

**Physics at the Perimeter Institute:
Topics in Theoretical Physics
&
Quantum Simulation Research**

- A Quick Guide -

Huan Q. Bui

Colby College

PHYSICS & MATHEMATICS
Statistics

Class of 2021

July 8, 2020

Preface

Greetings,

This guide is my notes from Perimeter Institute of Theoretical Physics Summer School, 2020. Topics include quantum information, thermodynamics, numerical methods, condensed matter physics, path integrals, and symmetries.

The text also includes reading notes for my research project at Perimeter on quantum simulation.

Enjoy!

Contents

Preface	2
1 Quantum Information & Thermodynamics	7
1.1 Introduction	8
1.2 Basic facts of Quantum Mechanics	9
1.3 Foundations of Quantum StatMech: Density Operator Formulation	12
1.4 Entanglement	16
1.5 Time evolution	18
1.6 Quantum Equilibration and Thermalization	20
1.6.1 Equilibration	20
1.6.2 Thermalization: Turtles All The Way Down!	21
1.7 Equilibration	27
1.7.1 Equilibration	27
1.8 Resource Theories and Quantum Information	30
1.9 Quantum Thermal Operations	31
1.10 Fluctuation Theorems and Quantum Information	32
1.11 Quantum Thermal Machines	33
1.12 Exercises	34
2 Numerical Methods & Condensed Matter Physics	45
2.1 Introduction	45
2.2 The Ising model: Motivation	47
2.3 Introduction to many-particle states and operators	47
2.3.1 Spin 1/2 Systems	47
2.3.2 Operators on Spin-1/2 Systems	48
2.4 Tensor Product Spaces	49
2.4.1 Born Rule for Probabilities	50
2.4.2 Expectation Values	50
2.5 The Two-site Ising Model	51
2.6 The N -site Ising Model:	54
2.7 Matrix product states	56

3 Path Integrals	57
3.1 Introduction to path integrals and the semi-classical limit	58
3.1.1 Motivation & Definition	58
3.1.2 Some properties	59
3.2 Propagator in real and imaginary time	61
3.2.1 Free Path Integrals	63
3.3 Perturbation theory	64
3.4 Non-perturbative physics and quantum tunneling	65
3.5 Topology and Particle Statistics	68
3.6 Applications	70
3.7 Exercises	71
4 Symmetries	83
4.1 Review: Equivalence Classes	84
4.1.1 Equivalence relations	84
4.1.2 Equivalence Classes	84
4.2 Definition of symmetry	87
4.3 Elements of group theory	87
4.4 Examples: Symmetry & Group Theory	89
4.5 Continuous and discrete symmetries	90
4.5.1 Discrete Symmetries	90
4.5.2 Continuous Symmetries	90
4.6 Infinitesimal symmetries	91
4.7 Noether's theorem	92
4.7.1 Noether's Theorem for fields in flat space (quais-symmetry) .	95
4.7.2 Noether's Theorem: Fields in Flat Space	95
4.8 Lie groups and Lie algebras	96
4.8.1 Matrix Exponential	96
4.8.2 Matrix Lie groups and associated Lie algebras	96
4.9 Representation theory	99
4.10 Symmetries in quantum mechanics	101
4.11 Exercises	104
5 Research Project: Quantum Simulation	109
5.1 Review: The density operator	110
5.1.1 Ensembles of Quantum States	110
5.1.2 General Properties of the Density Operator	111
5.1.3 The Reduced Density Operator	113
5.2 Quantum Computational Complexity: An Introduction	114
5.2.1 Introduction	114
5.3 Quantum Simulation: A Primer	115
5.3.1 Simulation in action	115
5.3.2 The quantum simulation algorithm	116
5.3.3 The algorithm	119
5.3.4 An illustrative example	119
5.3.5 Perspectives on quantum simulation	121

5.4	Problems in Quantum Simulations and Quantum Algorithms	122
5.4.1	Quantum Simulation: An Overview	122
5.4.2	Simulation challenges in molecular and materials science	124
5.4.3	Challenges for quantum algorithms in quantum simulation	126
5.5	Quantum Phase Estimation (QPE)	134
5.5.1	Performance and Requirements	136
5.5.2	The algorithm	138
5.6	Quantum Variational Eigensolver (QVE)	140
5.6.1	Quantum Expectation Estimation (QEE)	140
5.6.2	Quantum Variational Eigensolver (QVE)	143
5.7	Some Models in Quantum Statistical & Condensed-Matter Physics	145
5.7.1	The Ising Model	145
5.7.2	The Hubbard Model	149
5.8	Locality in Quantum Systems	150
5.8.1	Introduction & Notation	150
5.8.2	Locality & The Lieb-Robinson bounds	151
5.9	Quantum Approximate Optimization Algorithm (QAOA)	153
5.10	Quantum Variational Factorization (QVF)	154
5.11	Variational Thermal Quantum Simulation via Thermofield Double States	155
5.11.1	Thermofield Double States (TFD)	155
5.12	Generation of TFD and Critical Ground States with a Quantum Computer	156
5.13	Efficient variational simulation of non-trivial quantum states	157
5.13.1	Introduction	157
5.13.2	Non-trivial quantum states	158
5.13.3	Variational Quantum-Classical Simulation (VQCS)	159
5.13.4	Using VQCS to prepare nontrivial quantum states	161
5.14	Cluster States & Measurement-based QC	166
5.14.1	How it works	166
5.14.2	Simulating quantum circuits in the cluster-state model	168
5.14.3	Properties	174
5.15	Mesurement-based quantum computing: Jozsa	178
5.15.1	Review	178
5.15.2	Adaptive Measurements	179
5.15.3	Cluster-state computation/1WQC	181
5.15.4	Selected extra features: Cluster state model	187
5.16	QAOA with Cluster States	189
5.16.1	Universal gates with CC	189
5.16.2	Cluster States for TFIM state preparation	189
5.17	Variational Imaginary Time Ansatz (VITA)	190
5.17.1	Introduction	190
5.17.2	The Ansatz	190
5.17.3	VITA on TFIM	191
5.18	The Jordan-Wigner Transformation	192
5.18.1	Fermonic canonical commutation relations	192

5.18.2 Jordan-Wigner transformation	201
5.18.3 Summary: The Jordan-Wigner Transformation	203
5.18.4 Jordan-Wigner transformation in QAOA	205
5.19 Quantum Critical State from GHZ?	206

Part 1

Quantum Information & Thermodynamics

Instructor: Alioscia Hamma

The aim of this course is to understand the thermodynamics of quantum systems and in the process to learn some fundamental tools in Quantum Information. We will focus on the topics of foundations of quantum statistical mechanics, resource theories, entanglement, fluctuation theorems, and quantum machines.

1.1 Introduction

Let's unpack the following three terms: Thermodynamics, Information, and Quantum.

Thermodynamics: Thermodynamics is not concerned so much with how I make a more efficient engine as the most important question: "What is possible?" Any theory (relativity, string, quantum field, etc) must be thermodynamically valid. This way of think makes thermodynamics appear quite fundamental – which indeed it is.

Information: Information theory was so abstract that it was viewed as a mathematical subfield for the longest time. But because of the information revolution, information theory becomes physical. Physics tells us what kind of information theory is allowed. Information theory becomes a new way to understand Thermodynamics and what's possible.

Quantum information theory provides even newer insights.

Now, quantum theory as it is has unanswered questions. Consider the Schrodinger cat, which is in a superposition of begin dead and alive. We know that for spins, $(|0\rangle + |1\rangle)/\sqrt{2}$ exists (coherently), but we don't see cat states. We believe that cat states are possible to make, but we also believe it's hard to create. However, we don't know for sure that such a Schrodinger cat state is possible.

With this, let us start.

1.2 Basic facts of Quantum Mechanics

Some mathematical basics:

Definition 1.2.1 (f.d. Hilbert spaces). Consider finite-dimensional Hilbert spaces $\mathcal{H} \sim \mathbb{C}^d$. This is just a vector space over the complex numbers. We can also write $\mathcal{H} = \text{span}\{|1\rangle, \dots, |d\rangle\}$. We can also call this a d -level system. We will use the Dirac notation $|\psi\rangle$ to denote elements of this Hilbert space.

Definition 1.2.2 (Linear Operators). Linear operators are mappings linear $A : \mathcal{H} \rightarrow \mathcal{H}$ defined by

$$A = \sum_{ij} A_{ij} |i\rangle \langle j| \quad (1.1)$$

where the set $\{|i\rangle\}$ form the basis of the Hilbert space, and A_{ij} 's are the matrix elements of A .

Definition 1.2.3 (Positive Operators). An operator A is positive if $\langle \psi | A | \psi \rangle \geq 0$ for all $|\psi\rangle \in \mathcal{H}$.

Definition 1.2.4 (Hermitian Operators). A is Hermitian if $A^\dagger = A$.

Theorem 1.2.1 (Spectral Theorem). A Hermitian operator A has a spectral resolution $|A_i\rangle$ which forms an orthonormal basis of \mathcal{H} , with which we can write

$$A = \sum_i a_i |A_i\rangle \langle A_i| = \sum_i a_i P_i^A \quad (1.2)$$

Objects such as $|A_i\rangle \langle A_i|$ are projectors.

Definition 1.2.5 (Schatten p -norm). The p -norm of an operator A is defined to be

$$\|A\|_p \equiv \left(\text{tr} \left[\{A^\dagger A\}^{p/2} \right] \right)^{1/p}. \quad (1.3)$$

Example 1.2.1. For $p = 1$, we get the trace norm:

$$\|A\|_1 = \text{tr} \sqrt{A^\dagger A}. \quad (1.4)$$

For $p = 2$, we get the usual Frobenius norm:

$$\|A\|_2 = \sqrt{\text{tr}(A^\dagger A)}. \quad (1.5)$$

For $p = \infty$, we get the operator norm:

$$\|A\|_\infty = \sup_{|\psi\rangle} \frac{\|A|\psi\rangle\|}{\|\psi\|}. \quad (1.6)$$

In fact, there are multiple equivalent definitions of the operator norms. (Consider functional analysis for more details).

In quantum mechanics, we have states and observables. States are Hermitian positive operators: $\psi \geq 0$ such that $\text{tr}(\psi) = 1$, i.e., $\psi \in \mathcal{B}(\mathcal{H})$ with unity trace ($\mathcal{B}(\mathcal{H})$ denotes the set of bounded operators). Note that in this formulation, states are operators, not vectors, and these represent physical systems. To look at measurements, consider the spectral resolution of an operator A :

$$A = \sum_i a_i P_i^A. \quad (1.7)$$

The eigenvalues of A are the possible outcomes of the measurement. We note that this is just the density operator/density matrix formulation of quantum mechanics. There's nothing really fancy going on here.

QM is a probabilistic theory

$$\Pr(a_i) = \text{tr}(\psi P_i^A) = p_i. \quad (1.8)$$

With this,

$$E[A] = \langle A \rangle = \sum a_i p_i, \quad \langle A^2 \rangle = \sum a_i^2 p_i, \quad \dots \quad (1.9)$$

Recall that

$$\text{Var}A = E[A^2] - E[A]^2 \geq 0. \quad (1.10)$$

Consider a state ψ , which is positive Hermitian operator with unity trace, i.e., $\psi \in \mathcal{B}(\mathcal{H})$. By the spectral theorem, we can write

$$\psi = \sum_i p_i |\psi_i\rangle \langle \psi_i|. \quad (1.11)$$

Since $\psi \geq 0$, we have that $p_i \geq 0$. To show this, suppose we want to compute p_i :

$$p_i = \langle \psi_i | \psi | \psi_i \rangle \geq 0, \quad (1.12)$$

so we're done.

Next, we also said that the trace of ψ is 1:

$$\text{tr}(\psi) = 1 \implies \sum_i p_i = 1. \quad (1.13)$$

In other words, we have a probability distribution.

Are we're used to the fact that states are normalized vectors: $\langle \psi | \psi \rangle = 1$, and that $p_i = |\langle \psi | A_i \rangle|^2$, and that $E[A] = \langle \psi | A | \psi \rangle$. Another thing, image that

we're given some a set of orthonormal states $|\psi_l\rangle$ with some probability q_l . Now we're asked $E[A]$ with $\{|\psi_l\rangle, p_l\}$. Well,

$$E[A] = \sum_l q_l \langle \psi_l | A | \psi_l \rangle. \quad (1.14)$$

This is what we're used to. Now, if we look at states as spectral resolutions (as operators):

$$\psi = \sum_l q_l |\psi_l\rangle \langle \psi_l|. \quad (1.15)$$

Then we will show as an exercise that

$$E[A] = \text{Tr}(\psi |A_l\rangle \langle A_l|) = \sum_l q_l \langle \psi_l | A | \psi_l \rangle. \quad (1.16)$$

So the operator formulation gives the same answer.

Next, consider the states $|\psi_l\rangle$ which give a probability distribution q_l . This produces a state/operator $\psi = \sum q_l |\psi_l\rangle \langle \psi_l|$. If this probability distribution is extremal, i.e., $q_l = (1, 0, \dots, 0)$, then

$$\psi = |\psi_l\rangle \langle \psi_l|. \quad (1.17)$$

This is what we call a **pure state**. Whenever we don't have an extremal probability distribution, we don't have a pure state. Suppose $\mathcal{H} = \text{span}\{|0\rangle, |1\rangle\}$. Then $\psi = (|0\rangle \langle 0| + |1\rangle \langle 0|)/\sqrt{2}$ is not a pure state.

Question: Is there a way to quantify whether a state is pure (or not)? Consider

$$\psi = \sum_i p_i |\psi_i\rangle \langle \psi_i|. \quad (1.18)$$

We know that $\sum p_i = 1$. So, $\sum_i p_i^\alpha \leq 1$ with $\alpha \geq 1$. Equality occurs iff the state is pure. We call

$$\frac{1}{d} \leq \sum_i p_i^2 \leq 1 \quad (1.19)$$

the **purity** of the state (when $\alpha = 2$).

Definition 1.2.6 (Composite Systems). Consider two copies of $\text{span}\{|0\rangle, |1\rangle\} = \mathbb{C}^2$. The “total” vector space of this composite system is given by $\mathcal{H} = \mathbb{C}^{\otimes 2}$.

In general, the composite vector space is just the tensor product of the vector spaces of the subsystems.

