher-level parity-check operators. The Shor code uses 9 qubits to encode 1 qubit. It is the concatenation of 3-qubit bit-flip code with 3-qubit phase-flip code. An architecture for universal quantum computing with concatenated codes is the Knill C_4/C_6 code. The C_4 code uses 4 qubits to encode 2 qubits, and C_6 code uses 6 qubits to encode 2 qubits. Both are ED stabilizer codes. The C_4/C_6 EC code uses 12 qubits to encode 2 qubits. The Shor code and C_4/C_6 code are both stabilizer codes that correct one error, so they would not provide a universal transversal gate set.

A universal scheme is the concatenation of the 15-qubit Read-Muller code and the 7-qubit Steane code, with the later as the outer code. A code is an outer code if the subsystem in it is from the inner code. The Read-Muller code has transversal T gate and CX gate, and Steane code has transversal S gate, H gate, and CX gate. For the concatenated code, the logical H gate can be done transversally on the Steane code, while each H gate is not transversal on the Read-Muller code. Local error on the Read-Muller code will propagate possibly inducing an error. Such an error can be corrected on the level of Steane code. Meanwhile, the logical T gate can be done non-transversally on the Steane code, while each gate is transversal on the Read-Muller code. Local error on the Steane code will propagate possibly inducing an error. Such an error can be corrected on the level of Read-Muller code. We see that the effective code distance is three, i.e., it can correct one error. So it is a $[[105,1,3]]$ non-stabilizer code. In addition, we can also use the Read-Muller code as the outer code. The two schemes are not equivalent, though, in terms of resource usages (e.g., circuit size or depth) since the non-transversal gates on the two codes require different amount of resources.

4.5 Quantum algorithms

Question 95. *How quantum computers solve problems?*

If you think a quantum computer is just a quantum circuit, then running the circuit and make desired measurements solve problems. However, the question is more than this. Quantum circuits are unitary, which are operators, but if you want to know an answer to some problems, you have to convert them to classical objects, i.e., numbers or bit-strings. Quantum superposition of bit-strings yields speedup on one hand, but it also leads to complications on the other hand.

An algorithm is a procedure that maps an input to an output. But an algorithm cannot come from nowhere, instead it has to be designed by something, could be ourselves, or just another algorithm. Therefore, there are four cases:

1. C-C: a classical algorithm that designs a classical algorithm.

2. C-Q: a classical algorithm that designs a quantum algorithm.

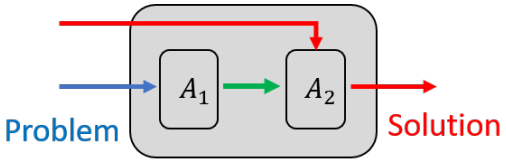

Figure 4.3: The complete model of quantum algorithm. The problem is the input (blue) to the first program A_1, and the solution is the output (red) of the second program A_2, which is the output from A_1 (shown as green) and has additional input data (red arrow).

3. Q-C: a quantum algorithm that designs a classical algorithm.

4. Q-Q: a quantum algorithm that designs a quantum algorithm.

The C-C type is the usual framework in classical computation. The C-Q type is the usual framework in quantum computation. How about the other two? Remember that the second algorithm is the one that are used to solve problems. So the Q-C type is like a quantum-enhanced classical computation, which is yet not widely used. The Q-Q type is not well understood, neither.

The whole algorithm can be shown in Fig. 4.3. The problem is the input to the first algorithm, and the solution is the output from the second algorithm, which itself, or in encoded form, is the output from the first algorithm. The input to the second algorithm is called 'ancilla' state.

Question 96. *How does a Q-Q algorithm work?*

From operator-state duality, we know that an operator can be represented as a state via vectorization. This serves as a stored-program scheme by converting an algorithm, which is a circuit U, into data, which is a quantum state $|\psi_U\rangle$. This is a generalization of quantum teleportation. So we find that, in the Q-Q algorithm, the first quantum algorithm Q_1 is the circuit that yields the state $|\psi_U\rangle$, and the second quantum algorithm Q_2 is the teleportation scheme that enables the action $U|\phi\rangle$ on 'ancilla' states $|\phi\rangle$, and the output is extracted from $U|\phi\rangle$. Where is the input as the problem? It cannot be the input to Q_1. As the result, there must still be a classical algorithm C_0 that takes the problem as input and designs Q_1, which further designs Q_2. This applies to all four types of frameworks of algorithms, namely, the chain has to start from a classical algorithm. The reason is anthropic: we are classical, unfortunately. This sounds not very exciting since we love quantum so we hope everything is quantum. But, we do not have to care about the classical starting point all the time; instead, the Q-Q part is the interesting part that matters.

Furthermore, there are different schemes for algorithm Q_1. Besides operator-state duality, there are also other ways to represent states. There are at least two of them: the state $|\psi_U\rangle$ can be written as $W|0\rangle$ for the trivial state $|0\rangle$ and a unitary W, which

is not unique, and also can be written as a matrix-product state. Therefore, Q_1 can be the circuit for W, or the circuit to prepare the matrix-product state, which is a quantum Turing machine.

The Q-Q algorithm applies to quantum simulation, which is to approximate a unitary U by another one U'. The gate U' can be stored in the three ways described above. As we will see, quantum simulation is usually done in the C-Q framework, in which the gate U' is directly executed. Comparing Q-Q to C-Q, the difference is that in Q-Q algorithm the gate U' is firstly stored as a quantum program, so that its action can be executed on demand, while in C-Q algorithm the gate U' is stored in encoded version $[U']$ as the output of the classical algorithm. The Q-Q framework serves as a foundation for stored-program quantum computers.

In quantum computing, most algorithms are the C-Q type. There is a classical algorithm C that takes the problem as input and yields the quantum circuit Q as output, and the measurement on the final state $|\psi_f\rangle = U|\psi_i\rangle$ with initial state $|\psi_i\rangle$ leads to classical answer, combined with some possible classical side-processing, to solve the original problem. In the following, we will focus on the C-Q type algorithms.

Question 97. *Are there different types of C-Q algorithms?*

It turns out there are. We can ignore the classical design algorithm C for a moment, and just focus on the quantum algorithm Q, which aims to solve problems. As the problems can be either quantum or classical, we find there are three types:

1. C-C: given a classical object, find another classical object.

2. Q-C: given a quantum object, find another classical object.

3. Q-Q: given a quantum object, find another quantum object.

The C-C type solves classical problems by encoding the problem in a quantum way. Examples include factoring, quantum phase estimation, amplitude amplification, quantum sampling etc. It deals with classical objects, namely, numbers, bit-strings, or probability vectors, but solves the problems by quantum circuits. The Q-C type converts quantum objects to classical objects, which apply to cases when a function of operators need to be evaluated. Examples are algorithms to compute eigenvalues, trace, determinant, or permanent of matrices, expectation values of quantum observable, partition function, etc. The Q-Q type solves quantum problems. Examples are various quantum simulations, including digital and analog types, and quantum state generations.

All three types play central roles in quantum algorithms. Quantum simulations are of main interest to physicists, as originally boosted by Feynman in 1980s. Later on, the boosts for quantum computing comes from Shor's factoring algorithm and Grover's search algorithm, which are of C-C type, and the evaluation of topological invariants in the setting of topological quantum computing, which are of Q-C type.

In addition, note that classical simulations of quantum processes, which is the field of numerical physics, do not belong here. It is instead classical computation. Also there is no C-Q type quantum algorithms since it doe not make much sense to use quantum objects to approximate classical ones.

4.5.1 Computational complexity

Computational complexity are the notion that physicists do not care much about, but they are interesting. They are like the principles for computing processes, just like the principles in physics, e.g., the laws in thermodynamics. Physicists like to find new things in nature, mathematicians like to prove theorems, with natural assumptions, while computer scientists like to solve problems found or created by physicists and mathematicians. The theory of computational complexity makes computation a mature science.

Question 98. *What are 'problems' in computer science?*

Problems can be classified, for instance, there are decision problems, search problems, counting problems, optimization problems, and function problems etc. The decision problems are the most common ones. Note that a problem class is a set of various problems. Complexity are measured by space, time, or other notions such as depth, width, etc. It turns out problems can be grouped according to complexity, hence complexity classes. A complexity class is a set of problems of equivalent complexity. There are separations between complexity classes. Some well known ones are as follows.

- P: all decision problems that can be solved easily, i.e., by classical computers using polynomial time.
- BPP: all decision problems that can be solved by stochastic computers using polynomial time with bounded error.
- BQP: all decision problems that can be solved by quantum computers using polynomial time with bounded error.
- NP: all decision problems whose solutions can be verified by classical computers using polynomial time, i.e., the solutions can be solved easily by non-deterministic computers using polynomial time.
- QMA: all decision problems whose solutions can be verified by quantum computers using polynomial time.

The relations among those classes are still unclear. It is certain that

$$P \subseteq BPP \subseteq BQP \subseteq QMA,\ P \subseteq BPP \subseteq NP \subseteq QMA \tag{4.51}$$

The most famous problem is whether P equals NP. This problem relates to many others, such as whether P equals BPP, BPP equals BQP, NP equals BQP, etc.

There is a feature of 'complete problem' in a class: a problem is complete iff it is as hard as every other problem in this class under reduction equivalence. There are P-complete and BQP-complete problems. There are no BPP-complete problems, however. The model of stochastic (or probabilistic) computation is not well understood, in general. A well-known QMA-complete problem is the k-local Hamiltonian problem, namely, to find the size of the gap of a k-local Hamiltonian, which is a Hamiltonian containing polynomial number of k-local terms. Another QMA-complete problem is the satisfiability problem, which is to determine whether there is a global state, given the local states of a many-body system.