1.3 Foundations of Quantum StatMech: Density Operator Formulation

The next question is Entropy. Consider this thought experiment by Von Neumann. Consider N particles in $\mathbb{C}^{d_i} = \mathcal{H}_i$. Then $\mathcal{H} = \bigotimes^N \mathcal{H}_i$. For simplicity, we can just look at a system of N 2-level systems/qubits. Consider the state of each of these particles:

$$|\psi_i\rangle = \sum_k w_k^i |e_k\rangle \in \mathcal{H}_i \quad (1.20)$$

Think of these particles in a box. Let's put them in a box, all of them in the same state. For example, let's assume that $|\psi_i\rangle = (1/\sqrt{2})(|+\rangle + |-\rangle)$. Then

$$|\psi\rangle = \bigotimes_i |\psi_i\rangle. \quad (1.21)$$

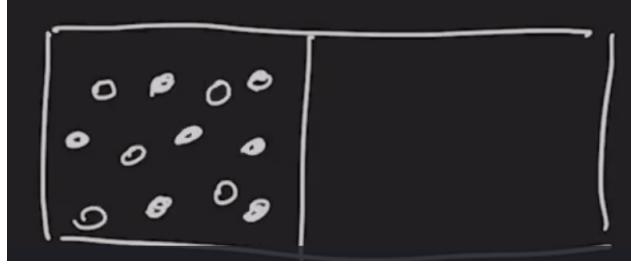


Figure 1.1: From Alioscia Hamma's lecture

I call that $S(|\psi\rangle) = 0$. We call this a quantum gas of non-interacting particles with zero entropy. We want to measure whether the state is in $|+\rangle\langle+|$ or $|-\rangle\langle-|$. Well, some of them will be $+$, some will be $-$. well,

$$n_+ = N|w_+|^2 = \frac{N}{2} = n_- \quad (1.22)$$

After the measurement, the state will be

$$\psi = |w_+|P_+ + |w_-|P_- = |w_+||+\rangle\langle+| + |w_-||-\rangle\langle-|. \quad (1.23)$$

How did the entropy change because of the measurement? Suppose we have a box like the picture on the right. It has a movable wall on the left and a fixed but semi-transparent wall in the middle. We also have another wall in the middle that can be moved. What Von Neumann thought was that what if we moved both the movable walls so that the volume containing the particles. With the semi-transparent wall, we are left with the picture on the left.

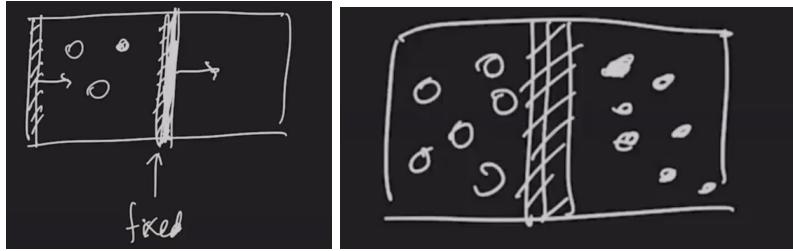


Figure 1.2: From Alioscia Hamma's lecture

Note that nothing happened to the gas, so $Q, W = 0$, and everything is reversible. At this point, we have separated the gases, but have not communicated with the environment yet. Now, we do a *isothermal compression*. This is the environment is interacting with our gas. Now, the + gas will be compressed $V \mapsto V'_+ = |w_+|^2 V$, and the - : $V \mapsto |w_-|^2 V$. In this compression, the gases exchange work and heat with the environment. What is the entropy exchanged?

$$\Delta S^{\text{env}} = \Delta S_+^{\text{env}} + \Delta S_-^{\text{env}} = -n_+ k \log \frac{V'_+}{V} - n_- k \log \frac{V'_-}{V}. \quad (1.24)$$

Now remember that $n_{\pm} = |w_{\pm}|^2 N$ and $V_{\pm} = |w_{\pm}|^2 V$, so

$$\Delta S^{\text{env}} = -Nk|w_+|^2 \log |w_+|^2 - Nk|w_-|^2 \log |w_-|^2. \quad (1.25)$$

Now, I do a unitary rotation, $|+\rangle\langle+| \mapsto w_+|+\rangle + w_-|- \rangle$. This is reversible. Once we've done it for all particles, we're exactly in the initial state. Now, because everything we've done so far is reversible, we have that

$$\Delta S^{\text{env}} + \Delta S^{\text{gas}} = 0 \implies \Delta S^{\text{env}} = -\Delta S^{\text{gas}} = -(S_1 - S_2) = -S_2. \quad (1.26)$$

So,

$$S_2 = Nk|w_+|^2 \log |w_+|^2 + Nk|w_-|^2 \log |w_-|^2. \quad (1.27)$$

Definition 1.3.1 (Von Neumann Entropy). The entropy of a state ψ is given by

$$S(\psi) = -Nk \sum_i |w_i|^2 \log |w_i|^2 \quad (1.28)$$

Proposition 1.3.1.

$$S(1/d) = \log d \quad (1.29)$$

$$S_{\alpha} = \frac{1}{1-\alpha} \log \text{Tr}(\psi^{\alpha}). \quad (1.30)$$

Let's look at a different definition of Von Neumann entropy, given in Mike and Ike:

Definition 1.3.2 (Von Neumann Entropy). The entropy of a quantum state ρ is given by the formula

$$S(\rho) = -\text{tr}(\rho \log \rho). \quad (1.31)$$

In this formula, the logarithm is take to base two. If λ_x are the eigenvalues of ρ then Von Neumann's entropy definition can be re-written as

$$S(\rho) = -\sum_x \lambda_x \log \lambda_x, \quad (1.32)$$

where we define $0 \log 0 = 0$. For calculations, it is usually this formula that is most useful.

For example, the completely mixed density operator in a d -dimensional space, $\mathbf{1}/d$, has entropy $\log d$, like above.

Just for clarity, let's look at the definition of pure states. Pure states are states of the form $\rho = |\psi\rangle\langle\psi|$. **Statistical mixtures of pure states form mixed states** of the form $\sum q_j |\psi_j\rangle\langle\psi_j|$. By definition, pure states are projectors, and thus they are idempotents, i.e., $\rho = \rho^2 = \rho^n$ for any $n \geq 2$.

So, just the recap, we know that quantum mechanics (under the density operator formulation) has **states**, which are $\psi \geq 0, \text{tr } \psi = 1$, and observables $A = \sum_{ij} A_{ij} |i\rangle\langle j|$. Next, all Hermitians \mathcal{H} are diagonalizable and the spectral theorem says $\mathcal{H} = \sum_k h_k |\mathcal{H}_k\rangle\langle\mathcal{H}_k|$ where $|\mathcal{H}_k\rangle\langle\mathcal{H}_k|$ is a projector and h_k 's are the eigenvalues, which are also the possible outcomes. Further, $\psi = \sum_{ij} \psi_{ij} |i\rangle\langle j| = \sum_k p_k |e_k\rangle\langle e_k|$. The p_k 's here are the eigenvalues of ψ . And notice that because $\text{tr } \psi = \sum_k p_k \langle e_k | e_k \rangle = 1$. Thus we have a probability distribution.

Definition 1.3.3 (Pure states). A state ψ is pure if and only if the probability distribution p_k is extremal, i.e., $(0, 0, 1, \dots, 0)$, i.e., if and only if $\psi = |\psi\rangle\langle\psi|$ is a rank-1 projector. This is true if and only if $\sum p_k^2 = 1$. The quantity $\sum p_i^2 = \text{tr}(\psi^2)$ is called the **purity** of the state ψ .

The connection with experiments is that, given an operator A :

$$\Pr(a_i) = \text{tr}(\psi |A_i\rangle\langle A_i|). \quad (1.33)$$

Then

$$E[A] = \sum_i \Pr(a_i) a_i = \text{tr}(\psi A). \quad (1.34)$$

If ψ is pure, then $E[A] = \langle \psi | A | \psi \rangle$.

What happens after measurements? What we learned in QM is that $a_i \mapsto |A_i\rangle$ following a measurement. But what does it mean to say “if we get a_i we get A_i .” What this really means is that if we measure a_i with probability $\text{Pr}(a_i)$ then we are in $|A_i\rangle$ with probability $\text{Pr}(a_i)$.

The act of measurement changes the purity of the state. Suppose we have a pure state $|\psi\rangle$. After measurement, suppose we have $\psi' = \sum_k \text{Pr}(a_i) |A_k\rangle \langle A_k| = \sum_k \text{Tr}[\psi P_k] P_k$. The state changes – sometimes called the collapse of the wavefunction. This is due to the fact that ψ , before measurement, gives some amplitudes about what potentially can happen. Note that the new state ψ' is a statistical ensemble of pure states P_k .

So, measurements change the purity of the state. We can ask something like, given a basis $\{|i\rangle\}$, what is the operation D that for X nondiagonal, DX is diagonal. This process is called **dephasing**. What can also ask changes the purity? We can answer the second question: **entanglement**.

1.4 Entanglement

Note: There is no “action at a distance.”

Now, to define entanglement, we need the notions of “here” and “there.” Consider a system made up of two parts. The mathematical description is through the tensor product:

$$\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B. \quad (1.35)$$

Let $A \in \mathcal{B}(\mathcal{H}_A)$ – A is a bounded operator in $\mathcal{H}(A)$. Suppose $\mathcal{H}_A = \text{span}\{|0\rangle, |1\rangle\}$, and $A = \sum_{i,j \in \{0,1\}} A_{ij} |i\rangle \langle j|$. Suppose $\mathcal{H}_B = \text{span}\{|0\rangle_B, |1\rangle_B\}$. Then

$$\mathcal{H} = \mathcal{H}_A \times \mathcal{H}_B = \text{span}\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}. \quad (1.36)$$

Now, $A \in \mathcal{B}(\mathcal{H}_A)$. Since the original space is $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. To generalize this A so that it acts on the whole of \mathcal{H} , we generalize by

$$A = \tilde{A} \otimes \mathcal{I}_B. \quad (1.37)$$

With this, with $\psi \in \mathcal{H}$ and $\psi_A \in \mathcal{H}_A$

$$E[A] = \text{tr}(A\psi) = \text{Tr}(\tilde{A}\psi_A). \quad (1.38)$$

What is ψ_A ? It turns out that

$$\psi_A = \text{tr}_B \psi. \quad (1.39)$$

Suppose

$$|\psi\rangle = \left(\frac{|0\rangle + |1\rangle}{\sqrt{2}} \right) \otimes |1\rangle. \quad (1.40)$$

Then

$$\begin{aligned} \psi_A &= \text{Tr}_B \psi \\ &= \text{Tr}_B \left[\left(\frac{|0\rangle + |1\rangle}{\sqrt{2}} \right) \left(\frac{|0\rangle + |1\rangle}{\sqrt{2}} \right)^\dagger \otimes |1\rangle \langle 1| \right] \\ &= \left(\frac{|0\rangle + |1\rangle}{\sqrt{2}} \right) \left(\frac{\langle 0| + \langle 1|}{\sqrt{2}} \right) \\ &= |+\rangle \langle +|. \end{aligned} \quad (1.41)$$

Is this a pure state? To check, we compute the purity:

$$\psi^2 = |+\rangle \langle +| |+\rangle \langle +| = \langle +| + \rangle = 1, \quad (1.42)$$

so ψ is pure.

As an exercise we can show that with any **pure separable state** $|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$, $\text{tr}_B \psi$ is also a pure state.

Next, consider again $|\psi\rangle = (|01\rangle + |10\rangle)/\sqrt{2}$. Is this state pure? The answer is YES because it's already given to us as a ket, i.e., the state is NOT a statistical ensemble of pure states. To show this explicitly, the density matrix is given by

$$\rho = \frac{1}{2} [|01\rangle \langle 01| + |01\rangle \langle 10| + |10\rangle \langle 01| + |10\rangle \langle 10|] \quad (1.43)$$

then

$$\text{tr}_B (\rho) = \dots = \frac{1}{2} [|0\rangle \langle 0|_A + |1\rangle \langle 1|_A]. \quad (1.44)$$

The purity of this is

$$\text{tr}(\rho^2) = \sum_i p_i^2 = \frac{1}{4} + \frac{1}{4} \neq 1. \quad (1.45)$$

So $\text{tr}_B \psi$ is no longer pure. Thus, **tracing out sometimes changes the purity**.

Definition 1.4.1 (Purity and Entanglement). We say that if the reduced/-marginal state is pure then the original state is also pure. If the reduced/-marginal state is not pure then the original state is entangled.

This is the essence of entanglement. There is no action at a distance. Rather, this is just an update in our probability distribution.

Now, something is still missing here. In physics, we're usually given interactions, not just states and operators. The next thing to learn is time evolution. Let's do that

1.5 Time evolution

Let a Hermitian $H \in \mathcal{B}(\mathcal{H})$ be given. Then the time evolution of ψ is given by

$$\psi \mapsto e^{-iHt}\psi e^{iHt}. \quad (1.46)$$

For operators, we let the operators evolve, by definition,

$$X \mapsto X = e^{iHt}Xe^{-iHt}. \quad (1.47)$$

Mathematically, this works (the check for why this has to be the case is easy to check). When we compute the expectation value

$$E_t[A] = \text{Tr}(A\psi(t)) = \text{Tr}(A(t)\psi) \quad (1.48)$$

The term in the middle is the Schrödinger representation, and the term of the RHS is the Heisenberg representation. Next, recall that any unitary can be written as $U = e^{-iHt}$ where H is some Hermitian. This means we can just look at unitaries. With this, the time evolution of X is given by

$$U(X) = U^\dagger X U. \quad (1.49)$$

This is called the *adjoint action*. Now, a very important property of unitary evolution is that it is **norm-preserving**, i.e.,

$$\|U(X)\| = \|X\|. \quad (1.50)$$

Recall **entanglement/ Rényi entropy**:

$$S_\alpha = \frac{1}{1-\alpha} \log \text{Tr}(\psi^\alpha). \quad (1.51)$$

Note that the argument of the logarithm looks like the **purity** of something. Usually, we look at a pure state ψ and look at ψ_A . If ψ_A is mixed then $S_\alpha(\psi_A) \geq 0$.

Proposition 1.5.1.

$$S(\psi) = S(\mathcal{U}(\psi)) = S(\mathcal{U}^\dagger \psi \mathcal{U}), \quad (1.52)$$

i.e., entropy is conserved under unitary evolution.

Now, there arises some trouble. We know for a fact that in time, entropy increases by the second law of thermodynamics. But we just claimed that \mathcal{U} preserves entropy. In classical mechanics, we actually have a similar theorem called the **Liouville Theorem**. This is the problem in statistical mechanics. The rest of this course is targeting this problem.

Another interesting thing is “Chaos.” In classical mechanics, we have the butterfly effect. In quantum mechanics, consider $|\psi_1\rangle$ and $|\psi_2\rangle$ very close to each other: $|\langle\psi_1|\psi_2\rangle| = 1 - \epsilon$. Now,

$$\mathcal{U}|\psi_1\rangle = \psi_1(t), \quad \mathcal{U}|\psi_2\rangle = |\psi_2(t)\rangle \quad (1.53)$$

It's easy to see that

$$|\langle\psi_1(t)|\psi_2\rangle| = 1 - \epsilon. \quad (1.54)$$

So, there is immunity to small changes in initial condition.

Finally, there is a trick to compute purity. Suppose $\psi \in \mathcal{H}$. We want to compute $\text{Tr } \psi^2$. Now, consider the following: Suppose we make a replica, $\psi \mapsto \psi \otimes \psi \otimes \dots \otimes \psi$. Consider the operator $T^{(2)}$ – a permutation operator of order 2. Let $\{|i\rangle\}, \{|j\rangle\}$ be bases of \mathcal{H} . Then $\{|ij\rangle\}$ is a basis for \mathcal{H} . $T^{(2)}$ is defined by $|ij\rangle \mapsto |ji\rangle$. Then we can show in the exercise that

$$\text{Tr } \psi^2 = \text{Tr} [T^{(2)} \psi \otimes \psi]. \quad (1.55)$$

Next, suppose ψ is pure but ψ_A isn't. What is the purity of ψ_A ? Well, we want some operator \mathcal{O} such that

$$\text{Tr } \psi_A^2 = \text{Tr} [\mathcal{O} \psi^{\otimes 2}]? \quad (1.56)$$

Again, this will be addressed in the exercises.

After these, we will look at how unitary operations make entropies grow.

1.6 Quantum Equilibration and Thermalization

Microscopic laws are time-reversal, but equilibration isn't. So, how are these possible? How can we find the real laws of physics transition from microscopic to macroscopic? We know that unitary microscopic evolution is time-reversal. How do we get non-time-reversal laws in the macroscopic scale?

Consider a non-degenerate Hamiltonian $\mathcal{H} = \sum_n E_n P_n = \sum_n E_n |n\rangle\langle n|$. Why do we make this assumption? Well, why does a Hamiltonian have a degeneracy in the first place? The answer is Noether's theorem. A degeneracy arises when when we evolve the system, some aspects of the system is conserved. An example will be a Hamiltonian with zeros everywhere except for a block on the diagonal.

Now, why do we make the assumption that \mathcal{H} is non-degenerate? First, we first note that this assumption is quite strict, because the set of non-degenerate Hamiltonians actually has measure zero! Almost all Hamiltonians are degenerate in some way. We can also have Hamiltonians with degenerate gaps! This is also not very typical of a Hamiltonian.

But in any case, we make the non-degenerate assumption to simplify our problems. Okay, so consider the Hamiltonian above. Then recall that $\phi_t = e^{-i\mathcal{H}t}\phi e^{i\mathcal{H}t}$ and $A_t = e^{i\mathcal{H}t}Ae^{-i\mathcal{H}t}$. We define

$$A_t = \sum_n p_n \langle n | A | n \rangle = \text{Tr}[A\phi_t] = \text{Tr}[A_t\phi]. \quad (1.57)$$

We have the second equality because of the commuting nature of the trace. This is also the relationship between the Schroedinger picture and Heisenberg picture.

1.6.1 Equilibration

Definition 1.6.1 (Equilibration).

$$\lim_{t \rightarrow \infty} \mathcal{U}_t \phi = \lim_{t \rightarrow \infty} e^{-i\mathcal{H}t} \phi e^{i\mathcal{H}t} = \phi_{eq} \quad (1.58)$$

if equilibration happens.

ϕ_{eq} is called the steady state, a state that is the fixed point of our unitary evolution:

$$\mathcal{U}\phi = \phi. \quad (1.59)$$

This means ϕ is firstly an eigenstate. So, if ϕ is $|E_n\rangle$ then it is an eigenstate. In fact, a steady state need not be an eigenstate. It turns out that ϕ_{eq} can, for example, be a Gibbs state:

$$\phi = \frac{e^{-\beta\mathcal{H}}}{Z}. \quad (1.60)$$

This state does not change in time if **it commutes with the Hamiltonian**:

$$\dot{\phi} = \frac{d}{dt}\phi_t = \dots = -i[\phi, \mathcal{H}] = 0. \quad (1.61)$$

So, **steady states are functions of the Hamiltonian**, or, more generally, a probability distribution over the eigenstates.

Now, we want to answer the question whether equilibration is possible? The short answer is NO, because of time reversal. In other words, consider a pure state $|\psi\rangle$. Suppose it evolves to the Gibbs state. This state is no longer pure because the probability distribution is no longer extremal. This means equilibration via unitary evolution is not possible in general.

How do we show this rigorously? Consider the distance:

$$\lim_{t \rightarrow \infty} \|\phi_t - \phi_{eq}\| = 0 \quad (1.62)$$

if equilibration is possible. Now, by definition,

$$\lim_{t \rightarrow \infty} \|\mathcal{U}_t\phi - \mathcal{U}_t\phi_{eq}\| = 0, \quad (1.63)$$

since $\mathcal{U}_t\phi_{eq} = \phi_{eq}$. So,

$$\lim_{t \rightarrow \infty} \|\mathcal{U}_t\phi - \mathcal{U}_t\phi_{eq}\| = \|\phi - \phi_{eq}\| = 0. \quad (1.64)$$

This happens if and only if $\phi = \phi_{eq}$ to begin with. This shows that equilibration in quantum mechanics is not possible. But equilibration DOES happen! So how does equilibration happen?

What we're forgetting when looking at time evolution of quantum systems so far isn't the full Hamiltonian! We're forgetting the environment, which has its own Hamiltonian. Consider the following: $\phi = \phi_{env} \otimes \phi_{sys}$ obeying the Hamiltonian $\mathcal{H} = \mathcal{H}_{env} + \mathcal{H}_{sys} + \mathcal{H}_{int}$ where \mathcal{H}_{int} is the interaction Hamiltonian. Then

$$\psi_{sys}(t) = \text{tr}_{env} \psi(t) = \text{tr}_{env} (\mathcal{U}_t \phi_{sys} \otimes \psi_{env}) \quad (1.65)$$

Note that ψ is not entangled, but $\psi_{sys}(t)$ can be an entangled state. Now since tracing out is not purity-non-preserving operation, this corresponds to a evolution that is not unitary!

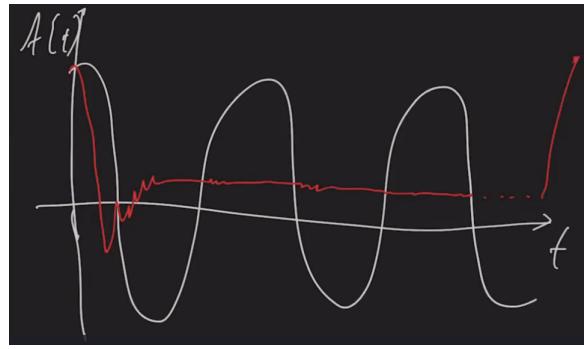
1.6.2 Thermalization: Turtles All The Way Down!

A small system thermalizes in some bath, which thermalizes in a larger bath, and so on. So, to avoid problems, we will restrict ourselves to unitary time evolution and closed systems. What do we do to get unstuck? We want just the small system to equilibrate. But we also know that ϕ_t never equilibrates!

What are we missing?

The way out is to change our way of thinking. We should not think of ϕ are a “thing” that equilibrates. Rather, we should look at how expectations $\langle A_t \rangle = \text{tr}[A\phi]$ equilibrate! We should always think about observables, because these are things we can observe! So now, the question becomes whether $\langle A_t \rangle$ can equilibrate.

Theorem 1.6.1 (Von Neumann). $\langle A_t \rangle$ is quasi-periodic, and so $\lim_{t \rightarrow \infty} A_t$ does not exist.



Almost always we’re very close to the “equilibrium” value, even though $\langle A_t \rangle$ (red) is quasi-periodic. So, consider

$$\Pr[|A_t - A_{eq}| > \epsilon] \ll 1. \quad (1.66)$$

This is called equilibration in probability, or that A_t has typical values $\sim A_{eq}$. Now, what is A_{eq} ? How do we compute the typical values of A_t ?



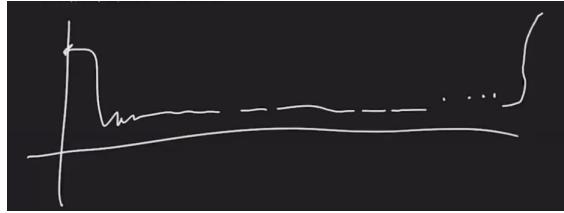
Intuitively, it makes sense to define A_{eq} as the time average:

$$A_{eq} = \bar{A} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T A_t dt. \quad (1.67)$$

It turns out that

$$A_{eq} = \bar{A} = \text{tr}(A\bar{\phi}). \quad (1.68)$$

This means the **typical** value of A_t depends on $\bar{\phi}$.



Okay, but notice that A_t is quasi-periodic, which means it has to go back to its initial value at some point: Right, but we want, also, that the probability that A_t returns to its initial value is small. This, of course, depends on the probability that ϕ_t returns to ϕ , which is

$$\overline{|\langle \phi | \phi_t \rangle|} \quad (1.69)$$

This is called the **return probability**, or the **Loschmidt echo**.

Now, just to summarize. So far, we have seen that if our quantum system is open then time evolution is unitary. And if the time evolution is unitary then the quantum system does not reach an equilibrium state. And so the only thing that is compatible with QM is that time average-expectation values equilibrate in probability. In other words, $A_t \rightarrow \bar{A}$ in probability, i.e.,

$$\Pr(|A_t - \bar{A}| > \epsilon) \ll 1. \quad (1.70)$$

If we can ensure this can happen then it is enough. Now,

$$\bar{A} = \text{Tr}[A\bar{\phi}] \quad (1.71)$$

with

$$\bar{\phi} = \sum_n P_n \phi P_n \quad (1.72)$$

where $\mathcal{H} = \sum_n E_n P_n$. We call $\bar{\phi}$ the **dephasing** in the basis of \mathcal{H} :

$$D_{\mathcal{H}}\phi = \bar{\phi}. \quad (1.73)$$

$D_{\mathcal{H}}$ basically erases all the off-diagonal elements of ϕ . The resulting quantity has matrix elements $\phi_{kk} = p_k = |\langle \phi | E_n \rangle|^2$.

Now, recall the Loschmidt echo:

$$\mathcal{L}_t = |\langle \phi | \phi_t \rangle|^2. \quad (1.74)$$

In the exercise, we show that

$$\overline{\mathcal{L}_t} = \text{tr } \bar{\phi}^2 = \text{tr } [(D_{\mathcal{H}}\phi)^2]. \quad (1.75)$$

which is the purity of $D_{\mathcal{H}}\phi$. If ϕ is pure then $\text{tr } \phi^2 = 1$. Then $\text{Tr } \bar{\phi}^2 = \sum_k p_k^2 \neq 1$. So, $D_{\mathcal{H}}\phi$ is no longer pure. Is something fishy going on? From $\phi \rightarrow \phi_t$ are use a unitary operator, so the purity remains the same. The change in purity is due to the time average, which is not a unitary evolution.

$\bar{\phi}$ is NOT an actual state. So why are we using this state at all? The point of using this fictitious state is so that we get a correct answer: $A_t = \text{tr}[A\phi_t] \rightarrow \text{tr}[A\bar{\phi}]$. So, now we would like to understand what makes the fluctuation $\text{tr}[A\phi_t] \approx \text{tr}[A\bar{\phi}]$ small.

Since, for pure ϕ , we have

$$\bar{\mathcal{L}}_t = \overline{|\langle \phi | \phi_t \rangle|^2} = \text{Tr}[\bar{\phi}^2] = p_n^2 \quad (1.76)$$

which is the probability of going back to $\text{tr } \phi$. Does this probability depend on the initial state ϕ or not? The answer is YES, of course.

Now, if ϕ is not pure then we want to look at the inner product of operators:

$$\langle A | B \rangle = \text{Tr}[A^\dagger B]. \quad (1.77)$$

which is just the HS inner product. In this case, the Loschmidt echo becomes

$$\mathcal{L}_t = |\langle \phi | \phi_t \rangle|^2 = \text{Tr}[\phi^\dagger \phi_t]. \quad (1.78)$$

In any case, we showed that, $\bar{\mathcal{L}}_t = \sum_n p_n^2$ which is just the purity $D_{\mathcal{H}}\phi$, which depends on ϕ .

Example 1.6.1. Consider $|\phi\rangle = (1/\sqrt{2})(|E_0\rangle + |E_1\rangle)$, corresponding to the Hamiltonian $\mathcal{H} = \sum_n E_n |E_n\rangle \langle E_n|$. Then

$$|\phi_t\rangle = \frac{1}{\sqrt{2}} [e^{-iE_0 t} |E_0\rangle + e^{-iE_1 t} |E_1\rangle]. \quad (1.79)$$

Then

$$\langle \phi | \phi_t \rangle = \frac{1}{2} [e^{-iE_0 t} + e^{-iE_1 t}] \quad (1.80)$$

And so the time average is $1/2$. So this is something with low probability of return, on average.

One can show that if the original state is very spread out in the basis states of the Hamiltonian, then the return probability is very small, which means equilibration occurs very quickly. Now, how likely is it for $\sum_n p_n^2$ to be very small, i.e., ϕ is extremely spread out in the basis states of the Hamiltonian? NOT VERY LIKELY! Let's show this mathematically.