Another feature that shall be kept in mind is that the definition of complexity classes are very delicate to the 'bounded-ness' of cost: e.g., there is a class PP which is different from BPP since its success probability cannot be boosted.

4.5.2 Examples of quantum algorithms

The usefulness of quantum computers relies on quantum algorithms that can solve problems. There are now hundreds of quantum algorithms, here we only review some of the primary ones.

Hadamard trick

Question 99. *How to understand the Hadamard gate?*

The Hadamard gate switches between Pauli X and Z. The Hadamard gate is also a *reflection* besides Pauli operators, since their eigenvalues are ± 1. The Toffoli gate is classically universal. With Hadamard gate, it becomes universal for quantum computing. Apparently, Hadamard gate is the magic for quantum computing. A usual way to understand its power is that it generates interference.

Now suppose we encode the problem and solution as quantum states, and a quantum state takes the form

$$|\psi\rangle = \sum_i \psi_i |i\rangle. \tag{4.52}$$

The basis $\{|i\rangle\}$ are usually bit-strings. If they are not, i.e., $|\psi\rangle = \sum_i \psi_i |\phi_i\rangle$, there always is a way to convert the state to the standard form in the equation. The information we need to manipulate are the coefficients (or amplitudes) ψ_i. They are complex numbers $\psi_i = |\psi_i| e^{i\theta_i}$. The size $|\psi_i|$ is also called amplitude, the angle θ_i is called phase. The evolution is to change both the amplitude and phase. Two different tasks are to change them or estimate them. To estimate amplitudes, we can perform measurement since $|\psi_i|^2$ is probability. To change phase, we can apply unitary evolution on the state directly. That is to say, the nontrivial tasks being left are phase estimation and amplitude change. These are the quantum phase estimation algorithm

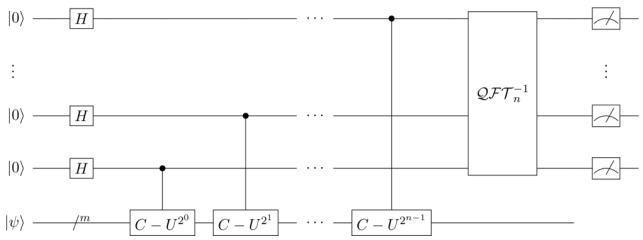

Figure 4.4: Quantum phase estimation algorithm.

and quantum amplitude amplification algorithm, which is a slight generalization of search algorithm.

Quantum phase estimation

The quantum phase estimation algorithm is stated as follows. Given a unitary operator $U \in SU(N)$ and an eigenstate $|\psi\rangle$ of it, $U|\psi\rangle = e^{i2\pi\theta}|\psi\rangle$, the problem is to estimate the angle θ with high probability within accuracy ε. It assumes the state $|\psi\rangle$ is a known eigenstate and can be prepared, and quantum controlled-U^n gates can be performed. The algorithm contains three stages: Hadamard array, controlled-U^n array, and quantum Fourier transform. It uses qubits ancilla which are measured at the end. See Fig. 4.4.

It uses the so-called 'phase kick back' trick: although the phases are on the state $|\psi\rangle$, each phase factor can be kicked back to the ancillary qubits so that the phase information are encoded by them. Furthermore, there is no entanglement despite the controlled-U^n array, but there are correlations between the ancilla and the state. Such correlations are 'classical' since they are not entanglement, which is quantum correlation.

To convert phases into amplitudes of ancillary qubits, the quantum Fourier transform (QFT) is applied. The QFT are qudit generalization of Hadamard gate, which can be defined as

$$QFT : |x\rangle \mapsto \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \omega_x^k |k\rangle, \omega_x = e^{-i2\pi x/N}. \tag{4.53}$$

For n-qubit, the Fourier transform does not generate entanglement, although it needs entangling gates for gate synthesis. The QFT further generates classical correlations among the ancillary qubits. Now the value of θ is encoded in the bit values of ancillary qubits, which are measured to extract it. Depending on whether $2^n\theta$ is an integer or not, the readout may have a success probability smaller than one, which can be boosted by the increase of the number of ancillary qubits.

It turns out quantum phase estimation is powerful. It plays central roles in Shor's factoring algorithm, and also the abelian hidden subgroup problems. However, its assumptions also yield limits, namely, it is difficult to find eigenstates of operators

and to perform controlled operations. Finding eigensystems are the tasks for many algorithms. Finally, note that it is different from quantum estimation, which does not assume the unknown parameter as a phase, and the main scheme is via many rounds of measurements.

Search by reflections

Householder reflection is a generalization of Hadamard gate. Householder reflection $R_\psi = 2|\psi\rangle\langle\psi| - \mathbb{1}$ about a state $|\psi\rangle$. A reflection is self-inverse, i.e., hermitian and unitary, $R_\psi^2 = \mathbb{1}$, and its eigenvalues are ± 1. Pauli operators and Hadamard gate are reflections.

Gorver's search algorithm uses reflections for search. A target to be found is given with known form. Given a target $|t\rangle$, the search algorithm starts from trivial state $|0\rangle$ with $\langle t|0\rangle \neq 0$ generically, and choose an orthogonal state $|t^\perp\rangle$, then Grover operator G as a product of two reflections $G = R_0 R_{t^\perp}$, and the algorithm is iterative G^L for a certain step L such that the final state approximates $|t\rangle$ well.

The number of steps $L \in O(\sqrt{N})$ for $N = \log D$, D as the Hilbert space dimension of state $|t\rangle$. The Grover algorithm is optimal since the lower bound is $\Omega(\sqrt{N})$. Furthermore, quadratic speedup for unstructured problems is typical due to superposition, while exponential speedup often occurs for structured problems which allow heavy classical pre- or post- processing.

Quantum gate synthesis

Quantum gate synthesis is a problem of quantum simulation. Given any n-qubit gate, using sine-cosine decomposition it can be exactly decomposed as product of Givens rotations. A Givens rotation can be decomposed as products of CX and single-qubit rotation. Any qubit rotation can be written as $R_z(\alpha)HR_z(\beta)HR_z(\gamma)$, so we only need to decompose rotation along z-axis $R_z(\alpha)$ for any angle α. One universal gate set is $\{H, T\}$. Using a binary rep of angle $\alpha \approx m\pi/2^n$ for an integer n and m, then we need to get angle $\pi/2^n$. It turns out $\{H, T\}$ cannot generate diagonal gate with angle $m\pi/2^n$, i.e., a gate $R_z(\alpha)$ has to be approximated by a non-diagonal gate. This is based on number-theoretic study.

Another common task is Hamiltonian simulation, which is to simulate the evolution $U = e^{itH}$ for a sparse or local Hamiltonian H. The problem looks trivial as one only needs to decompose U as a product of the form $U \approx U' = \prod_n U_n$ for each $U_n = e^{it_n H_n}$ with a simple term H_n (see Chapter 2). The nontrivial part is to determine how 'simple' the terms H_n could be. In the context of many-body system, there is a notion of locality and geometric locality, the terms H_n are expected to be geometrically local. It turns out this can be done rather easily since local terms H_n can be engineered in artificial systems, such as trapped ions or superconductors. This is why quantum simulators are already built for use.

Many others

There are quantum algorithms that are of interest to physicists: algorithms that are quantum versions of traditional numerical algorithms. These include, but not limited to, quantum annealing, quantum optimization, and quantum algorithms to compute partition function. These algorithms are common in numerical physics, and the motivation is to see whether quantum superposition can enhance them. Quantum annealing, which uses quantum fluctuations and quantum evolution itself, aims to find global minimum faster than classical computers. Quantum machine learning, as an optimization problem, is not well studied yet. Optimization problems are generically hard, and very different from decision problems. The quantum advantage for optimization problems needs more investigation.

Chapter 5

Condensed Matter Physics

Question 100. *What is condensed matter about?*

It studies the property of 'macro-objects' made of a large amount of parts which have certain symmetries. Given a macro-object, it can be treated as a 'black box' or 'white box'. When it is a black box, the detailed structure or property in it is not clear; however, itself as a whole can be studied under various probes, such as x-ray, neutron, light, heat, mechanical, chemical, electric etc. One central issue of study is phase transition. If a microscopic model can be established, it becomes a while box, and more control techniques can be developed, and it can be used for other purposes, including applications in other settings, design of new materials.

5.1 Symmetry

Question 101. *What are order parameters?*

Order parameter labels different 'phases' or 'orders'. Phases have symmetry properties. So order parameters must relate to symmetry. This is indeed the case. As symmetry is described by group theory, order parameters are 'charges' of symmetry, i.e., some characters of symmetry. Examples are density-wave, polarization (location-dependent), magnetization. Furthermore, order parameters shall be measurable in practice, and some measurable observable that are not 'charges' of symmetry are often used as order parameters, such as correlation functions and gap functions.

The symmetry of a model H is a group G such that

$$[H,G]=0. \tag{5.1}$$

That is, any $g\in G$ preserves H. To list all the symmetry of a model on the first hand is not an easy task. The other way is to define models that has a given set of symmetry. The interplay between symmetry and modeling guides both theorists and experimentalists to find new materials.

There are many symmetry to consider, since there are so many groups. For quantum system, it is known that a symmetry can be represented in either unitary or anti-unitary way. The most common anti-unitary symmetry is the time reversal, T, which plays central role in fermionic systems.

As a many-body system occupy a spacetime volume, i.e., a manifold, there is a notion of locality, i.e., neighborhood, such that interaction terms can be treated as local or nonlocal. A general way to define symmetry is by the notion of q-form: for d-dim system, a symmetry is q-form when it acts on $(d-q)$ manifold. So a global symmetry is a 0-form symmetry, a gauge symmetry is a d-form symmetry, and other intermediate symmetry are generally q-form.

A model is defined by its spectrum, i.e., ground states and excitations. Sometimes ground states have larger symmetry than the model itself, and these extra symmetry are called 'emergent symmetry' on ground subspace. Namely, a symmetry G' is emergent when

$$\langle\Omega|[H,G']|\Omega\rangle = 0 \tag{5.2}$$

for ground states $|\Omega\rangle$. Usually at critical point there are emergent symmetry.

Sometimes a symmetry is not exact, instead it becomes exact in a certain limit. These symmetry are called quasi-exact, or quasi, symmetry. The emergent symmetry is a quasi symmetry in the sense that it becomes exact when the temperature is very low such that the system stays on ground states. Another common limit is the large-system limit, i.e., thermodynamic limit.

Symmetry, by definition, is (anti-) unitary. Symmetry, also by definition, preserves order parameters. There is a slight generalization of symmetry that can change order parameters in a reasonable way, which is known as *duality*. Duality cannot be realized by a unitary process on the system itself, although it can be unitary with extra ancilla. Usually, duality transformation only preserves local dynamics but the global (topological) properties will be lost. For instance, the quantum Ising model has the duality between order (FM) and disorder (PM) parameters. Yet the FM has two-fold degeneracy, while PM does not. Duality, $\mathscr{D}$, which is a superoperator and independent of λ, applies to a parameterized Hamiltonian $H(\lambda)$ such that

$$\mathscr{D}(H(\lambda)) = H(\lambda'). \tag{5.3}$$

The new model $H(\lambda')$ shall take the same form except the new parameters λ'.

If we do not stick to the apparent form of a model, there are unitary equivalence relation between H_1 and H_2 when there exists a unitary U such that

$$H_2 = UH_1U^\dagger. \tag{5.4}$$

Such a unitary equivalence is different from duality since it is unitary (preserves spectrum) but it does not preserve the form of the model. An example of this is the Jordan-Wigner transformation that maps between spins and fermions.

There are also 'correspondence' relation. The correspondence between two models H_1 and H_2 is a weak equivalence

$$\mathrm{tr} e^{-\beta_1 H_1} \propto \mathrm{tr} e^{-\beta_2 H_2} \tag{5.5}$$

since it refers to classical variables instead of operators. The AdS/CFT and bulk-boundary correspondence are examples.

An even weaker, but powerful relation among models is renormalization.

Question 102. *What is the essence of renormalization?*

The essence is that it trivializes the form of a Hamiltonian H. By this, it means that to obtain the possible phases, the exact form of Hamiltonian does not matter. What matters are the set of all possible Hamiltonians, $\{H\}$. A renormalization, $\mathscr{R}$, is in general not invertible and does not preserve Hilbert space dimension, and it changes the form of the model. Repeated action of it shall drive the system to fixed points

$$\mathscr{R}^n(H) \to \{H_*\} \tag{5.6}$$

that label different phases. However, there is no principle for how to define $\mathscr{R}$ given a model H, and the existence of fixed points is not guaranteed. The usual way is to use coarse-graining to ignore short-distance details so only the long-distance features, which are 'universal', are extracted. This agrees with quantum field theory which also describes long-distance features, or low-energy sectors, of many-body systems.

Under renormalization, some interaction terms are relevant, and some are irrelevant, and some are marginal. Whether a term is relevant or not depend on scaling dimensions, and they can be obtained by renormalization, and also conformal field theory. A term with zero conformal spin and scaling dimension d is relevant if $d < D$, and irrelevant if $d > D$, and marginal if $d = D$ and depends on the sign of coupling g. Large scaling dimension implies that correlation functions decays fast. Here D is spacetime dimension. When conformal spin is not zero, there are also modified versions. This is an intriguing relation between scaling dimension and spatial dimension. There is also a similar relation in the setting of topological quantum computing: the logical gates are related to the spatial dimension of a topological stabilizer code.

Phases are defined and classified according to symmetry: what they are, how they are represented on the system, how they are dealt with by ground states, broken or preserved, how they act on excitations or defects, etc. When symmetry is preserved, it is called symmetry-protected or enriched phase. When symmetry is broken, it is called spontaneous symmetry-broken phase. Classification of phases is based on group and group cohomology, related to central extensions of group. The reason for using group cohomology $\mathscr{H}^n(G,F)$ is that the symmetry may only be represented projectively instead of linearly on the system. A projective representation is a linear representation up to a constant, which usually is the global phase of a quantum state. There is also an interesting relation between the level of cohomology n and the dimension of the system: if G is q-form, then $n = d + 1 - q$.

Changes between phases are called phase transition. Traditional classifications yield first order and second order phase transition depending on how the partition function changes. However, the 'order' of the transition does not tell much information. In quantum theory, a phase transition is labeled by closing of gap, or change of correlation function, or edge states. Quantum systems at critical points or regions are described by conformal field theory, or by matrix-product states with large bond dimensions. Critical models are defined by universal features that do not depend on short-distance details, and these universal features, such as scaling index, can be derived by renormalization or conformal field theory.

5.2 Ising model

The quantum Ising model is

$$H = -\sum_n \sigma_n^x - \lambda \sum_n \sigma_n^z \sigma_{n+1}^z. \tag{5.7}$$

We now use duality to find its critical point. The duality is defined by

$$\tau_n^x := \sigma_n^z \sigma_{n+1}^z, \tau_n^z := \prod_{m \le n} \sigma_m^x \tag{5.8}$$

on the dual lattice, the inverse duality is

$$\sigma_n^x := \tau_{n-1}^z \tau_n^z, \sigma_n^z := \prod_{m \ge n} \tau_m^x. \tag{5.9}$$

The model can now be written as

$$H = -\lambda \sum_n \tau_n^x - \sum_n \tau_n^z \tau_{n+1}^z. \tag{5.10}$$

The critical point is $\lambda_c = 1$, and large (small) λ corresponds to the low (high) temperature phase. The high-temperature phase has zero 'order' $\sum_n \sigma_n^z$ while a nonzero 'disorder' $\sum_n \tau_n^z$, and the opposite for the low-temperature phase.

Question 103. *How to understand the Ising duality?*

There are two types of excitations in Ising model: spinon and vison. Spinon is the flip of one energy term, while vison is for two. Use 0 and 1 for spin states, then a spinon is like 11000, and the term for 10 is flipped. A vison is like 00100, and the terms for 01 and 10 are flipped. A spinon (vison) is also called a kink (breather), and we can see that a vison is a bounded pair of spinons. Now the duality is to use the dual variables which labels the 'link' states: when the two states are (not) the same, we use 0 (1). So the dual state of a spinon 11000 is 0100, which is a vison, the dual

state of a vison 00100 is 0110, which is a spinon. A combination of spinon and vison is known as a Majorana fermion, since it maps to itself under duality.

The duality can be done by a quantum circuit with CX gates. Locally, the duality encodes the parity of a neighboring pair by a spin: it is 1 (0) if there is (no) domain wall. The parity measurement is done by applying CX from each of the control spin to the target spin, which is the dual site. For a system of N spins, there are also N links so N dual spins. The duality is not one-to-one, so it does not preserve orthogonality. For a system state $|\phi\rangle$, and the initial dual state $|\vec{0}\rangle$, the sequence of CX as U, then the final dual state is

$$|\tilde{\phi}\rangle = \langle\phi|U|\vec{0}\rangle|\phi\rangle. \tag{5.11}$$

The process is a *complementary* channel since it maps the system state to the ancilla state for the dual system. Crucially, it is not unitary. Only half of the dual space states are used, e.g., it maps both $|0000\rangle$ and $|1111\rangle$ to $|0000\rangle$. It does not preserve the degeneracy.

The critical point is described by Ising CFT with central charge $c = 1/2$, i.e., one Majorana fermion, which is a product of order and disorder operators. The order describes the phase when spinons dominant, so there are large domains, and the disorder describes the phase when visons dominant, so there are no large domains, hence 'disordered'. This is the case since visons have higher energies. The spinon and vison are massive. The fermion ψ is massless at the critical point. There are three primary fields: 1, a fermion ψ, and the Ising anyon σ. Denote spinon as s, vison as v, we have the fusion rules

$$s \times s = 1, v \times v = 1, s \times v = \psi, \psi \times \psi = 1, \tag{5.12}$$

and also the non-Abelian one

$$\sigma \times \sigma = 1 + \psi, \sigma \times \psi = \sigma. \tag{5.13}$$

The Ising anyon σ only appears at the critical point. Note that 'fusion' refers to the Hilbert space structure. It does not mean the two particles have to be together, or pass through each other.

It turns out they can be realized by defects. The defects are modification of the original Hamiltonian 'by hand'. Invertible defects can be realized by unitary operations. They appear at least in pairs and are created by a 'symmetry twist': a segment of a symmetry operator. The symmetry defect is a spinon, and a vison is a bounded pair of them, and a ψ is a spinon and a vison. A vison can be 'eaten' by a spinon in the sense that when a vison encounters a spinon, it will join it. So a ψ can be viewed as a spinon. The ψ defect now can be viewed as a 'symmetry defect' so a pair can be created by $\otimes_{i\le n\le j}\sigma_n^x$ at link $(i-1,i)$ and $(j,j+1)$. It flips the terms $\sigma_{i-1}^z\sigma_i^z$ and $\sigma_j^z\sigma_{j+1}^z$.