Consider a random state ϕ , then we have $\text{tr}((D_h\phi)^2) = \text{tr}\bar{\phi}^2 = \sum_n p_n^2$. What is the average of this over all the ϕ 's what do we get? Well, this is given by

$$\langle \text{tr}[(D_{\mathcal{H}}\phi)^2] \rangle_{\phi} \equiv \int \text{tr}[(D_{\mathcal{H}}\phi)^2] d\phi \quad (1.81)$$

where of course we require normalization $\int d\phi = 1$. Taking this integral requires knowing the Haar-measure. One can easily prove that

$$\langle \phi \rangle_{\phi} = \frac{\mathcal{I}}{d} \quad (1.82)$$

which is the microcanonical ensemble. Now to the harder problem. Call $\omega = D_h\phi$. What we want to compute now is

$$\langle \text{tr} \omega^2 \rangle_{\phi}. \quad (1.83)$$

Now, recall from the exercise to get

$$\text{tr} \omega^2 = \text{tr} \left(T^{(2)} \omega^{\otimes 2} \right). \quad (1.84)$$

With this and by linearity we can proceed:

$$\begin{aligned} \langle \text{tr} \omega^2 \rangle &= \langle \text{tr} \left(T^{(2)} \omega^{\otimes 2} \right) \rangle \\ &= \text{tr} \left[T^{(2)} \langle \omega^{\otimes 2} \rangle \right] \\ &= \text{tr} \left[T^{(2)} \langle (D_{\mathcal{H}}\phi)^{\otimes 2} \rangle \right] \\ &= \text{tr} \left[T^{(2)} D_{\mathcal{H}}^{\otimes 2} \langle \phi^{\otimes 2} \rangle \right]. \end{aligned} \quad (1.85)$$

Now, group theory says that

$$\langle \phi^{\otimes 2} \rangle = \frac{\mathcal{I}_{\mathcal{H}^{\otimes 2}} + T^{(2)}}{d(d+1)}. \quad (1.86)$$

Plugging this back into the exercise, we get (in the exercise) $2/(d+1)$. So, the purity of the state after we erase the diagonal elements, on average, is $2/(d+1)$. This means that if the Hilbert space is very large, then the purity is very close to zero.

Now, knowing the average of something does not mean much. It might not be the typical case. So is this typical, or does this have a lot of fluctuations, i.e., is the variance of $\bar{\mathcal{L}}_t$ small or large? What can we say about $\text{Var}(\bar{\mathcal{L}}_t)$. First, $\bar{\mathcal{L}}$ is nonnegative. Second, its average is very small. So its variance must also be small. This comes from the **Markov Inequality**:

$$X \geq 0, \text{Pr}(X \geq a) \leq E[X]/a. \quad (1.87)$$

We can make $a = 1/\sqrt{d}$, so that $\Pr(X \geq 1/\sqrt{d}) \leq 1/\sqrt{d}$.

Now, let's consider entanglement. Consider $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ and ϕ pure in \mathcal{H} . If ϕ is entangled then $\text{tr } \phi_A^2 < 1$. Next, recall that $S_2 = -\log \text{tr } \phi_A^2$. Now, we ask what $\langle -\log \text{tr } \phi_A^2 \rangle_\phi$ is, to see how typical entanglement is. Now,

$$\begin{aligned}\langle -\log \text{tr } \phi_A^2 \rangle_\phi &\geq -\log \langle \text{tr } \phi_A^2 \rangle_\phi \\ &= -\log \langle \text{tr } \phi_A^{\otimes 2} T_A^{(2)} \rangle_\phi \\ &= -\log \text{tr} [T_A^{(2)} \langle \phi_A^{\otimes 2} \rangle].\end{aligned}\quad (1.88)$$

In the exercise, we show that

$$\langle -\log \text{tr } \phi_A^2 \rangle_\phi = \log \frac{d_A + d_B}{d_A d_B + 1}. \quad (1.89)$$

Now, call the purity $p = \text{tr } \phi_A^2$. Then we show that

$$\langle p \rangle_\phi = \langle \text{tr } \phi_A^2 \rangle_\phi \frac{d_A + d_B}{d_A d_B + 1}. \quad (1.90)$$

Now, if A is our system of interest, then $d_B \gg d_A$ since B is the “universe” in this case. But d_A is also quite large by itself, so we have the typical thermodynamic situation: $d_B \gg d_A \gg 1$. With this,

$$\langle p \rangle_\phi \approx \frac{1}{d_A} = \frac{1}{2^{N_A}}. \quad (1.91)$$

And so

$$\log_2 d_A \geq S_2 \geq \log_2 d_A \quad (1.92)$$

And so, we have that **almost all states are maximally entangled**.

1.7 Equilibration

We introduce the notion of distance between states.

$$\begin{aligned} d(\rho, \sigma) &= \max_P |\mathrm{tr} P_\rho - \mathrm{tr} P_\sigma| \\ &= \frac{1}{2} \|\rho - \sigma\|_1 \\ &= \frac{1}{2} \mathrm{tr} \left[\sqrt{(\rho - \sigma)^\dagger (\rho - \sigma)} \right]. \end{aligned} \quad (1.93)$$

Consider a map Φ that is trace-preserving and maps a state to a state. For example, $\mathcal{U}_t, D_H, \mathrm{tr}_B$ are examples.

Theorem 1.7.1. Let a trace-preserving map Φ be given, then

$$d(\Phi(\rho), \Phi(\sigma)) \leq d(\rho, \sigma). \quad (1.94)$$

Theorem 1.7.2.

$$d(\rho_A, \sigma_A) \leq d(\rho, \sigma). \quad (1.95)$$

Theorem 1.7.3.

$$\|\rho - \sigma\|_1 \leq \sqrt{d} \|\rho - \sigma\|_2 = \sqrt{d \mathrm{tr}(\rho - \sigma)^2} \quad (1.96)$$

1.7.1 Equilibration

Let $\omega = D_H \phi = \bar{\phi}$ be given and $\omega_S = \mathrm{tr}_B \omega$, with $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_B$. Remember that $A_t \sim \bar{A}$ and $\mathrm{tr}(A\phi_t) \sim \mathrm{tr} A\omega$, and thus we want to show

$$\overline{\mathrm{tr}[A\phi_{St}]} \sim \overline{\mathrm{tr}[A\omega_S]} \quad (1.97)$$

This means that we want $\overline{d(\phi_{St}, \omega_S)}$ small.

Theorem 1.7.4.

$$\overline{d(\phi_{St}, \omega_S)} \leq \frac{1}{2} \sqrt{d_S^2 \mathrm{tr} \omega^2} \quad (1.98)$$

where $\omega = D_H \phi$.

Proof. We had lemma that says

$$d(\rho, \sigma) = \frac{1}{2} \mathrm{tr} \sqrt{(\rho - \sigma)^2} \leq \frac{1}{2} \sqrt{d \mathrm{tr}(\rho - \sigma)^2} \quad (1.99)$$

Because $\sqrt{\cdot}$ is concave, we can bring the average inside. So,

$$\overline{d(\phi_{St}, \omega_S)} \leq \frac{1}{2} \sqrt{d_S \mathrm{tr}(\phi_{St} - \omega_S)^2}. \quad (1.100)$$

Now,

$$\phi_{St} - \omega_S = \sum_{k \neq l} c_k c_l^\dagger e^{-i(E_k - E_l)t} \text{tr}_B |E_k\rangle \langle E_l|. \quad (1.101)$$

Thus,

$$\overline{\text{tr}(\phi_{St} - \omega_S)^2} = \sum_{k \neq l, m \neq n} \overline{\tau_{klmn}} \text{tr} (\text{tr}_B |E_k\rangle \langle E_l| \cdot \text{tr}_B |E_m\rangle \langle E_n|). \quad (1.102)$$

Explicitly,

$$T_{klmn} = c_k c_l^\dagger c_m c_n^\dagger e^{\overline{-i(E_k - E_l + E_m - E_n)t}}. \quad (1.103)$$

Now, since we have non-degenerate Hamiltonians, so

$$\overline{e^{-i(t)}} = \delta_{kn} \delta_{lm}. \quad (1.104)$$

Plug this in to get

$$\begin{aligned} \overline{\text{tr}(\dots)^2} &= \sum_{k \neq l} |c_k|^2 |c_l|^2 \cdot \text{tr} [\text{tr}_B |E_k\rangle \langle E_l| \cdot \text{tr}_B |E_l\rangle \langle E_k|] \\ &= \sum_{k \neq l} \text{tr}_B (\text{tr} |c_k|^2 P_k \cdot \text{tr} |c_l|^2 P_l). \end{aligned} \quad (1.105)$$

Now add the $k = l$ terms and subtract them. Recall that

$$\sum_k |c_k|^2 P_k = \omega \quad (1.106)$$

so that

$$\overline{\text{tr}(\dots)^2} = \text{tr}_B \omega_B^2 - \sum_k |c_k|^4 \text{tr} [\text{tr}_B P_k]^2. \quad (1.107)$$

The term on the right is upper bounded by one. So,

$$\overline{\text{tr}(\dots)^2} \leq \text{tr}_B \omega_B^2 \leq d_S \text{tr} \omega^2. \quad (1.108)$$

Plugging everything back in we get

$$\overline{d(\phi_{St}, \omega_S)} \leq \frac{1}{2} \sqrt{d_S^2 \text{tr} \omega^2} \quad (1.109)$$

so we're done. \square

Typically, $\text{tr} \omega^2 \sim 1/d = 1/(d_S d_B)$, and so typically,

$$\overline{d(\phi_{St}, \omega_S)} \leq \frac{1}{2} \sqrt{d_S/d_B} \ll 1 \quad (1.110)$$

because $d_S/d_B \ll 1$. Here “typically” is in the sense of the Haar-average. To prove this theorem, we want to compute

$$\begin{aligned} \langle \text{tr } \omega^2 \rangle_\phi &= \langle \text{tr}(D_H \phi)^2 \rangle_\phi \\ &= \text{tr} \langle (D_H \phi)^{\otimes 2} T^{(2)} \rangle_\phi \\ &= \text{tr } D_H^{\otimes 2} \langle \phi^{\otimes 2} \rangle_\phi T^{(2)} \\ &\sim \frac{1}{d}. \end{aligned} \tag{1.111}$$

So we have that

$$\langle \overline{d(\phi_{St}, \omega_S)} \rangle_\phi \leq \frac{1}{2} \sqrt{d_S^2 \langle \text{tr } \omega^2 \rangle}. \tag{1.112}$$

From here, we see that $\phi_{St} \sim \omega_S$ almost always. But what is $\omega_S = \text{tr}_B D_H \phi$, typically? We need a new canonical state that doesn’t depend on ϕ and establish the distance between ω_S and that state which we will call Ω_S .

Theorem 1.7.5.

$$\langle d(\omega_S, \Omega_S) \rangle_\phi \ll 1 \tag{1.113}$$

where $\omega_S = \text{tr}_B D_H \phi$ and $\Omega_S = \mathcal{I}_S/d_S$ is the canonical state.

1.8 Resource Theories and Quantum Information

1.9 Quantum Thermal Operations

1.10 Fluctuation Theorems and Quantum Information

1.11 Quantum Thermal Machines

1.12 Exercises

1. Density Operators Show that given the distribution of states $(q_j, |\psi_j\rangle)$ the probability of getting an outcome a_i when measuring $A = \sum_i a_i |a_i\rangle \langle a_i| \equiv \sum_i a_i P_i$ is given by

$$\Pr(a_i) = \text{tr}(\psi P_i) \quad (1.114)$$

where we have define the density matrix of the system as

$$\psi = \sum_j q_j |\psi_j\rangle \langle \psi_j| \quad (1.115)$$

Proof. This is not very hard:

$$\begin{aligned} \Pr(a_i) &= \langle a_i | \psi | a_i \rangle \\ &= \sum_j \langle a_i | q_j |\psi_j\rangle \langle \psi_j | a_i \rangle \\ &= \sum_j q_j |\psi_j\rangle \langle \psi_j | a_i \rangle \langle a_i | \\ &= \text{tr}(\psi P_i). \end{aligned} \quad (1.116)$$

Alternatively, we can do this like the following: Since the trace is linear, the probability of getting the outcome a_i is

$$\Pr(a_i) = \text{tr} \left[\sum_j q_j \psi_j P_i \right] = \text{tr} [\psi P_i]. \quad (1.117)$$

This is perhaps a better way to do the problem than the previous approach. In any case they're equivalent. \square

2. Partial Trace Mixed states appear naturally also when considering subsystems of a composite quantum system. Consider

$$A = \tilde{A}_S \otimes I_{\bar{S}} \quad (1.118)$$

in a Hilbert space with tensor product structure $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_{\bar{S}}$. It is a natural question to ask what is the state $\rho_S \in \mathcal{B}(\mathcal{H}_S)$ onto the subsystem \mathcal{H}_S that can return all the right expectation values when performing any measurement of operators with that support, namely, the \tilde{A}_S . We ask that

$$\text{tr}(A\rho) = \text{tr}(\tilde{A}_S \rho_S). \quad (1.119)$$

Show that the solution of this equation is

$$\rho_S = \text{tr}_{\bar{S}}(\rho) \quad (1.120)$$

where the operation of partial trace is defined by

$$\text{tr}_{\bar{S}} X = \text{tr}_{\bar{S}} \sum_{i_S i_{\bar{S}} j_S j_{\bar{S}}} X_{i_S i_{\bar{S}} j_S j_{\bar{S}}} |i_S i_{\bar{S}}\rangle \langle j_S j_{\bar{S}}| = \sum_{i_S j_S} X_{i_S j_S} |i_S\rangle \langle j_S|. \quad (1.121)$$

Proof. The definition of partial trace above is not very useful due to bad notation. So we will consider an equivalent definition:

$$\text{tr}_{\bar{S}}(\rho) = \sum_{i_{\bar{S}}} \langle i_{\bar{S}} | \rho | i_{\bar{S}} \rangle. \quad (1.122)$$

Note that this is an operator, not a number. We want to show

$$\text{tr}(A\rho) = \text{tr}\left(\tilde{A}_S \text{tr}_{\bar{S}}(\rho)\right). \quad (1.123)$$

Well, by the definition of the trace in terms of the inner product:

$$\begin{aligned} \text{tr}(A\rho) &= \text{tr}(\rho A) \\ &= \text{tr}\left(\rho(\tilde{A}_S \otimes I_{\bar{S}})\right) \\ &\equiv \text{tr}\left(\rho(\tilde{A} \otimes I)\right) \\ &= \sum_{i_S} \sum_{i_{\bar{S}}} \langle i_S i_{\bar{S}} | \rho(\tilde{A}_S \otimes I_{\bar{S}}) | i_S i_{\bar{S}} \rangle \\ &= \sum_{i_S} \sum_{i_{\bar{S}}} \langle i_S i_{\bar{S}} | \rho | \tilde{A}_S i_S \otimes i_{\bar{S}} \rangle \\ &= \sum_{i_{\bar{S}}} \sum_{i_S} \langle i_S | \langle i_{\bar{S}} | \rho | i_{\bar{S}} \rangle | \tilde{A}_S i_S \rangle \\ &= \text{tr}\left(\tilde{A}_S \text{tr}_{\bar{S}}(\rho)\right) \\ &= \text{tr}\left(\tilde{A}_S \rho_S\right). \end{aligned} \quad (1.124)$$

□

3. Rényl Entropy The Rényi entropy for Rényi index q is defined as

$$S_q(\rho) = \frac{1}{1-q} \log \text{Tr} \rho^q = \frac{1}{1-q} \log \left(\sum_i \lambda_i^q \right) \quad (1.125)$$

where the logarithm is again taken to be base two and λ_i are the eigenvalues of ρ .