Question 104. *How to make Ising anyons σ?*

The Ising anyon σ is non-Abelian, $\sigma \times \sigma = 1 + \psi$. It is called a 'duality' defect, but how to introduce in σ? Recall that

- spinon: states like $\dots 10 \dots$ or $\dots 01 \dots$, and 10 fuses with 01, and it moves by a single flip, and it separates different domains. Spinons appear in pairs.
- vison: states like $\dots 010 \dots$ or $\dots 101 \dots$, and 010 fuses with 010, and it moves by two neighboring flips, and it separates the same domain. Visons do not have to appear in pair.

At the critical point, it is likely there are equal number of spinons and visons. We find a duality defect is a three-site object: it uses the middle spin as the parity of the other two spins. It has four states: 000, 011, 101, 110, so it can be vacuum, a spinon, or vison. The middle spin can be introduced in as an external dual spin, so it is an 'extrinsic' defect. It connects to the system by two CX gates. Now to make it a defect, we have to replace one spin by the dual spin, say, the first spin. To make this clear, the process is

$$000(000)000 \to 000(00)000, 000(011)111 \to 000(11)111, \tag{5.14}$$
$$111(100)000 \to 111(00)000, 111(101)111 \to 111(01)111. \tag{5.15}$$

Now what is the fusion $\sigma \times \sigma$? We need to see the Hilbert space of two Ising anyons. It is clear to see that there are ground state (vacuum) and spinons, visons. This is the physical meaning for $\sigma \times \sigma = 1 + \psi$. What is the fusion $\sigma \times \psi$? As σ contains ψ, so it is clear that $\sigma \times \psi = \sigma$.

But how to move the Ising anyon? It cannot be moved by on-site operators, like spin flips. It turns out it can be moved with entangling gates or tuning of Hamiltonian parameters, which is not on-site (i.e., transversal) operations, though. This is the reason that non-abelian anyon braiding is not transversal, instead they are linear-depth local unitary circuits. We can summarize as

- Ising anyon: states like $\dots 000 \dots$, $\dots 011 \dots$, $\dots 101 \dots$, $\dots 110 \dots$, cannot move. Created as extrinsic defect. Can eat a Majorana fermion.

5.3 Ising world

Here we will survey various interesting things related to Ising. It turns out if you get familiar with the Ising world, you will appear as an expert on many-body physics.

5.3.1 Fermionization

We did not really solve the Ising model, namely, we do not have its spectrum yet. This can be obtained using the Majorana fermion ψ. Each spin is mapped to two

Majorana fermions. We define

$$\psi_{n,1} := \sigma_n^z \tau_{n-1}^z, \psi_{n,2} := i\sigma_n^z \tau_n^z, \tag{5.16}$$

then $i\psi_{n,1}\psi_{n,2} = -\sigma_n^x$, $i\psi_{n,2}\psi_{n+1,1} = -\sigma_n^z\sigma_{n+1}^z$. It is easy to check that they are indeed fermions: they anticommute. With Fourier transform $\psi_k \propto \sum_n \psi(n)e^{ikn}$, and the Dirac fermion $\gamma_k = \psi_k + i\psi_{-k}$, the model is $H = \sum_k E(k)(\gamma_k^\dagger \gamma_k - 1/2)$ for

$$E(k,\lambda) = \sqrt{1+\lambda^2 - 2\lambda\cos k}. \tag{5.17}$$

This is equivalent to the Kitaev chain: the trivial phase is the PM, and the nontrivial symmetry-protected topological (SPT) phase with edge Majorana modes is the FM. The two edge Majorana modes span a two-dimensional space, corresponding to the two-fold degenerate ground space of FM. They pair up to form an occupied Dirac fermion or not. The trivial phase has a unique ground state: all Majorana pair belonging to the same spin pair up to form an occupied Dirac fermion, which is the plus eigenstate $|+\rangle$ of X_n.

Question 105. *Is the FM phase topological or not?*

We know that, in the spin picture, the FM phase is SSB. In the fermion picture, it is the so-called SPT phase, protected by fermion parity symmetry Z_2^f. It is topological in the sense that the fermion parity is a global property, but it is also SSB. The Ising Z_2 symmetry is Z_2^f, the operator $\otimes_n X_n$ is mapped to $\otimes_n \psi_{n,1}\psi_{n,2}$. It is topological also because there will be Majorana zero mode, which behaves as the non-abelian Ising anyon σ under braiding. Note that the Majorana zero mode is not a Majorana fermion! We know that it is difficult to make an Ising anyon σ, while here it is also difficult to create a Majorana zero mode, which shall have zero energy, i.e., no fluctuation. Under fluctuation, the two edge Majorana fermions do not pair, and they are just simply Majorana fermions instead of anyons.

5.3.2 Bosonization

The Ising model can also be viewed as a system of bosons instead of fermions! This is like a magic, but this is true. The reason is of course mathematical (due to scaling dimension), but the physical reason is that, the fermions, bosons, and also spins describe different d.o.f of the same system.

Question 106. *What are the spins, fermions, and bosons for the same system?*

Roughly, the spins describe local sites in space, the fermions describe currents across the system, and the bosons describe global features as if the system is taken as a whole as a single system.

In bosonization, the fermion reduces to two species near the Fermi points: a right mover R and a left mover L. They are further written as 'vertex operators' as

exponents of bosons Φ and Θ, for $\Phi = \varphi + \bar{\varphi}$, $\Theta = \varphi - \bar{\varphi}$. The bosonic fields Φ and Θ are compactified, i.e., similar with angular variables, so they are like the angles to specify the coordinates of a vector. The interaction between the two movers will map to nonlinear terms of Φ, sometimes also for Θ.

We will discuss bosonization in more details for the XXZ model, which is more general than the Ising model. The bosonized form of the Ising model is

$$H = \int dx \Pi^2 + (\partial_x \Phi)^2 + m_\lambda \cos \sqrt{4\pi}\Phi, \tag{5.18}$$

for a mass m_λ depending on λ. $m_1 = 0$. This is the sine-Gordon model with $\beta = \sqrt{4\pi}$. The mass term has scaling dimension 1, so it is relevant. Depending on the sign of m_λ, we will have PM and FM phases.

The Ising model is usually mapped to the phi-four theory, namely, a Landau effective field theory. It can be viewed as a truncated sine-Gordon model by Tayor-expanding the cosine term to the order Φ^4, while the coefficients of Φ^2 and Φ^4 terms are treated independently. We will further learn that the sine-Gordon field theory is very powerful to describe 1D quantum systems. It is also worthy to mention that the usual mean field theory fails to predict the phase transition of 2D classical Ising model (i.e., 1D quantum Ising model), but with certain modifications, it can lead to the phi-four theory.

5.3.3 Gauging

Question 107. *Is gauging a fundamental process?*

As field (e.g., electromagnetic field) usually has local (gauge) symmetry, the process of gauging of a system is to couple it to external field. Gauging can promote a global symmetry to a local gauge symmetry. If the global symmetry is SSB, then the gauged model is SPT, hence gauging will reduce ground-state degeneracy.

The gauged Ising model is the cluster model

$$H = -\sum_n \sigma_n^y \sigma_{n+1}^y - \lambda \sum_n \sigma_n^z \sigma_{n+1}^x \sigma_{n+2}^z. \tag{5.19}$$

How is this constructed? The scheme is to couple the system and its dual together. Recall that there is the Ising duality. Now we modify $\sigma_n^z \sigma_{n+2}^z$ to $\sigma_n^z \sigma_{n+1}^x \sigma_{n+2}^z$, and modify σ_{n+2}^x to $\sigma_{n+1}^z \sigma_{n+2}^x \sigma_{n+3}^z$ to the model, which generates the local Z_2 gauge symmetry. All terms commute. The gauged model (without the YY terms) has a unique ground state: it is a cluster state $|C\rangle = \prod_n CZ_n |+\rangle^{\otimes N}$. Also the cluster state has SPT order by $Z_2 \times Z_2$ global symmetry: $XIXI\cdots$ and $IXIX\cdots$. We notice that one of the global Z_2 symmetry is the original Z_2 symmetry of the Ising model that is SSB. The model is a coupled 'matter-gauge' system, see Fig. 5.1.

The YY terms comes from a further application of the Ising duality. From the duality, we find the two terms will map to each other, so the critical point is $\lambda_c = 1$,

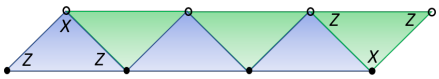

Figure 5.1: Gauging of 1D Ising model to 1D cluster model. Black (white & black) dots are spins in Ising (cluster) model, Ising terms ZZ changes to ZXZ (blue) as matter terms, X changes to ZXZ (green) as gauge terms.

on the two sides are FM with SSB order and SPT order. So we see that the SPT phase is the dual of the FM phase.

We know that FM also separate from PM by the Ising duality. So how about the SPT phase and PM? As already mentioned, there is a direct transition from PM to SPT order. The required action is to map $\sigma_n^z\sigma_{n+1}^x\sigma_{n+2}^z$ and σ_{n+1}^x to each other, and this is done by $U=\prod_n CZ_n$, i.e., a sequence of neighboring CZ gates. This is the way to prepare a cluster state, by the way. Furthermore, it is clear to see 'higher-order' cluster model can also be defined. The PM phase has trivial SPT order and does not break the global Z_2 symmetry.

We could apply the gauging procedure further. We now want to gauge a parallel array of Ising chains. We can put a gauge spin at the center for each four spins. Now the system is on a square lattice with one spin on each edge. We modify ZZ term from a chain to $ZZZZ$ with two additional Z acting on two gauge spins. The new weight-four Z terms, called plaquette terms, do not couple the original chains. To couple them, we add $XXXX$ as vertex terms, and this is the local Z_2 gauge symmetry. Now it is clear that the gauged model is the toric code. We see that Ising Z_2 symmetry becomes the 1-form symmetry acting on the original chains, the Ising local flux insertion (by loops of Z oeprators) extends to a new 1-form symmetry. The SSB order becomes the TOP order of toric code.

Note that, in the original formula of gauge models, the gauge symmetry terms are not included in the model. However, nowadays we usually include them to reduce ground state degeneracy.

We see that toric code SSB the 1-form symmetry. Can this be further gauged? The answer is yes, and it turns out the resulting model is the 2D cluster model. The 1-form Z_2 symmetry now is preserved, so it has SPT order by this symmetry. The ground state is unique.

5.3.4 Defects

Question 108. *Given a model, how to find the possible primary defects?*

You can introduce wild defects by whatever ways you want, but we want to keep it simple. So we look for primary defects. There are many kinds:

1. symmetry defects: created by a segment of symmetry operator on a ground state. Depending on how the symmetry behaves, e.g., broken, preserved, sym-

metry defects also behave differently. For SSB order, symmetry defects are domain walls. We say a symmetry defect (domain wall) 'implement' a symmetry operation. For SPT order, symmetry defects are excitations which are projective representations of the symmetry.

2. duality defects: when the system has a duality, there are duality defects. The duality maps excitations of the system to each other, e.g., electric and magnetic. A duality defect implements the duality.

The 1D Ising model has a symmetry defect ψ as a Majorana fermion (a combination of spinon s and vison v), and a duality defect σ as the Ising anyon. The duality is also known as the high-low temperature duality. A vison v is created by a disorder operator Z at a single site, which is a flux insertion, and a spinon s pair is by the order operator $\otimes_n X_n$. The braiding of s and v is simply to apply $\vec{X} = \otimes_n X_n$ in the presence of Z, and it is clear that they are mutually semion:

$$\vec{X}Z = -Z\vec{X}. \tag{5.20}$$

This is actually the same as the toric code (we can also see this later).

The toric code has a duality defect as the Ising anyon and a symmetry defect as a Majorana fermion $\psi = e \times m$ (a combination of spinon and vision). The duality is due to translation symmetry and maps between electric (spinon) and magnetic (vison) charges. The symmetry defect is created by a segment of 1-form symmetry (Wilson loop) that are SSB, and they are Majorana fermions. The toric code can be un-gauged to obtain 1D Ising models. Naturally, the 1D Ising model also appears as the boundary of the surface code. The Ising anyon, as a defect, cannot be easily moved, so it is very difficult to perform braiding.

5.3.5 XXZ model

Now we move on to models with higher symmetry, e.g. U(1) or SU(2). The XXZ model is

$$H = -\sum_n \sigma_n^x \sigma_{n+1}^x + \sigma_n^y \sigma_{n+1}^y + \Delta \sigma_n^z \sigma_{n+1}^z + h\sigma_n^z, \tag{5.21}$$

and it has U(1) symmetry of rotations around the z-direction. It includes the Heisenberg model for $\Delta = 1$ and $h = 0$, which has the full SU(2) symmetry. Note that the model can be defined for arbitrary spins. A famous result by Haldane is that the Heisenberg model is gapped for integer spins, while gapless for half-integer spins. We will understand this distinction using the valence-bond solids construction and matrix-product states.

The phase diagram of the spin-1/2 XXZ model is well studied. There are three phases: FM, AF, and XY phase, see Fig. 5.2. We can see this roughly: when Δ is very big, we will have the AF phase, when Δ is very small, we have the FM phase, while

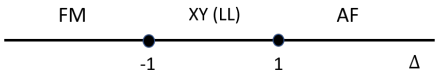

Figure 5.2: The phase diagram of the XXZ model.

if Δ is close to zero, we have a phase by the XX and YY terms, so the XY phase. The gapless XY phase exists for $-1 < \Delta \leq 1$, the easy plane anisotropy regime, and the spins are ordered on the XY planes. The gapless phase is also known as Luttinger liquid, which can be further gapped out be adding new terms. In all, the model can be described by sine-Gordon field theory.

We now solve the model by fermionization via Jordan-Wigner transformation (JWT). It maps spin operators to spinless (i.e., no index) fermions

$$\sigma_i^z := n_i - 1/2,\ \sigma_i^+ := (-1)^i \psi_i^\dagger (-1)^{\sum_{j=1}^{i-1} n_j}, \tag{5.22}$$

for $n_i := \psi_i^\dagger \psi_i$ as number operator. The long tiles are the distinct feature of JWT, and it does not preserve locality in the spin picture and the fermion picture. We see that $\sigma_i^+ \sigma_{i+1}^- \mapsto -\psi_i^\dagger \psi_{i+1}$ as free hopping term, $\sigma_i^z \sigma_{i+1}^z \mapsto (n_i - 1/2)(n_{i+1} - 1/2)$ as on-site interaction term. It is clear to see when $\Delta = 0$, it is a free fermion and gapless. The spectrum is the simplest one: $E(k) = \cos k$. The Fermion 'surface' now is only two points: $k = \pm\pi/2$.

Now we use perturbation theory to include Δ. Near the two Fermion surface, we linearize the spectrum then the model only contains two terms: a left mover L and a right mover R. In real space, the model is $H_0 = -i \int dx (R^\dagger \partial_x R - L^\dagger \partial_x L)$, which is the same as a free boson $H_0 = \int dx \{\Pi^2 + (\partial_x \Phi)^2\}$ for boson $\Phi(x)$ and the conjugate momentum $\Pi(x)$

$$[\Phi(x), \Pi(y)] = i\delta(x-y). \tag{5.23}$$

The momentum $\Pi(x) = i\partial_\tau \Phi$, and a dual field Θ is defined such that $\Pi = \partial_x \Theta$ and

$$[\Phi(x), \Theta(y)] = i\theta_H(x-y). \tag{5.24}$$

The observable of the theory are vertex operators $e^{im\Phi}$, also called spin-wave, or electric operators, and $e^{in\Theta}$, called vortex, or magnetic operators. They satisfy the vertex-operator algebra (VOA).

The spin operator maps to

$$\sigma_x^z = \rho(x) + (-1)^x M(x). \tag{5.25}$$

The first term $\rho(x) = \partial_x \Phi(x)$ relates to the magnetization, and the second term $M(x) = \sin\sqrt{2\pi}\Phi(x)$ relates to AF order. Note we have ignored some constants in front of operators for simplicity. In terms of fermions, $\rho(x) = R^\dagger R + L^\dagger L$, $M(x) = R^\dagger L + L^\dagger R$, and $R^\dagger L \propto e^{i\sqrt{2\pi}\Phi}$ as a vertex operator. The ladder operators are

$$\sigma^\pm(x) = e^{\pm i\sqrt{2\pi}\Theta(x)}[(-1)^x + \cos\sqrt{2\pi}\Phi(x)]. \tag{5.26}$$

The dimer operator $\varepsilon = (-1)^n \vec{\sigma}_n \cdot \vec{\sigma}_{n+1}$ describes the order when neighboring spins form singlets. It holds

$$\varepsilon \sim \cos\sqrt{2\pi}\Phi, \sigma^z \sim \sin\sqrt{2\pi}\Phi, \sigma^x \sim \sin\sqrt{2\pi}\Theta, \sigma^y \sim \cos\sqrt{2\pi}\Theta. \tag{5.27}$$

We see that Φ and Θ are like angle parameters for the system: they are periodic, so-called 'compactified'. Indeed, this is related to the non-abelian bosonization form of $SU(2)$-WZW g matrix

$$g = \varepsilon \mathbb{1} + i\sigma^a n^a \tag{5.28}$$

for spin operators σ^a.

The bosonized form of the XXZ model is

$$H = \int dx \Pi^2 + (1 + 4\Delta/\pi)(\partial_x \Phi)^2 + \Delta \cos\sqrt{8\pi}\Phi. \tag{5.29}$$

We see that the magnetization is renormalized by Δ, and there is an additional cosine term, also known as 'umklapp term'. The scaling dimension of this term is 2, so it is marginally relevant, which means the value of Δ matters. We can see from this form that, when $\Delta \gg 1$, the ground state has $\sqrt{8\pi}\Phi = \pi, 3\pi$. This corresponds to the dimer order as $\varepsilon = 0$, and AF order as $M(x) = \pm 1$, which means the phase is AF. Further, there is no Θ term in the model, and this means that the system is not ordered in the XY plane, i.e., the σ^x and σ^y sectors are random. However, it is not easy to see the features when $\Delta \approx 1$. At $\Delta = 1$, it is gapless, while it is gapped for $\Delta > 1$.

We can map the model to Majorana form. The cosine term is like $R^\dagger L - L^\dagger R$. Each fermion, R and L, can be written as two Majorana fermions η_R^1, η_R^2, and η_L^1, η_L^2. The two Majorana fermions have the same mass. We can see that the gap is the mass term of Majorana fermions.