1. Show that for the completely mixed density operator $\rho = I/d$ in d -dimensional Hilbert space, the von Neumann entropy is given by

$$S(I/d) = \log d. \quad (1.126)$$

Proof. This is very easy. The eigenvalues of I/d are $1/d$. The von Neumann entropy is given as

$$S_{vn} = - \sum \lambda_x \log \lambda_x = - \sum \frac{1}{d} \log \frac{1}{d} = \log d. \quad (1.127)$$

□

2. Show that

$$S_0(\rho) = \lim_{q \rightarrow 0+} S_q(\rho) = \log(\text{rank } \rho) \quad (1.128)$$

where $\text{rank } \rho$ is the rank of the density matrix ρ .

Proof. This is also quite easy. When $q \rightarrow 0+$, the definition says the leading factor becomes +1. Inside the logarithm, we just get a sum of λ_i^0 . Now, of course these values only approach 1 if they're not zero. So, the sum is exactly the rank of the matrix ρ . Thus,

$$S_0(\rho) = \log \text{rank } \rho. \quad (1.129)$$

□

3. Show that

$$\lim_{q \rightarrow 1} S_q(\rho) = S_1(\rho) \quad (1.130)$$

where $S_1(\rho)$ is the von Neumann entropy.

Proof. This is quite hard. The proof is contained in this [paper](#) by Muller-Lennert in 2013.

But we can also do this using l'Hopital's rule:

$$\begin{aligned} \lim_{q \rightarrow 1} S_q(\rho) &= \frac{1}{\ln 2} \lim_{q \rightarrow 1} \left[\frac{\log(\sum_i \lambda_i^q)}{1-q} \right] \\ &= \frac{1}{\ln 2} \lim_{q \rightarrow 1} \left[\frac{\frac{1}{\sum_i \lambda_i^q} \sum_i \frac{d}{dq} \lambda_i^q}{-1} \right] \\ &= \frac{1}{\ln 2} \lim_{q \rightarrow 1} \left[-\frac{\sum_i \lambda_i^q}{\sum_i} \lambda_i^q \ln \lambda_i \right] \\ &= -\frac{1}{\sum_i \lambda_i} \sum_i \frac{\ln \lambda_i}{\ln 2} \\ &= -\sum_i \lambda_i \log \lambda_i \\ &= S_{vn}(\rho). \end{aligned} \quad (1.131)$$

□

4. Consider the Rényi entropy for the case where the Rényi index is an integer n with $n \geq 2$. If ρ corresponds to a pure state, calculate $S_n(\rho)$. Show that if ρ_{pure} corresponds to a pure state, then

$$S_n(\rho_{\text{pure}}) = 0. \quad (1.132)$$

Proof. By definition, pure states are idempotents. This means $\text{tr}(\rho) = \text{tr}(\rho^n) = 1$ always, by normalization. Now, diagonalize $\rho, \rho^2, \dots, \rho^n$ we see that the condition $\text{tr}(\rho^n) = 1$ implies $\sum \lambda_\rho^n = 1$. So, by the definition of the Renyi entropy, we have that

$$S_n(\rho_{\text{pure}}) = 0. \quad (1.133)$$

□

4. Von Neumann Entropy Compute the von Neumann entropy of the Gibbs state.

Proof. The Gibbs state is given by

$$\rho = \frac{e^{-\beta \mathcal{H}}}{\text{tr}(e^{-\beta \mathcal{H}})} \quad (1.134)$$

The von Neumann entropy of ρ is given by

$$S_{vn}(\rho) = -\sum \lambda_x \log \lambda_x \quad (1.135)$$

where λ_x are the eigenvalues of ρ . Suppose the Hamiltonian \mathcal{H} has the spectrum $\{\sigma_i\}$, then by the spectral mapping theorem we have that the spectrum of $e^{-\beta \mathcal{H}} = \{e^{-\beta \sigma_i}\}$, and so it follows that

$$\text{tr}(e^{-\beta \mathcal{H}}) = \sum e^{-\beta \sigma_i} = \mathcal{Z} \quad (1.136)$$

where of course \mathcal{Z} is the partition function. With these, the von Neumann entropy is given by

$$S_{vn} = -\frac{1}{\mathcal{Z}} \sum e^{-\beta \sigma_i} \log \frac{e^{-\beta \sigma_i}}{\mathcal{Z}} = \log \mathcal{Z} + \beta \langle E \rangle \quad (1.137)$$

□

5. Swap Operator. One of the problems of the entanglement entropies is that to know them one has to know all the eigenvalues of the reduced density operator, which amounts to knowing the wave-function. In other words, entanglement is not the expectation value of an observable. Define the swap operator as the operator on two copies of the original Hilbert space, namely $T^{(2)} \in \mathcal{H} \otimes \mathcal{H}$. This operator is defined from the action on the basis states as

$$T^{(2)} |\phi_i \phi_j\rangle = |\phi_j \phi_i\rangle. \quad (1.138)$$

1. **Purity as expectation value.** Given a state $\sigma \in \mathcal{B}(\mathcal{H})$, show that

$$\text{tr } \sigma^2 = \text{tr} (T^{(2)} \sigma^{\otimes 2}). \quad (1.139)$$

Proof.

$$\begin{aligned}
\text{tr} \left(T^{(2)} \sigma^{\otimes 2} \right) &= \sum_{i,j} \langle ij | T^{(2)} \sigma^{\otimes 2} | ij \rangle \\
&= \sum_{i,j} \langle ij | T^{(2)} (\sigma |i\rangle \otimes \sigma |j\rangle) \\
&= \sum_{i,j} \langle ij | T^{(2)} (\sigma |j\rangle \otimes \sigma |i\rangle) \\
&= \sum_{i,j} \langle ij | \sigma^2 |ji\rangle \\
&= \sum_i \langle i | \sigma^2 |i\rangle \\
&= \text{tr} (\sigma^2). \tag{1.140}
\end{aligned}$$

□

2. **Trace of $T^{(2)}$.** Compute the trace of $T^{(2)}$.

Proof. Just use $\sigma = \mathcal{I}$. Then the $\text{tr}(T^{(2)})$ is just the dimension of the space. □

3. **Frobenius norm.** Show that for any linear operator X , the Frobenius norm reads

$$\|X\|_2^2 = \text{tr} \left(T^{(2)} X \otimes X^\dagger \right). \tag{1.141}$$

Proof. This one is similar to part 1:

$$\begin{aligned}
\|X\|_2^2 &= \sqrt{\text{tr}(X^\dagger X)} \\
&= \sum_i \langle i | X^\dagger X | i \rangle \\
&= \sum_{i,j} \langle i | X^\dagger | i \rangle \langle j | X | i \rangle \\
&= \sum_{i,j} (|i\rangle \otimes |j\rangle) (X^\dagger |j\rangle \otimes X |i\rangle) \\
&= \sum_{i,j} \langle ij | T^{(2)} (X |i\rangle \otimes X^\dagger |j\rangle) \\
&= \text{tr} \left(T^{(2)} X \otimes X^\dagger \right). \tag{1.142}
\end{aligned}$$

□

6. Unitary Evolution

1. **Norm invariance.** Show that the Frobenius norm and the Operator norm are preserved by unitary operations.

Proof. This is one easy. For the Frobenius norm, we just use the fact that $\text{tr}(AB) = \text{tr}(BA)$. Everything else follows. In general, things follow from the fact that unitaries preserve the inner product. \square

2. **Entropy invariance.** Show that all Entropies are preserved by unitary evolution.

Proof. Entropies depend on something that looks like $\text{tr} \psi^\alpha$ for some α and the state ψ . Again, because $\text{tr}(BA) = \text{tr}(AB)$, it's an easy argument to show the trace remains invariant. \square

7. **Entanglement and partial trace.** Show that, if a state $\sigma(t)$ is evolving in time, the purity of the marginal state $\sigma_A(t)$ is given by

$$\text{tr} \sigma_A^2(t) = \text{tr} \left[T^{(2)} \otimes I_B^{\otimes 2} U^{\otimes 2} \sigma^{\otimes 2} U^\dagger \right] \quad (1.143)$$

where U is the unitary evolution operator.

Proof. To do this we use the identity

$$\text{tr}(X_A \rho_A) = \text{tr} (X \otimes I_B \rho) \quad (1.144)$$

then recall that $\sigma(t) = U \sigma U^\dagger$ and take the tensor power two. \square

8. **Time evolution by Hamiltonian.** Be $\mathcal{U}(X) = e^{iHt} X e^{-iHT} \equiv X_t$ the time evolution generated by a non-degenerate Hamiltonian $H = \sum_n E_n P_n$.

1. **Time averaged operator** Define the time average as $\bar{X} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T X_t dt$. Show that

$$\bar{X} = \sum_n P_n X P_n. \quad (1.145)$$

Proof. This needs quite a lot of manipulations. The key is to expand X in the

energy eigenbasis:

$$\begin{aligned}
\bar{X} &= \frac{1}{T} \lim_{T \rightarrow \infty} \int_0^T e^{iHt} X e^{-iHT} dt \\
&= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T e^{iE_n t} \sum_{n,m} X_{nm} |n\rangle \langle m| e^{-iE_m t} dt, \quad X_{nm} = \langle n| X |m\rangle \\
&= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T e^{i(E_n - E_m)t} \sum_{n,m} X_{nm} |n\rangle \langle m| dt \\
&= \langle n|m \rangle \sum_{n,m} |n\rangle X_{nm} \langle m| \\
&= \langle n|m \rangle \sum_{n,m} |n\rangle \langle n| X |m\rangle \langle m| \\
&= \sum_n |n\rangle \langle n| X |n\rangle \langle n| \\
&= \sum_n P_n X P_n.
\end{aligned} \tag{1.146}$$

So we're done. \square

9. Quantum Measurement. Show that, after measuring on the state σ some observable $\Omega = \sum_i o_i |i\rangle \langle i| \equiv \sum_i o_i \omega_i$, the resulting state is

$$\sigma' = \sum_i \omega_i o_i \omega_i. \tag{1.147}$$

Proof. Well, measurement is basically a projection:

$$\begin{aligned}
\sigma' &= \sum_i |i\rangle \langle i| \sigma \\
&= \sum_i |i\rangle \langle i| \sum_j \sigma_j |i\rangle \langle j| \\
&= \sum_{i,j} |i\rangle \langle i| |j\rangle \sigma_j \langle j| \\
&= \sum_i |i\rangle \sigma_i \langle i| \\
&= \sum_i |i\rangle \langle i| \sigma |i\rangle \langle i| \\
&= \sum_i \omega_i \sigma \omega_i
\end{aligned} \tag{1.148}$$

where we have used the fact that the matrix element $\sigma_i = \sigma_{ii} = \langle i| \sigma |i\rangle$. \square

10. Time averages by Hamiltonian evolution. Be $\mathcal{U}(X) = e^{iHt} X e^{-iHt} \equiv X_t$ the time evolution generated by a non-degenerate Hamiltonian $H = \sum_n E_n P_n$

and $\phi_t = \mathcal{U}\phi = e^{-iHt}\phi e^{iHt}$ the time evolved state. Let p_n be the probability of being in the state P_n , namely $p_n = \text{tr}(\phi P_n)$.

- **Time averaged state.** Let ϕ be a pure state with amplitudes $|\phi\rangle = \sum_n a_n |E_n\rangle$. Show that

$$\bar{\phi} = \sum_n P_n \phi P_n = \sum_n p_n P_n = \sum_n |a_n|^2 P_n = D_H \phi \quad (1.149)$$

and compute its purity.

Proof. Well, we have already shown that $\bar{\phi} = \sum_n P_n \phi P_n$ earlier. So we only need to verify the subsequent equalities:

$$\begin{aligned} \bar{\phi} &= \sum_n P_n \phi P_n \\ &= \sum_n |n\rangle \langle n| \phi P_n \\ &= \sum_n |n\rangle \langle n| \sum_m a_m^\dagger a_m |m\rangle \langle m| P_n \\ &= \sum_n \sum_m a_m^\dagger a_m |n\rangle \langle n| |m\rangle \langle m| P_n \\ &= \sum_n \sum_m a_m^\dagger a_m |m\rangle \langle m| P_n \\ &= \sum_n \sum_m a_m^\dagger a_m P_m P_n. \end{aligned} \quad (1.150)$$

Now note that $P_m P_n = P_n \delta_{mn}$, so we get $= \sum_n |a_n|^2 P_n$ for free. With this we're done. The equality with $D_H \phi$ can readily be checked using its definition. Finally,

$$\text{tr}(\phi P_n) = \text{tr} \left(\sum_n |a_n|^2 P_m P_n \right) = |a_n|^2. \quad (1.151)$$

So we're done. The purity in this case is

$$\text{tr}(\bar{\phi}^2) = \sum_n |p_n|^2. \quad (1.152)$$

□

- **Average Loschmidt Echo.** Define the Loschmidt echo as $\mathcal{L}_t = |\langle \phi | \phi_t \rangle|^2$. Show that $\bar{\mathcal{L}}_t = \text{Tr} \bar{\phi}^2$.