Question 109. *What are the primary defects in XXZ model?*

We know that Ising model has two defects: the Majorana fermion and the Ising anyon. We also know that the XXZ model can be viewed as two coupled Ising models but with U(1) global symmetry. The Ising model has two phases: PM and AF, while the XXZ model has three phases: AF, FM, and XY. The PM and AF are dual of each other as disordered and ordered phases. There are two types of defects in the XXZ model, but whether they are Majorana fermions or Ising anyons are not clear.

5.3.6 Dimer models

We now study models with full SU(2) symmetry which generalize to higher-spin cases. Let's first see the so-called valence-bond solid (VBS), also known as Affleck-Kennedy-Lieb-Tasaki (AKLT) model for spin-1 chain

$$H = \sum_j \vec{S}_j \cdot \vec{S}_{j+1} + \frac{1}{3}(\vec{S}_j \cdot \vec{S}_{j+1})^2. \tag{5.30}$$

Figure 5.3: The valence-bond solid model. Dots are for virtual spins, with valence-bond between them. Circles are projections to physical spins.

How to read this model? The two terms commute. It turns out the first term is Heisenberg: it allows AF order like up-down, up-down etc, while the second term allows fluctuation: it allows zero-up, up-zero, zero-down, and down-zero, with zero playing the role of defect. The model is solvable: the ground state is the AKLT state, which is the 'father' of matrix-product states (MPS) with global $SO(3)$ symmetry, and excitations can be obtained from field theory, such as nonlinear sigma model or WZW model, or numerical simulations like DMRG.

This model has several notable features:

- The ground state has the so-called dilute AF order: it is a weighted superposition of dilute AF strings.
- The ground state has SPT order, so it has nontrivial edge states. The edge states live in the 'virtual' or 'bond' space.
- The Hamiltonian is a sum of projectors, and this construction generalizes to all MPS with small bond dimensions. Each on-site spin-1 can be viewed as a projection from products of two virtual spin-1/2, and each two-body Hamiltonian term is a projection onto spin-2. See Fig. 5.3.
- It can be used for quantum computing on a qubit to realize the group $SU(2)$.

In terms of MPS (Chapter 3), the on-site tensor are the three Pauli matrices σ^i. The dilute AF order can be seen in a rotated basis that specify the tensor as $(\sigma^+, \sigma^z, \sigma^-)$, for the physical spin states $(\uparrow, 0, \downarrow)$. From $(\sigma^\pm)^2 = 0$ and $[\sigma^+\sigma^-, Z] = 0$, we see that the ground state contains configurations like $\uparrow 000 \downarrow$, but not $\uparrow 0 \uparrow 0 \downarrow$, i.e., it is a dilute AF. The weight of each configuration is an inverse exponent of the number of $\uparrow\downarrow$ pairs.

With edges, there is a spin-1/2 d.o.f at each edge. This edge states are protected by the global symmetry. The SPT order follows from the fact that for $U \in SO(3)$ acting on the spin-1, there will be a $V \in SU(2)$ acting on the virtual spin-1/2, namely,

$$\sum_j u_{ij}\sigma^j = V\sigma^i V^\dagger. \tag{5.31}$$

As spin-1/2 is a 'projective' representation of $SO(3)$, the AKLT has nontrivial SPT order. Recall that half-integer spins and integer spins belong to two different congruence classes of representations of $SO(3)$. All half-integer spins are projective representations of $SO(3)$, and integer spins are linear representations of $SO(3)$. If the

virtual spins are integers, then they can be mapped to spin-zero case without changing the order parameter of the states, i.e., no phase transition, and the SPT order is trivial. This confirms the Haldane distinction between integer and half-integer spin Heisenberg models.

Question 110. *How powerful is SPT order?*

For quantum computing, we need to first couple one edge spin-1/2 to one additional spin-1/2 forming a singlet. Then there are at least two ways to use the state to realize quantum gates.

- One scheme is via local measurement, as we have discussed in Chapter 4. The gate sequence on AKLT state is a bit different from that on 1D cluster state, which is

$$|\text{out}\rangle = X(-\gamma)\sigma^{r_1 r_2} X(\gamma) Z(-\beta)\sigma^{t_1 t_2} Z(\beta) X(-\alpha)\sigma^{s_1 s_2} X(\alpha)|\text{in}\rangle. \tag{5.32}$$

  for $Z(\alpha) := e^{-i\alpha\sigma^z}$ and $X(\alpha) := e^{-i\alpha\sigma^x}$. The Pauli operators σ^i can be pulled out from the sequence by the $su(2)$ algebra. The input (output) state is carried by the left edge (additional spin-1/2).

- The other scheme is via global unitary operators, by treating the left edge as the logical space, while the rest as the encoded space. Then a global symmetry $G = \otimes_n U_n \otimes V$ for $U_n \in SO(3)$ acting on bulk site and $V \in SU(2)$ acting on the additional spin-1/2 coupled to the right edge will act as logical gate V, which is universal for $SU(2)$.

The above features can be generalized to other dimer or VBS models, and also to higher Lie group symmetry cases. We denote a VBS state as $|\Xi_{mn}\rangle$ for $m+n=2S$, and integers $m,n \geq 0$. The on-site spin-S is a projection from $2S$ spin-1/2, and there are m (n) bonds to the left (right) of this site. Alternatively, m bonds can be treated as a single bond of two spin-$m/2$. For PBC there are even number of sites.

Each $|\Xi_{mn}\rangle$ can be expressed as

$$|\Xi_{mn}\rangle = \sum_{i_1,\ldots,i_L} \text{tr}(A_{i_1} B_{i_2} \cdots A_{i_{L-1}} B_{i_L})|i_1 \cdots i_L\rangle \tag{5.33}$$

with two types of matrices A_i of size $(n+1)\times(m+1)$ and B_i of size $(m+1)\times(n+1)$. To take a certain states $|\Xi_{mn}\rangle$ as exact ground states, parent Hamiltonian can be constructed from projectors.

5.4 Topological phases

We have seen a lot from the Ising world. Now we explore what are the new things beyond the Ising world. There is a brand new area of topological orders or phases, with symmetry-protection or not. We will study TOP (topological) order, SPT (symmetry-protect topological) order, and SET (symmetry-enriched topological) order.

1. TOP order: SSB of q-form symmetry ($q > 0$). Ground state degeneracy depends on topology. Examples include toric code, double-semion model, superconductors. TOP order is described by TQFT for the bulk, CFT for the edge.

2. SPT order: ground state is unique, except edge states. No intrinsic anyonic excitations. The protection symmetry can be q-form. Examples include VBS models, Kitaev chain, 'topological' insulator, 'topological' superconductor. SPT order can be described by effective field theory, e.g., WZW model, sine-Gordon theory, nonlinear-sigma model. The edge state is described by CFT.

3. SET order: a combination of TOP and SPT order. Examples include fractional quantum Hall, quantum dimer models. The anyons are projective representations of the preserved symmetry. SET order is described by TQFT (with preserved symmetry) for the bulk, CFT for the edge.

The SSB of 0-form symmetry, i.e. global symmetry, is the traditional SSB order, which we will not discuss.

Question 111. *What do we care about a many-body system?*

As we studied for the Ising model, there are lots of things to care about, although the Ising model appears very simple.

There are too many models in many-body and condensed-matter physics, and there are excellent textbooks even just on one model. Here we select some of these models, and we study the issues listed below for a model:

- Hamiltonian; Ground states; Excitations; Dispersion; Quantum phase transition; Order parameters; Order (TOP or SPT); Symmetry; Thermal property; Transport; Field theory description; Edge states; Defects; Computation with it; Duality or gauging; Special cases; Related models.

We first make a quick remark on superconductor.

Question 112. *Does superconductor have topological order?*

The surprising answer is yes. Traditionally superconductor is treated as a seminal example of SSB order. However, it is realized that, in the presence of magnetic field, the traditional local order parameter (e.g., correlation function) is not gauge-invariant. The gauge symmetry here is U(1). With the U(1) gauge symmetry, we are forced to choose nonlocal order parameters, which turns out to be Wilson loops. Evidence of topological order include fractionalizations of charge, spin, and flux for excitations. There are spinon (quasiparticle), vortex (vison), and massive photon excitations. The spinon is neutral and of spin 1/2, while the underlying electrons are charged and form singlets or triplets. A vortex (vison) can carry half-quanta flux instead of integer

ones. The presence of gauge symmetry is essential, and the magnetic field is said to be 'dynamical', i.e., as intrinsic d.o.f. As such, type-I superconductor can still be viewed as an example of SSB order, described by BCS mean-field theory.

Conventional type-II 2D and 3D superconductor can be described by the so-called 'BF field' theory, which shows that there are quasiparticle and vortex excitations, which are self-boson but mutually semion. That is, when a quasiparticle moves around a vortex, it will get a π-phase shift. This is the key feature of Z_2 topological order, also shared by, e.g., toric code and quantum spin liquid. There are two sets of Pauli algebra of Wilson loops, and the GSD on a torus is 4. Therefore, two logical qubits can be encoded for the purpose of topological quantum computing.

Another very appealing feature is that there is a finite-temperature phase transition, i.e., there is a T_c. This is a strong evidence that superconductor might be a *self-correcting* quantum memory. The self-correction means that below T_c, the quantum state would not change by thermal noises, just like 2D Ising model that can encode a classical bit. However, the quantum gates that can be performed are limited, same as the set of gates on toric code, as we can see below. To promote the computational power of superconductor, we might need some sort of 'anyonic superconductors'.