Proof. Well, just crank:

$$\begin{aligned}
\overline{\mathcal{L}_t} &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \langle \phi | \phi_t \rangle \langle \phi_t | \phi \rangle dt \\
&= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T dt \left[\sum_n a_n^\dagger \langle E_n | e^{-iHt} \phi e^{iHt} \rangle \right] \times \text{c.c.} \\
&= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T dt \left[\sum_n \sum_m a_n^\dagger a_m \langle E_n | E_m \rangle e^{-i(E_n - E_m)t} \right] \times \text{c.c.} \\
&= \sum_{n,m} a_n^\dagger a_n a_m^\dagger a_m \\
&= \text{tr} \left(\sum_n a_n^\dagger a_n |n\rangle \langle n| \sum_m a_m^\dagger a_m |m\rangle \langle m| \right) \\
&= \text{tr} \left(\left[\sum_n |a_n|^2 P_n \right]^2 \right) \\
&= \text{tr}(\bar{\phi}^2).
\end{aligned} \tag{1.153}$$

In the case of mixed states, we can define the Loschmidt echo as

$$\text{Tr}(\phi^\dagger \phi_t) \tag{1.154}$$

where ϕ and ϕ_t are now operators. \square

- **Typicality in time (Challenge).** Be H a non-degenerate Hamiltonian with also non-degenerate gaps. Define $\omega = D_H \phi$. Show that

$$\overline{D(\phi_{St}, \omega_S)} \leq \frac{1}{2} \sqrt{d_S^2 \text{tr } \omega^2}. \tag{1.155}$$

Proof. We'll wait until the next lecture to do this. For now we don't know what the subscript S stands for. \square

11. Average in \mathcal{H} .

- **Haar-average Loschmidt Echo.** Show that

$$\langle \bar{\mathcal{L}}_t \rangle_\phi = \frac{2}{d+1}. \tag{1.156}$$

Proof. Well in lecture we have shown that

$$\begin{aligned}
\langle \bar{\mathcal{L}}_t \rangle_\phi &= \text{tr} \left[T^{(2)} D_{\mathcal{H}}^{(2)} \frac{\mathcal{I}^{\otimes 2} + T^{(2)}}{d(d+1)} \right] \\
&= \frac{1}{d(d+1)} \left[\text{tr} \left[T^{(2)} \left[D_{\mathcal{H}}^{\otimes 2} \mathcal{I}^{\otimes 2} + \mathcal{D}_{\mathcal{H}}^{\otimes 2} T^{(2)} \right] \right] \right].
\end{aligned} \tag{1.157}$$

Now,

$$D_{\mathcal{H}}^{\otimes 2} \mathcal{I}^{\otimes 2} = \sum_k P_k \mathcal{I} P_k = \sum_k P_k P_k = \sum_k P_k = \mathcal{I}^{\otimes 2}. \quad (1.158)$$

And,

$$\begin{aligned} D_{\mathcal{H}}^{\otimes 2} T^{(2)} &= \sum_{n,m} |nm\rangle \langle nm| T^{(2)} |nm\rangle \langle nm| \\ &= \sum_{n,m} |nm\rangle \langle nm| mn \rangle \langle nm| \\ &= \sum_n |nn\rangle \langle nn| \\ \implies \text{tr}(D_{\mathcal{H}}^{\otimes 2} T^{(2)}) &= d. \end{aligned} \quad (1.159)$$

With these, we can continue:

$$\begin{aligned} \langle \bar{\mathcal{L}}_t \rangle_\phi &= \text{tr} \left[T^{(2)} D_{\mathcal{H}}^{(2)} \frac{\mathcal{I}^{\otimes 2} + T^{(2)}}{d(d+1)} \right] \\ &= \frac{1}{d(d+1)} \left[\text{tr} \left[T^{(2)} \left[D_{\mathcal{H}}^{\otimes 2} \mathcal{I}^{\otimes 2} + D_{\mathcal{H}}^{\otimes 2} T^{(2)} \right] \right] \right] \\ &= \frac{1}{d(d+1)} \left[\text{tr} [T^{(2)}] + d \right] \\ &= \frac{2}{d(d+1)} d \\ &= \frac{2}{d+1} \end{aligned} \quad (1.160)$$

where we have used the fact that the original Hilbert space is d -dimensional. Using an identity we saw earlier, we can show that $\text{tr}(T^{(2)}) = d$. With this we're done. \square

- **Markov Inequality.** Use Markov inequality to upper bound

$$\langle \Pr(|A_t - \bar{A}| > \epsilon) \rangle_\phi \quad (1.161)$$

Proof. to be continued... \square

- **Haar-average purity.** Show that

$$\langle \text{tr} \phi_A^2 \rangle_\phi = \frac{d_A + d_B}{d_A d_B + 1}. \quad (1.162)$$

Proof. This is a generalization of the first item. To avoid running into trouble with $\text{tr}^{\otimes 2}$, we use the identity

$$\phi_A^2 = \text{tr}_A(T^{(2)} \phi_A^{\otimes 2}) = \text{tr} T^{(2)} \otimes \mathcal{I}_B \phi^{\otimes 2}. \quad (1.163)$$

With this,

$$\begin{aligned}
\langle \text{tr} \text{tr} T^{(2)} \otimes \mathcal{I}_B \phi^{\otimes 2} \rangle_\phi &= \text{tr} \text{tr} \langle T^{(2)} \otimes \mathcal{I}_B \phi^{\otimes 2} \rangle_\phi \\
&= \text{tr} \text{tr} T^{(2)} \otimes \mathcal{I}_B \langle \phi^{\otimes 2} \rangle_\phi \\
&= \text{tr} \text{tr} T^{(2)} \otimes \mathcal{I}_B \frac{\mathcal{I}_{A \otimes B} + T^{(2)}}{d_A d_B (d_A d_B + 1)} \\
&= \frac{d_A d_B (d_A + d_B)}{d_A d_B (d_A d_B + 1)} \\
&= \frac{d_A + d_B}{d_A d_B + 1}.
\end{aligned} \tag{1.164}$$

□

- **Purity of the dephased state.** Let $\omega = D_{\mathcal{H}} \phi$. Show that

$$\langle \text{tr} \omega^2 \rangle_\phi < \frac{2}{d}. \tag{1.165}$$

Proof. Well we have that $\omega = \bar{\phi} = D_{\mathcal{H}} \phi$. As we have seen, $\text{tr} \omega^2 = \text{tr} \bar{\phi}^2 = \langle \bar{\mathcal{L}}_t \rangle_\phi$ with the property

$$\langle \bar{\mathcal{L}}_t \rangle_\phi = \frac{2}{d+1} < \frac{2}{d}. \tag{1.166}$$

□

- **Typicality of canonical state (challenge).** Define $\omega = D_{\mathcal{H}} \phi$. Also define $\Omega_S \equiv \langle \text{tr}_B \Omega \rangle_\phi = \langle \omega_S \rangle_\phi$. Be $d(\cdot, \cdot)$ the trace distance. Show that

$$\langle d \left(\omega_S, \frac{\mathcal{I}_S}{d_S} \right) \rangle_\phi \leq \frac{1}{2} \sqrt{\frac{1}{d_B}} \tag{1.167}$$

Proof. to be continued... □

- **General canonical state.** Consider a subspace $\mathcal{H}_R \subset \mathcal{H}$ and let Π_R the projector on it. If we take the Haar average on \mathcal{H}_R one has

$$\langle \phi^{\otimes 2} \rangle_\phi = \frac{\Pi_R^{\otimes 2} (\mathcal{I}^{\otimes 2} + T^{(2)})}{d_R (d_R + 1)} \tag{1.168}$$

Then show that, in this subspace, the bound of the previous exercise holds as

$$\langle d \left(\omega_S, \frac{\Omega_S}{d_S} \right) \rangle_\phi \leq \frac{1}{2} \sqrt{\frac{d_S}{d_R}} \tag{1.169}$$

Also show that, in the same subspace, the bound on the Haar-average purity reads

$$\langle \text{tr} \omega^2 \rangle_\phi < \frac{2}{d_R}. \tag{1.170}$$

Part 2

Numerical Methods & Condensed Matter Physics

Instructor: Aaron Szasz

This course has two main goals: (1) to introduce some key models from condensed matter physics; and (2) to introduce some numerical approaches to studying these (and other) models. As a precursor to these objectives, we will carefully understand many-body states and operators from the perspective of condensed matter theory. (However, I will cover only spin models. We will not discuss or use second quantization.)

Once this background is established, we will study the method of exact diagonalization and write simple python programs to find ground states, correlation functions, energy gaps, and other properties of the transverse-field Ising model and XXZ model. We will also discuss the computational limitations of exact diagonalization. Finally, I will introduce the concept of matrix product states, and we will see how these can be used with algorithms such as the density matrix renormalization group (DMRG) to study ground state properties for much larger systems than can be studied with exact diagonalization.

2.1 Introduction

What is *condensed matter*? This is the study of many particles interacting with each other. This could be studying crystals, liquid, etc. These are all condensed matter. We will be studying this using computer programs.

The goals of this course are

- Learn some numerical methods.
- To learn important models in condensed-matter physics.

46 PART 2. NUMERICAL METHODS & CONDENSED MATTER PHYSICS

In particular, we will be studying the quantum Ising model, or equivalently the transverse field Ising model (TFIM) given by the Hamiltonian

$$\mathcal{H} = -J \sum_i S_i^z S_{i+1}^z - h \sum_i S_i^x \quad (2.1)$$

We will first cover some background material required for the rest of the course. Then, we will look at **exact diagonalization** and use it to solve the Ising model. On the third lecture, we will continue with *exact diagonalization*, and depending on how much time we have we will study the *XXZ* model. On the fourth lecture, we will look at **matrix product states (MPS)**. On the fifth day, we will use MPS to solve the Ising and XXZ models more efficiently than using exact diagonalization.

2.2 The Ising model: Motivation

The Ising model is important because it models magnetism and features quantum phase transition, which is an important topic we will cover later.

Consider a one-dimensional chain of N atoms. Each atom has one electron, and that the electrons cannot move due to Coulomb repulsion. So, we have 1 fixed electron per site. For each atom, we solve a local SE and get some set of orbitals.

We will assume that each electron is in the p^z orbital and forget about it.

We note that this description allows for only one degree of freedom, which is the spin of our electron. Two neighboring electrons can have the same or opposite spins. Our model is given by the Hamiltonian

$$\mathcal{H} = -J \sum_i S_i^z S_{i+1}^z \quad (2.2)$$

the product $S_i S_{i+1}$ is 1 if the electrons are parallel, and -1 if antiparallel. When $J > 0$, the model favors all electrons pointing in the same direction. Now, if the model only has an external field

$$-h \sum_i S_i^x \quad (2.3)$$

which favors all electrons point in the direction of the external field $+x$.

The goal is, given the Hamiltonian:

$$\mathcal{H} = -J \sum_i S_i^z S_{i+1}^z - h \sum_i S_i^x \quad (2.4)$$

What is the most favorable configuration of the electron?

2.3 Introduction to many-particle states and operators

2.3.1 Spin 1/2 Systems

This appears in many introductory quantum mechanics courses, but we'll review this anyway.

In general, a state of a spin-1/2 system can be written as a linear superposition of two spin states

$$|\psi\rangle = a |\uparrow\rangle + b |\downarrow\rangle \quad (2.5)$$

where $|\psi\rangle$ is normalized and $|\uparrow\rangle, |\downarrow\rangle$ form an ONB. This means that $\langle \uparrow\downarrow | \uparrow\downarrow \rangle = 1$ and $\langle \uparrow | \downarrow \rangle = 0$.

2.3.2 Operators on Spin-1/2 Systems

An \mathcal{O} is specified by its action on the basis state:

$$\mathcal{O}|\uparrow\rangle = \mathcal{O}_{11}|\uparrow\rangle + \mathcal{O}_{12}|\downarrow\rangle, \quad \mathcal{O}|\downarrow\rangle = \mathcal{O}_{21}|\uparrow\rangle + \mathcal{O}_{22}|\downarrow\rangle. \quad (2.6)$$

To write in a matrix representation, we will assign $|\uparrow\rangle = (1, 0)^\top$ and $|\downarrow\rangle = (0, 1)^\top$. With this, we can find operators as well as other states in this basis.

An aside: Consider a 2×2 matrix $\mathcal{O} = [\mathcal{O}_{ij}]$. Note that $\mathcal{O}(1, 0)^\top$ returns the first column of \mathcal{O} , and so on. This means any matrix \mathcal{O} can be written as the output of \mathcal{O} when fed the standard basis states.

Suppose we want to find the first column of \mathcal{O} . From the definition above, we can easily see that

$$\mathcal{O} = \begin{bmatrix} \mathcal{O}_{11} & \mathcal{O}_{12} \\ \mathcal{O}_{21} & \mathcal{O}_{22} \end{bmatrix} \quad (2.7)$$

We care a lot about Hermitian operators: $\mathcal{O} = \mathcal{O}^\dagger$. For 2×2 matrices, any Hermitian operator can be written as

$$\mathcal{O} = a\mathcal{I} + b\sigma^z + c\sigma^x + d\sigma^y \quad (2.8)$$

where σ^i are the Pauli matrices. Now, we want to write this in the language of states and operators. Recall that

$$\sigma^z|\uparrow\rangle = |\uparrow\rangle, \quad \sigma^z|\downarrow\rangle = -|\downarrow\rangle \quad (2.9)$$

and so on: σ^x flips the arrow, and σ^y flips and multiplies $|\uparrow\rangle$ by i and $|\downarrow\rangle$ by $-i$.

Now, matrices can appear different under different bases. Consider the basis given by

$$|+\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle + |\downarrow\rangle) = (1, 0)^\top, \quad |-\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle - |\downarrow\rangle) = (0, 1)^\top \quad (2.10)$$

This is a new ONB, by inspection.

What do the Pauli matrices become under this change of basis? Well, the first column of the new σ^z is given by

$$\sigma^z \begin{bmatrix} 1 \\ 0 \end{bmatrix}_* = \frac{1}{\sqrt{2}}\sigma^z(|\uparrow\rangle + |\downarrow\rangle) = \frac{1}{\sqrt{2}}(|\uparrow\rangle - |\downarrow\rangle) = |-\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}_* \quad (2.11)$$

Similarly, we can show that $\sigma^z|\pm\rangle = |\mp\rangle$. In this new basis, $\sigma_{new}^z = \sigma^x$. We can also find the form for the other matrices, but let's not worry about that. The gist is that $|\uparrow\downarrow\rangle$ are the z -eigenstates, and $|\leftarrow\rightarrow\rangle$ are the x -eigenstates.

How the matrices change can be found via the usual change of basis that we all have seen in linear algebra.

2.4 Tensor Product Spaces

Suppose there are two space of spin-1/2 particles. The question we want to answer is: What are the allowed states of the 2-spin system? The answer is this: any state in the “tensor product space,” which is the vector space whose basis is the following four vectors: $\{|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle\}$. Each basis vector is an ordered pair, the first being a basis vector in A , the second in B .