5.4.1 Toric code

The Hamiltonian for the toric code of qubits on the square lattice is

$$H = -A\sum_v X_v - B\sum_p Z_p \tag{5.34}$$

for the vertex term $X_v = \prod_{i\in v}\sigma_i^x$ and plaquette term $Z_p = \prod_{i\in p}\sigma_i^z$, and σ_i^x, σ_i^z act on qubits living on edges of the lattice. Here v (p) labels the vertex (plaquette). $A, B > 0$.

The model is commuting and each term X_v and Z_p are called stabilizers. A ground state $|G\rangle$ is a stabilizer state with $X_v|G\rangle = |G\rangle$ and $Z_p|G\rangle = |G\rangle$. It can be viewed as a so-called 'string-net condensation'; i.e., it is an equal-weight superposition of loops of 1 in the background of 0. The loops do not have 'tension' in the sense that its size can be changed (by stabilizers) without causing energy. The model is gapped, so it is an 'insulator'.

On a torus there are four ground states since $\prod_p Z_p = \prod_v X_v = \mathbb{1}$. So it encodes two logical qubits. It is the most famous example of a topological stabilizer code. The GSD is 4^g for genus g, i.e., number of 'holes', so it encodes $2g$ qubits. Denote the two directions of a torus as x and y, then the logical Pauli operators are X_{ℓ_x}, Z_{ℓ_y}, and X_{ℓ_y}, Z_{ℓ_x}, for ℓ denoting nontrivial loops, and X_{ℓ_x} a loop of X operators along direction x etc. The GSD can be viewed as SSB of symmetry defined by X_{ℓ_x} and X_{ℓ_y}, which are known as 1-form symmetry.

There are two types of excitations: 'electric' charges e and 'magnetic' fluxes m. They all appear in pairs and can be separated far away from each other without causing energy. A charge-pair (flux-pair) is created by Z_i (X_i) on a site i. They

form a duality similar with the duality in Maxwell's equation: the electric field and magnetic field. Actually the original model is the Z_2 gauge theory, as a discrete model of Maxwell's equation without sources. The duality can be implemented by a global Hadamard operations, which is a weak symmetry of the model. The excitations are bosons, while the composite one $\psi = e \times m$ is nothing but the Majorana fermion we already know.

A generalization of the model is

$$H = H_0 - h_x \sum_i \sigma_i^x - h_z \sum_i \sigma_i^z, \tag{5.35}$$

for H_0 as the bare model and other terms as external fields. We can expect that when h_x or h_z is big enough, the model will be driven to a trivial PM phase in x-direction or z-direction. This is indeed the case. The transition to the x-PM phase (z-PM phase) is flux (charge) condensation. The transition is topological since there is no symmetry breaking. The original Z_2 gauge model treats the vertex terms X_v as 'gauge symmetry', and do not contain h_z term. While one can also treat plaquette terms Z_p as gauge symmetry. The toric code is a topological order, or 'deconfined phase', according to the cohomology classification

$$\mathscr{H}^3(Z_2, U(1)) = Z_2, \tag{5.36}$$

for the first Z_2 as the gauge symmetry, and the second as the classification. The toric code is the nontrivial one with edges states, while PM is the trivial 'confined phase' with no edge states. A confined phase can be viewed as a phase whose ground state is only a particular configuration of the ground state for the deconfined or topological phase. Excitations become confined when their creation need energy, and they will create other configurations in the topological ground state.

Observable are gauge-invariant and are known as Wilson loops, W_x and W_z. In the deconfined phase, on ground states the values for W_x and W_z are nontrivial, while in the confined PM phases, they can be trivial. At finite temperature, the thermal value of W_x and W_z obey the so-called perimeter law in the deconfined phase, and the area law in the confined phase, for perimeter and area of the loops. The reason is that the loop W_z (W_x) has tensions in the x-PM (z-PM) phase, while they do not in the topological phase.

The Z_2 topological order can be described by the $U(1)^2$ Chern-Simons gauge theory, with Lagrangian

$$\mathscr{L} = \int d^3x \, \varepsilon^{\mu\nu\rho} a_\mu \partial_\nu b_\rho, \tag{5.37}$$

for a and b as the gauge fields corresponding to the electric and magnetic charges. Wilson loops are of forms $e^{i\int_\ell a(b)_\mu dx_\mu}$. The edge physics is described by Ising CFT. Another field theory which is equivalent to the Chern-Simons theory in this case is the so-called BF field theory, which makes clear that there are two species of excitations. The BF field theory is also more suitable to describe 3D toric code, for which the

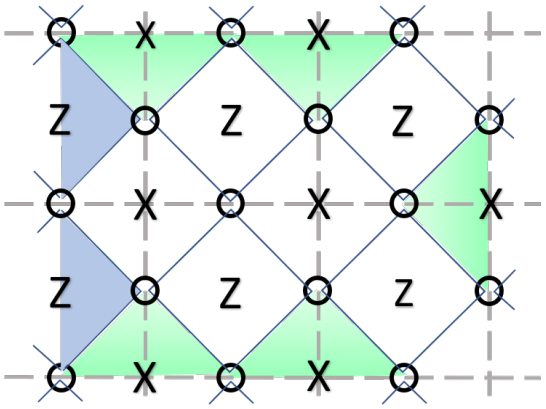

Figure 5.4: Toric code with edges. Bulk have weight-four stabilizers, while edges have weight-three stabilizers. Circles are qubits.

magnetic excitations become line-like instead of point-like. We will introduce the BF field theory for topological insulators, which can also be described by it.

Question 113. *What are the edges of toric code?*

The edge physics is interesting. A ground state can be written as a tensor-network state, with tensors as a generalization of the GHZ state. This implies that the edges states may be a GHZ state, and this is true. It turns out there are two types of edges: one as the edge that can condense fluxes, and the other for charges. The former is known as 'smooth' edge, and the later as 'rough' edge. See Fig. 5.4. On the edges the stabilizers need to be modified. It is not hard to see that we shall delete one local X or Z operators from the corresponding stabilizers, i.e., the stabilizers for edges are of weight three. A smooth edge can be described by the 'Ising' model $H = -\sum_n X_{n-1}X_nX_{n+1}$, and the global symmetry is $\otimes'_n Z_n$, which only acts on every other sites. A rough edge can be described by the 'Ising' model $H = -\sum_n Z_{n-1}Z_nZ_{n+1}$, and the global symmetry is $\otimes'_n X_n$, which also only acts on every other sites. We can see the smooth (rough) edge supports charge (flux) excitations while condense fluxes (charges). The symmetry operators play the roles of logical gates.

What sits at the interface between the two types of edges? That degree of freedom is called a 'defect'. We already know that the duality defect in Ising model is the Ising anyon, which is non-abelian.

What happens when two types of edges are close to each other, i.e., fused together? The fused edge will be described by

$$H = -\sum_n X_{n-1}X_nX_{n+1} + Z_nZ_{n+1}Z_{n+2}, \tag{5.38}$$

which is nothing but the 1D cluster model, as the gauged Ising model. This is reasonable since the two edges are dual of each other, and a gauged model is a composite of the system and its dual.

Question 114. *How to use the toric code for quantum computing?*

It turns out there are many ways.

- We can use the whole system as a single code.
- We can use holes, defects, or edges on the system to encode qubits.
- If we can perform local projective measurements on each site, then we can use the code as resource for the teleportation-based model.

The three schemes above are consistent in the sense that they can only do Clifford gates naturally. The limitation of toric code, and abelian topological order in general, comes from the fusion and braiding structures of anyons.

One way to actually move the defects is by local measurements. The two defects sit at the end of an edge. The three-weight stabilizer terms at the end of an edge can be viewed as a hole, i.e., the absence of the bulk four-weight stabilizer terms. As there are two types of edges, we can have two types of holes: X-cut and Z-cut holes. Two holes encode one qubit. We can use one additional qubit as an ancilla for each vertex and plaquette term to record the parity of it. This needs entangling gates between the ancilla and qubits around it. To braid holes, we can use measurement to first enlarge a hole until it encircles another hole, and then bring it back to the original position. Braiding of X-cut and Z-cut holes will induce an entangling gates on them: the CX gate.

When the Z_2 topological order is realized in some physical systems like quantum spin liquids, local measurements cannot be easily done. It is a great technical challenge to create anyons (or defects), locate them, and braid them.

5.4.2 Fractional quantum Hall effect

Hall effects are a family of phenomena of electrons in the presence of strong magnetic effects and low temperatures. Here we discuss the famous fractional quantum Hall (FQH) effect, which was discovered by the observation of fractional quantized plateaus of the Hall conductance. It requires very strong external magnetic field, which is not an easy task. As such, other effects like spin Hall or anomalous Hall aim to substitute the field by a certain intrinsic processes like spin-orbit coupling.

FQH effect generalizes the integer quantum Hall (IQH) effect by electron-electron interactions. The IQH effect favors disorder, while FQH effect favors pure samples since electrons have to interact with each other. In IQH effect, the resistivity ρ_{xy} shows plateaux, while ρ_{xx} shows peaks at the edges of plateaux. At each plateau, labelled by integer $\nu \in \mathbb{Z}$, the magnetic field $B_\nu \propto 1/\nu$, $\rho_{xy} \propto 1/\nu$. This has to be explained by the filling of *Landau levels*. The Landau model here describes free electrons in magnetic field and electric field. It turns out electrons undergo circular motions, and behave just like harmonic oscillators. The gap value of the harmonic

oscillator $\Delta \propto B$, so the role of B is essentially to tune the value of gap. When B increases, Δ gets bigger and fewer levels can be filled by electrons. In other words, increasing B is like cooling the electrons to the lowest Landau level, i.e., the ground states. The FQH effect is similar except that the value ν, the *filling factor*, can be fractional numbers.

Question 115. *Is FQH effect fully understood?*

By far there are still controversy on the microscopic details of FQH effect, such as the Hamiltonian and systematically determine and explain the origin of ν. In Hall family, Hamiltonian is not usually mentioned since it does not play much role for transport phenomena. In general, the Hamiltonian, such as the Trugman-Kivelson model, describes a collection of electrons with Coulomb interaction under external fields. It is difficult to exactly solve such models.

As a generalization of the single-electron wave function for the lowest Landau level, the Laughlin states for N interacting electrons

$$\psi(m) \propto \prod_{i<j}(z_i - z_j)^m e^{-\sum_{i=1}^N |z_i|^2/4\ell_B^2} \tag{5.39}$$

describes the universal features of FQH at $\nu = 1/m$, $m \in 2\mathbb{Z}+1$ for fermions, for ℓ_B as the so-called magnetic length. Physically, m is the relative angular momentum between electrons i and j, and also measures the distance between electrons. We know that Coulomb interaction tends to repulse electrons, and when quantized, electron behaves like a localized wavepacket or 'droplet' of a certain size m. The exponent factor above is viewed as a uniform charged background, the 'center of charge' of the electrons. If we view the 2D sample as a lattice, then electrons will fill in these sites, but we do not know which site an electron sits. Instead, the ground state has topological order: it is a superposition of all possible configurations of filling electrons to the space.

The FQH system can be described by the $U(1)^n$ Chern-Simons gauge theory with Lagrangian

$$\mathscr{L} = \int d^3x \, K_{IJ} \, \varepsilon^{\mu\nu\rho} a^I_\mu \partial_\nu a^J_\rho, \tag{5.40}$$

for n species of gauge fields a^I, which are abelian anyons. Its edge state is chiral Luttinger liquid described by

$$\mathscr{L} = \int d^3x K_{IJ} \partial_x \phi_I \partial_t \phi_J - V_{IJ} \partial_x \phi_I \partial_x \phi_J, \tag{5.41}$$

with potential nonlinear terms like $\cos \phi_I$ that can induce gaps. Exponents of chiral bosons ϕ_I are vertex operators, following from bosonization theory. An edge is chiral since electrons can only propagate in one direction.

Question 116. *What is the order of FQH state?*

The FQH system has topological order. For abelian case, the GSD is ν^g for genus g. For non-abelian case, the GSD can also be determined but is a bit technical. For $\nu = 1/3$ on a torus, the GSD can encode a qutrit. Different from the toric code, here the Wilson loops W_x and W_y of the same anyon do not commute, instead,

$$W_x W_y = e^{i2\pi/3} W_y W_x, \tag{5.42}$$

which behave as the qutrit Pauli operators X and Z (see Chapter 2).

A FQH ground state can also be written as MPS form. As the x-direction is along the electric field, the k_x is a good quantum number. The system is effectively like a 1D system along the y-direction. We take the x-direction as periodic with radius r, then along the y-direction there are a series of 'orbits'. Then the FQH state can be viewed as a 1D chain of orbits. As we know the edges are described by a CFT, i.e., chiral Luttinger liquid, then the translation-invariant tensors in the MPS act on the CFT space. The tensor A^m in a local occupation basis $|m\rangle$ contains the vertex operator V^m (and other less trivial part) for $V = \frac{1}{L}\oint V(z)dz$, $V(z) = e^{i\phi(z)/\sqrt{\nu}}$ for Laughlin state with $\nu = p/q$, for $L = 2\pi r$ as the perimeter of the periodic direction. The vertex operator V will be more complicated for other states; e.g., for Moore-Read states, there is a chiral Majorana field part.

The system have two global symmetry: charge conservation and momentum conservation. The net particle number is $C = \sum_j^{L_y} qN_j - p$, momentum is $K = \sum_j^{L_y} j(qN_j - p)$ for j as the orbital index along the y-direction, N_j as particle numbers on that orbit. C and K commute, and both shall commute with the Hamiltonian. The exponent $T_y = e^{i\frac{2\pi}{L_y}K}$ acts as a twist operator, also known as 'large gauge transformation', and is physically a flux insertion operation. On a torus, the y-direction is also periodic, and the flux T_y is equivalent to W_x, and hence does not commute with W_y. Equivalently, a flux T_x is equivalent to W_y. The two flux operators T_x and T_y hence do not commute, which can also be used to show the GSD.

Due to the preserved global U(1) charge symmetry, the FQH state has SET order. The abelian anyons are fractional (projective) representations of the U(1) symmetry.

In addition, we mention that the SET features of FQH states are similar with 2D quantum dimer models, which also have abelian topological order yet with global $SU(2)$ symmetry. The 2D quantum dimer models are extensions of the toric code, e.g., there are spinon and vison as excitations. Due to the global symmetry, we can use twist operators to partially substitute Wilson loops to play the role of logical operators.

For FQH with other values of ν, such as 5/2, there could be non-abelian anyons, which are of interest for anyonic quantum computing. We have discuss this topic in Chapter 4. Although anyons have not been definitely confirmed in FQH systems, there is always a hope to achieve this and also build quantum computing hardware with them.

What would be the competing phase other than the FQH topological phase? Similar with the toric code, there shall be confined phases. A such phase is known as

'Wigner crystal', for which the electrons are ordered with certain positions. This phase can be achieved under the so-called thin-torus limit, which effectively reduces the 2D system to a 1D system. A Wigner crystal state can be viewed as a particular configuration in the topological FQH state, which is a superposition of various configurations as a condensation, similar with the case of toric code.

5.4.3 Topological insulator

Topological insulator (TI) has conducting edge states but the bulk is still insulating. TI has SPT order protected by time-reversal symmetry, which leads to the edge states, and it does not have TOP order so there is no anyonic excitations in the bulk.

TI has intriguing relations with superconductors and quantum Hall effects. On one hand, TI can be viewed as two copies of integer quantum Hall states with opposite chirality, and on the other hand, TI and superconductors can both be described by the BF theory, which is a topological field theory.

With spin-orbit coupling, 2D TI is also called spin-Hall states since a charge current can induce a spin current in the orthogonal direction. No external magnetic field is required. On the contrary, there shall be no magnetic field since it will break the time-reversal symmetry, which is crucial for the SPT order of TI. The edge state contains two 'channels': a spin-up one and a spin-down one, and the electrons of them move in opposite directions. This is known as spin-momentum locking or helical states.

Question 117. *A similarity with toric code?*

You may notice there is a similarity with the edge states of toric code. Indeed this is the case. For toric code, there are two types of edge states. The edge states can be fused together to form a composite edge, and defects between the edges state behave like Ising anyon. The number of edge states is the same as the number of species of excitations, and for them it is two.

The similarity arises from the BF field theory, which can describe both of them. However, TI only has SPT order, so the TOP order is trivial: there is no topological GSD. Excitations of TI are electrons, which are not fractionalized, hence not anyon, neither. This is analog with the Chern-Simons description of integer quantum Hall states, which also have trivial TOP order. However, there may exist fractional TI at least by theory, as the analog of FQH states.

The model of 2D BF theory is defined by

$$\mathscr{L} = \frac{k}{2\pi}\varepsilon^{\mu\nu\lambda} a_\mu \partial_\nu b_\lambda + \frac{1}{\pi}\varepsilon^{\mu\nu\lambda} A_\mu \partial_\nu b_\lambda + \frac{1}{\pi}\varepsilon^{\mu\nu\lambda} B_\mu \partial_\nu a_\lambda . \tag{5.43}$$

The first term is the so-called 'BF' term, which encodes the statistics of the gauge fields a_μ and b_λ. For 3D case, we shall replace b_λ by $b_{\lambda\rho}$, and the pre-factor of the BF term becomes $\frac{k}{4\pi}$. Here $k = 2$ for TI, and it determines the mutual statistics of a_μ

and b_λ: when a_μ and b_λ braid, the Berry phase is $2\pi/k = \pi$, which means they are mutual-semion yet self-bosons.

The fields a_μ and b_λ, however, are fictitious 'partons' of electron: a_μ is spinon (with spin 1/2) with no charge, while b_λ is holon (with charge e) but no spin. Note in TI there is no actual spin-charge separation. The A_μ and B_μ are external fields acting as sources for holon and spinon, respectively.

Question 118. *What does TI bring to electrodynamics?*

The BF field theory describes the excitations, edge states, and also the bulk response to external fields, leading to the so-called 'axion electrodynamics'. The axion terms are direct couplings between the electric field and the magnetic field. TI acts like 'axion media', a hint for which is the interplay between spin current and (charge) current, i.e., the spin-Hall effect. By tracing out the bulk excitations, the BF model reduces to

$$\mathscr{L}_{2D} = \frac{1}{\pi}\varepsilon^{\mu\nu\lambda}A_\mu\partial_\nu B_\lambda,\ \mathscr{L}_{3D} = \frac{1}{8\pi}\varepsilon^{\mu\nu\lambda\rho}\partial_\mu A_\nu\partial_\lambda A_\rho, \tag{5.44}$$

which describe the response bulk fields. The axion term of 3D TI is of the form $\theta\vec{E}\cdot\vec{B}$, and furthermore, the value of $\theta = \pi$ is of topological nature.