So, a general state in the tensor product space is just some superposition of these vectors. What is the inner product in this new space? Suppose we have

$$|\psi_A\rangle \otimes |\psi_B\rangle, \quad |\phi_A\rangle \otimes |\phi_B\rangle. \quad (2.12)$$

Then the inner product between these is defined as

$$\langle \langle \psi_A | \otimes \langle \psi_B | \rangle (|\phi_A\rangle \otimes |\phi_B\rangle) = \langle \psi_A | \phi_A \rangle \cdot \langle \psi_B | \phi_B \rangle. \quad (2.13)$$

One important consequence is that if the A part is orthogonal or the B part is orthogonal then the inner product in the tensor space is also zero (so, also orthogonal). This is quite easy to see.

Finally, we want to know how to apply operators. And operator on the tensor product space looks like

$$\mathcal{O} = \sum_i \mathcal{O}_i^A \otimes \mathcal{O}_i^B. \quad (2.14)$$

The way each component of \mathcal{O} acts is the following:

$$(\mathcal{O}^A \otimes \mathcal{O}^B)(|\psi_A\rangle \otimes |\psi_B\rangle) = (\mathcal{A}|\psi_A\rangle) \otimes (\mathcal{B}|\psi_B\rangle). \quad (2.15)$$

By linearity, we can work out how \mathcal{O} acts in general.

From this definition, we can see some useful facts. Suppose we want to multiply two operators:

$$(\mathcal{A} \otimes \mathcal{O}^B)(\tilde{\mathcal{O}}^A \otimes \tilde{\mathcal{O}}^B)(|\psi_A\rangle \otimes |\psi_B\rangle) = \dots = (\mathcal{O}^A \tilde{\mathcal{O}}^A \otimes \mathcal{O}^B \tilde{\mathcal{O}}^B)(\psi_A \psi_B). \quad (2.16)$$

So, we have

$$(O^A \otimes O^B)(H^A \otimes H^B) = O^A H^A \otimes O^B H^B. \quad (2.17)$$

Next, suppose we have $O^A \otimes I^B$ and $I^A \otimes O^B$. By looking at the commutation, we can easily check that these commute. This fact can be summarized as the following: “Operators on different sites commute.”

2.4.1 Born Rule for Probabilities

Now we want to look at how to get probabilities. Suppose

$$|\psi\rangle = a|00\rangle + b|01\rangle + c|10\rangle + d|00\rangle. \quad (2.18)$$

Then $P(00) = |a|^2$, $P(01) = |b|^2$. Then $P(0_A) = |a|^2 + |b|^2$. Finally, suppose we want to find expectation value:

$$\langle\sigma_A^z\rangle = \langle\sigma_A^z \otimes \mathcal{I}_B\rangle = 1 \cdot P(0_A) + (-1)P(1_A) = |a|^2 + |b|^2 - |c|^2 - |d|^2. \quad (2.19)$$

With this, in the state $|\psi\rangle$, we're going to calculate $\langle\sigma_A^z \otimes \mathcal{I}_B\rangle$ directly:

$$\langle \dots | \sigma_A^z \otimes \mathcal{I}_B | \dots \rangle = \dots = |a|^2 + |b|^2 - |c|^2 - |d|^2 \quad (2.20)$$

as expected.

2.4.2 Expectation Values

Another useful way to compute expectation values is to use matrix representations. First, we assign basis vectors. Then find matrix for $\sigma_A^z \otimes \mathcal{I}_B$. Then find matrix/vector representation of $|\psi\rangle$. Finally, multiply out.

At this point, we will frequently use the short hand $|\psi_A\rangle |\phi_B\rangle = |\psi\phi\rangle$. With this, we introduce the representation:

$$\begin{aligned} |\uparrow\uparrow\rangle &= (1, 0, 0, 0), & |\downarrow\uparrow\rangle &= (0, 0, 1, 0) \\ |\uparrow\downarrow\rangle &= (0, 1, 0, 0), & |\downarrow\downarrow\rangle &= (0, 0, 0, 1) \end{aligned} \quad (2.21)$$

We want to find the matrix for $\sigma_A^z \otimes \mathcal{I}_B$. To do this, we find column-by-column by letting it act on the standard basis vectors. After this is done, we find that

$$(\sigma_A^z \otimes \mathcal{I}_B) = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & -1 & \\ & & & -1 \end{bmatrix} \quad (2.22)$$

Finally, we can verify that

$$\langle\psi| \sigma_A^z \otimes \mathcal{I}_B |\psi\rangle = |a|^2 + |b|^2 - |c|^2 - |d|^2. \quad (2.23)$$

Example 2.4.1. We will look at the 4×4 matrices in this basis for the following operators:

$$\sigma_A^y \otimes \sigma_B^z = \begin{bmatrix} & -i & & i \\ & i & & -i \\ & & & \\ & & & \end{bmatrix} \quad (2.24)$$

$$\sigma_A^z \otimes \sigma_B^y = \begin{bmatrix} & -i \\ i & & \\ & & i \\ & -i & \end{bmatrix} \quad (2.25)$$

A shortcut to do these calculations is of course the Kronecker product:

$$\sigma_A^y \otimes \sigma_B^z = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} & -i & i \\ & & \\ i & & \\ & -i & \end{bmatrix} \quad (2.26)$$

$$\sigma_A^z \otimes \sigma_B^y = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \otimes \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = \begin{bmatrix} & -i & i \\ & & \\ i & & \\ & -i & \end{bmatrix} \quad (2.27)$$

In general,

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B \\ a_{21}B & a_{22}B \end{bmatrix}. \quad (2.28)$$

This is always true if the choice of basis is correct. They are correct in the sense that we can construct the tensor product basis by taking Kronecker products of the basis vectors of the factor space.

The moral of the story is this. The matrix for $\mathcal{O}_A \otimes \mathcal{O}_B$ is the Kronecker product of matrices for \mathcal{O}_A and \mathcal{O}_B .

2.5 The Two-site Ising Model

Recall the Hamiltonian for the model:

$$\mathcal{H} = -J \sum_i S_i^z S_{i+1}^z - h \sum_i S_i^x. \quad (2.29)$$

where

$$S_i^z = \frac{\hbar}{2} \sigma_i^z. \quad (2.30)$$

We can rewrite the model as

$$\mathcal{H} = -J \left(\frac{\hbar}{2} \right)^2 \sim_i \sigma_i^z \sigma_{i+1}^z - h \frac{\hbar}{2} \sum_i \sigma_i^x \quad (2.31)$$

Let the constants absorb $\hbar/2, \hbar^2/4$ respectively, then the model becomes

$$\mathcal{H} = -J \sum_i \sigma_i^z \sigma_{i+1}^z - h \sum_i \sigma_i^x. \quad (2.32)$$

Next, we want to make this precise in the sense of tensor products. For the two-site model, we have precisely

$$\mathcal{H} = -J \sigma_0^z \otimes \sigma_1^z - h(\sigma_0^x \otimes \mathcal{I}_1 + \mathcal{I}_0 \otimes \sigma_1^x). \quad (2.33)$$

We can now convert this into a 4×4 matrix, in the programming exercises.

Before any programming though, let's do this by hand. First, let us find the matrix for $\sigma_0^z \otimes \sigma_1^z$. Once we do that, we will find the eigenstates and the eigenvalues of H when $h = 0$. Furthermore, we will show that in each of the eigenstate, $\langle \sigma_0^x \otimes \mathcal{I}_1 + \mathcal{I}_0 \otimes \sigma_1^x \rangle = 0$. Next, we will do the opposite find the matrix for $\sigma_0^x \otimes \mathcal{I}_1 + \mathcal{I}_0 \otimes \sigma_1^x$. Find the eigenvalues and eigenvectors when $J = 0$. Then, we will show that in each of these states, $\langle \sigma_0^z \otimes \sigma_1^z \rangle = 0$. Then, we will find the 4×4 matrix for \mathcal{H} .

Well,

$$\sigma^z \otimes \sigma^z = \text{diag}(1, -1, -1, 1). \quad (2.34)$$

Obviously the eigenstates are exactly the vectors in the ZZ basis: $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ with eigenvalues $1, -1, -1, 1$ respectively. The corresponding eigenvalues in the Hamiltonian are $-J, J, J, -J$.

Next, to see that $\langle \sigma_0^x \otimes \mathcal{I}_1 + \mathcal{I}_0 \otimes \sigma_1^x \rangle = 0$ for each eigenvector, just use the fact that $\sigma^z |\uparrow\rangle = |\downarrow\rangle$ and $\sigma^x |\downarrow\rangle = |\uparrow\rangle$ and that the inner product $\langle \uparrow|\downarrow \rangle = 0$. This part is easy.

Next,

$$\sigma_0^x \otimes \mathcal{I}_1 + \mathcal{I}_0 \otimes \sigma_1^x = \begin{bmatrix} \sigma^x & \mathcal{I} \\ \mathcal{I} & \sigma^x \end{bmatrix}. \quad (2.35)$$

The eigenvectors in this case will be $\{|++, |-+\rangle, |+-\rangle, |--\rangle\}$. The eigenvalues are $\{1, 0, 0, -1\}$ respectively. These correspond with the eigenvalues $-2h, 0, 0, 2h$ in the Hamiltonian. Verifying the expectation is easy, so again I won't do that.

Finally, the 4×4 matrix for H is given by

$$\mathcal{H} = -J \begin{bmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & 1 \end{bmatrix} - h \begin{bmatrix} 1 & 1 & 1 & \\ 1 & 1 & 1 & \\ 1 & 1 & 1 & \\ 1 & 1 & 1 & \end{bmatrix} = \begin{bmatrix} -J & -h & -h & \\ -h & J & -h & \\ -h & -h & J & -h \\ -h & -h & -h & -J \end{bmatrix} \quad (2.36)$$

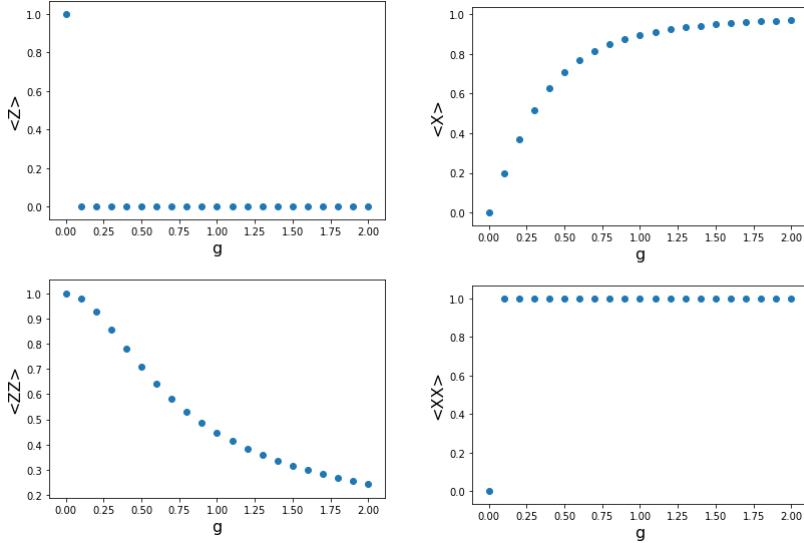
We will solve this and n -site Ising models next. First, we want to find the eigenstate energies for any J, h and some properties of the ground state (GS). To do this, we use programming. To find properties of the ground states, we want to find expectation values. For example, we want to find $\langle \sigma_0^z + \sigma_1^z \rangle$, which is the tendency of spins to point along z in a particular direction. We can also look at $\langle \sigma_0^z \otimes \sigma_1^z \rangle$ which is the tendency of the two spins to point parallel along z .

Definition 2.5.1 (Correlation Function). The quantity $\langle \mathcal{O}_A \mathcal{B} \rangle$ is called a correlation function.

Before we start, we note that \mathcal{H} has two parameters J and h . But by the form of \mathcal{H} we can scale things to that it only has one parameter $g = h/J$, which is the relative strength factor. This is nice if we want to solve the model for every parameter point. Varying one parameter is much better than two. With this, we will solve

$$\mathcal{H} = -\sigma_0^z \otimes \sigma_1^z - g (\sigma_0^x \otimes \mathcal{I}_1 + \mathcal{I}_0 \otimes \sigma_1^x). \quad (2.37)$$

Here's what we get:



We see jumps in $\langle Z \rangle = \langle \sigma_0^z + \sigma_1^z \rangle / 2$ and $\langle \sigma^x \otimes \sigma^x \rangle$. Why is that? Let's start with the first one. We ask two questions: why is it zero and why does it jump from 1 to 0? When $g = 0$, there are two possible ground states $\uparrow\uparrow$ and $\downarrow\downarrow$. This means the ground state looks like $|\psi\rangle = a |\uparrow\uparrow\rangle + b |\downarrow\downarrow\rangle$. The expectation value $\langle \sigma_0^z + \sigma_1^z \rangle / 2$ in this case will be $(|a|^2 - |b|^2)$. This means we'll get a random answer at $g = 0$. Python randomly picks a, b when it does the numerical calculation for the eigenvalues. This means the $g = 0$ value is meaningless. Now, when $g \neq 0$, then because the eigenstates respect the symmetries of the Hamiltonian,

when we flip all spins, the Hamiltonian doesn't change. This means the ground state has $|\uparrow\rangle$ and $|\downarrow\rangle$ parts that are equal. In comparison, $\langle \sigma^z \otimes \sigma^z \rangle$ is a better measure of the tendency of spins to point along z .

The case with x occurs for the same reason. $\sigma^x \otimes \sigma^x$ has $z - y$ symmetry where we get a degeneracy in the eigenvalues. But we also note that $\langle \sigma^x \otimes \sigma^x \rangle$ is constant, which means $[\mathcal{H}, \sigma^x \otimes \sigma^x] = 0$, which says that $\sigma^x \otimes \sigma^x$ is a conserved quantity.

2.6 The N -site Ising Model:

The Hamiltonian is given by

$$\mathcal{H}/J = - \sum_i \sigma_i^z \sigma_{i+1}^z - g \sum_i \sigma_i^x. \quad (2.38)$$

We want to solve this model. To do this, we need to write the matrix for the Hamiltonian. We follow the steps below:

- Find a basis.
- Find a matrix for \mathcal{H}
- Find the ground state
- Calculate expectation values, especially $\langle \sigma_0^z \otimes \sigma_{N-1}^z \rangle$.

With N sites, we have 2^N basis states. Let's pick the standard basis states as the basis for the space, where $(1, 0, \dots, 0)^\top = |\uparrow \dots \uparrow\rangle = |0\rangle_{2^N-1}$ and so on. Now, we need a function to turn the index $(0, 1, 2, \dots)$ into the list of spins. Suppose that $N = 3$. Then we can use binary arithmetics to get back the list of spins.

Now how do we find the matrix \mathcal{H} ? There are a few methods to do this:

- **Kronecker product:** Suppose we want to do the following $\sigma_0^x \otimes \sigma_1^z \otimes \sigma_2^x$. Then this looks like

$$\begin{bmatrix} 0 & \sigma^x \\ \sigma^x & 0 \end{bmatrix} \otimes \sigma_2^x = \begin{bmatrix} 0 & \dots \\ \dots & 0 \end{bmatrix} \quad (2.39)$$

which is an 8×8 . In practice, we start out by making a list of operators the looks like

$$[\sigma^z \sigma^z, \mathcal{I}, \mathcal{I}, \dots, \mathcal{I}] = [\sigma^z, \sigma^z] + [\mathcal{I}] * (N - 2) \quad (2.40)$$

The second step will be repeated Kronecker product. We initiate a matrix as the first element in the list of operators. Then for each operator in the list, the new matrix is the old matrix Kronecker-ing the next one.

- **Column-by-Column approach:** In this method, we take the index of the basis states and turn it into a spin list, apply some operation on it to get a spin list with some coefficient, then get back the vector representation.
 - The first column of the matrix is the operators applied to the first basis state. To do this, we convert the basis vector into a list of spins. Let the operator act on this list of spins to get a new list of spins. Then convert the list of spins to numbers again.
- **Use some knowledge specific to operators.** This will be the fastest way to get the Hamiltonian. In particular we know that the ZZ term is diagonal, which the diagonal elements being the expectation values of $\langle \sum_i \sigma_i^z \sigma_{i+1}^z \rangle$ in each basis state. To construct the matrix, we convert the column index to a list of spins, and calculate and put the expectations on the diagonal.

2.7 Matrix product states

Not all quantum states are product states, i.e., not all states can be written as a tensor product of two states. These states are called entangled states.

Product states are useful. Consider a model with N sites. $|\psi\rangle = \bigotimes |\psi\rangle$. Then

$$\begin{aligned}\langle \sigma_i^z \rangle &= \langle \psi_0 | \mathcal{I} | \psi_0 \rangle \cdot \dots \langle \phi_i | \sigma_i^z | \phi_i \rangle \cdot \dots \langle \psi_{N-1} | \mathcal{I} | \psi_{N-1} \rangle \\ &= \langle \psi_i | \sigma_i^z | \phi_i \rangle.\end{aligned}\tag{2.41}$$

We just reduced the problem from 2^N coefficients to 2 coefficients.

The idea of matrix product states is to write entangled states in a way that lets us do these same efficient calculations. For example, when $N = 2$,

$$(a_0 |0\rangle + b_0 |1\rangle) \otimes (a_1 |0\rangle + b_1 |1\rangle) = (A_0 |0\rangle + B_0 |1\rangle) \otimes (A_1 |0\rangle + B_1 |1\rangle)\tag{2.42}$$

where now the “coefficients” are not matrices. With this formulation, any state can be written as matrix product state. Note that these matrices aren’t necessarily squares. They can be row/column matrices.

Part 3

Path Integrals

Instructor: Dan Wohns

The goal of this course is to introduce the path integral formulation of quantum mechanics and a few of its applications. We will begin by motivating the path integral formulation and explaining its connections to other formulations of quantum mechanics and its relation to classical mechanics. We will then explore some applications of path integrals.

3.1 Introduction to path integrals and the semi-classical limit

Path integrals is a different yet equivalent formulation of quantum mechanics. The path integral formulation of quantum mechanics does make a connection between classical and quantum physics more clear. It also has a number of advantages, some which we will touch on later. In essence, path integrals can allow for perturbation approaches to solving problems in quantum mechanics. Path integrals also generalize quantum field theory. Solving SE is much more difficult in higher dimensions. Using a differential-equation approach to QFT is often hard. Path integrals allow for easier approaches to the same problems in QFT.

3.1.1 Motivation & Definition

Consider the double slit experiment.

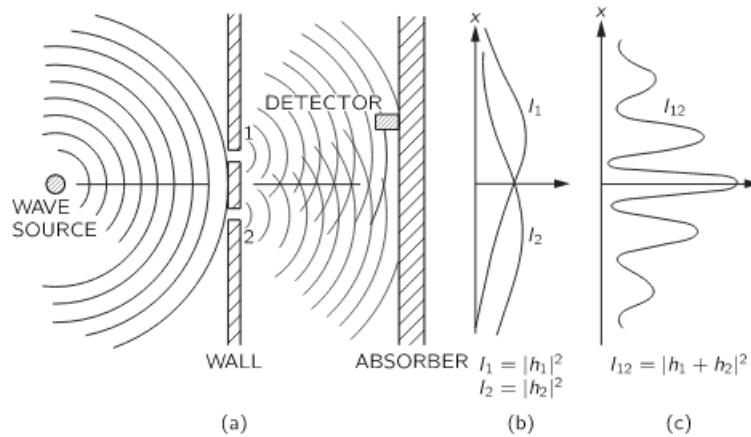


Fig. 1–2. Interference experiment with water waves.

Figure 3.1: From Feynman's lectures of Physics

We know that the amplitude/propagator $K = A_1 + A_2$ leads to interference.

The basic idea behind the path integral comes from considering what happens when there are many many many more slits. This suggests that the propagator should be the sum over all paths of some functional $A[\gamma]$ that is the amplitude associated with each path:

$$K = \sum_{\gamma} A[\gamma] \quad (3.1)$$

where

$$K(q_f, t_f, q_i, t_i) = \langle q_f | e^{i\mathcal{H}(t_f - t_i)/\hbar} | q_i \rangle \quad (3.2)$$

is the propagator.

When there are infinitely many holds of infinitesimal size, the particle can be anywhere (since there's essentially no more holes), so we have the picture In

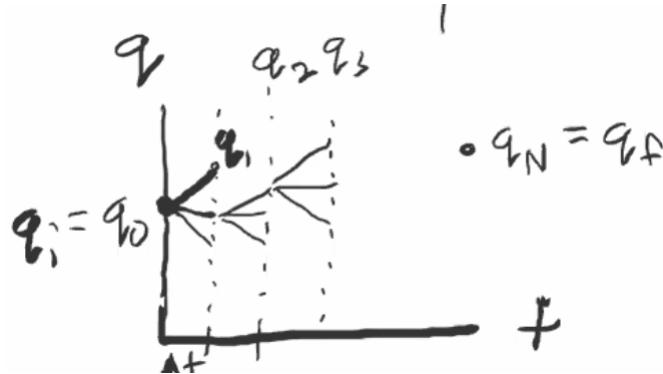


Figure 3.2: From PI lecture by Dan Wohns

this case, the sum over paths becomes an integral

$$\sum_{\gamma} \rightarrow \lim_{N \rightarrow \infty} \int dq_1 dq_2 \dots dq_{N-1} \quad (3.3)$$

3.1.2 Some properties

Proposition 3.1.1 (Composition). Consider the path γ_{12} composed of the path γ_1 followed by the path γ_2 . Then

$$A[\gamma_{12}] = A[\gamma_1]A[\gamma_2]. \quad (3.4)$$

This implies (as we will see in the exercises) that

$$K(q_f, t_f, q_i, t_i) = \int dy K(q_f, t_f, y, t_{int})K(y, t_{int}, q_i, t_i). \quad (3.5)$$

Proposition 3.1.2 (Classical Limit). When $\hbar \rightarrow 0$ and the action $S \gg \hbar$, then the classical path becomes most important.

For most of the course, we will be consider actions that look like

$$S = \int_{t_i}^{t_f} dt \frac{1}{2} m \dot{q}^2(t) - V(q(t)). \quad (3.6)$$

The classical path satisfies the Euler-Lagrange equation:

$$\frac{\delta S}{\delta q(t)} = 0. \quad (3.7)$$

This turns out to be

$$m\ddot{q} + V'(q) = 0 \quad (3.8)$$

which is just Newton's second law of motion.

With these, we have

$$\int \mathcal{D}q(t)e^{iS[q(t)]/\hbar}. \quad (3.9)$$

The composition property suggests that we should have an exponential, so that we can convert multiplication to summation. This says that all paths contribution with equal weight. So, the classical path is as important as the quantum path in the action. However, only paths such that $\delta S = 0$ contribute more to the propagator.

Consider the paths $q(t)$ and $q'(t) = q(t) + \eta(t)$ where $\eta(t)$ is some small perturbation, then

$$S[q'(t)] = S[q(t)] + \int dt \eta(t) \frac{\delta S[q]}{\delta q(t)} + \mathcal{O}(\eta^2). \quad (3.10)$$

What happens in the general case is that

$$e^{iS[q]/\hbar} + e^{iS[q']/\hbar} \approx e^{iS[q]/\hbar} \left(1 + \exp \left[\frac{i}{\hbar} \int dt \eta(t) \frac{\delta S[t]}{\delta q(t)} \right] \right) \quad (3.11)$$

So, when $\hbar \rightarrow 0$, we get destructive interference. But if $\delta S/\delta q = 0$, we get constructive interference. So, even though all paths are equally important, it's also true that the classical path are "more important" than quantum paths, because they interfere constructively.

3.2 Propagator in real and imaginary time

Last time, we guessed the form of the propagator. This time, we will derive in some detail the propagator in real and imaginary time.

What we're trying to compute is the propagator:

$$K(q_f, t_f, q_i, t_i) = \int \prod_j^{N-1} dq_j K(q_{j+1}, t_{j+1}, q_j, t_j) \equiv \int \prod_j^{N-1} dq_j K_{q_{j+1}, q_j}. \quad (3.12)$$

using the composition property that we discussed last time. Now this is starting to look like a path integral.

Now, we know that

$$\begin{aligned} K_{q_{j+1}, q_j} &= \langle q_{j+1} | e^{i\mathcal{H}\delta t/\hbar} | q_j \rangle \\ &\approx \langle q_{j+1} | q_j \rangle + \langle q_{j+1} | \frac{i\delta t \mathcal{H}}{\hbar} | q_j \rangle + \dots \\ &= \delta(q_{j+1} - q_j) + \text{Computed in Exercise} \\ &= \int \frac{dp_j}{2\pi\hbar} e^{ip_j(q_{j+1}-q_j)/\hbar} + \text{Computed in Exercises} \end{aligned} \quad (3.13)$$

Putting everything together we get

$$K_{q_{j+1}, q_j} = \int \frac{dp_j}{2\pi\hbar} e^{ip_j(q_{j+1}-q_j)/\hbar} \left(1 - \frac{i\delta t \mathcal{H}}{\hbar} + \dots \right) \quad (3.14)$$

The term on the right hand side looks like a Taylor expansion for $e^{-i\delta\mathcal{H}t/\hbar}$. So, if we assume that $\mathcal{H} = p_j^2/2m + V(q_j)$ then this integral becomes a Gaussian integral

$$\begin{aligned} K_{q_{j+1}, q_j} &= \int \frac{dp_j}{2\pi\hbar} e^{i\delta t/\hbar(p_j \dot{q}_j - p_j^2/2m - V(q_j))} \\ &= \sqrt{\frac{m}{2\pi i\delta t \hbar}} e^{(i\delta t/\hbar)(mq_j^2/2 - V(q_j))}. \end{aligned} \quad (3.15)$$

Gaussian integrals are important in the study of path integrals, because they're among the integrals that we can evaluate exactly.

Staring at this for a bit, we notice that the exponential looks like part of the action. So, by multiplying together all these infinitesimal propagators, we get the entire action:

$$K = \left(\frac{m}{2\pi i\hbar\Delta t} \right)^{N/2} \int \prod_{j=1}^{N-1} dq_j \exp \left[\frac{i\Delta t}{\hbar} \sum_{j=0}^{N-1} \left(\frac{m\dot{q}_j^2}{2} - V(q_j) \right) \right]. \quad (3.16)$$

In the $N \rightarrow \infty$ limit,

$$K = \int \mathfrak{D}[q(t)] e^{iS(t)/\hbar} \quad (3.17)$$

Now we want to make the connection to imaginary time and statistical physics. Recall that

$$K = \int \mathfrak{D}q(t) \exp [iS[q(t)]/\hbar]. \quad (3.18)$$

Now, the i leads to oscillations in the integrand, which leads to questions about convergence of the integral. So, it would be much more convenient if i wasn't there. This means we want to consider **imaginary time**. Why do we want to do this? Mathematically, it's generally easier. More important, if we make the exponential real, we can make connections to statistical physics. So we shall see a connection between quantum mechanics in imaginary time and statistical physics. Thirdly, in quantum mechanics we're interested in the ground states energy of Hamiltonians. Sometimes, it is possible to compute ground state energies using quantum mechanics in imaginary time.

To this end, consider $T = -iT_E$ where E stands for Euclidean. Once we do that, we can computer the Euclidean propagator:

$$K_E = \langle q_f | e^{-T_E \mathcal{H}/\hbar} | q_i \rangle \quad (3.19)$$

with $t \rightarrow -i\tau$. If we do that, then

$$iS[q(t)] = i \int_0^T dt \left(\frac{m}{2} \dot{q}(t)^2 - V(q(t)) \right) \mapsto - \int_0^{T_E} d\tau \left(\frac{m}{2} \dot{q}(\tau)^2 + V(q(\tau)) \right) \quad (3.20)$$

where the time derivative becomes derivative over τ , so the relative sign between the kinetic and potential energy changes. With this, we have the Euclidean action:

$$iS[q(t)] = -S_E[q(\tau)]. \quad (3.21)$$

After this replacement, we can compute the Euclidean propagator:

$$K_E = \int \mathfrak{D}[q(t)] e^{-S_E[q(\tau)]/\hbar}. \quad (3.22)$$

With this we can compare this to the expression for the partition function:

$$\mathcal{Z}[\beta] = \sum_j e^{-\beta E_j} \quad (3.23)$$

by identifying $\beta = T_E/\hbar = 1/k_B T$ where now T is temperature. It turns out that (which we will verify in the exercises)

$$\mathcal{Z} = \int dq K_E(q, 0, q, \beta \hbar). \quad (3.24)$$

Note that the ground state energy E_0 can be found this way:

$$E_0 = \lim_{\beta \rightarrow \infty} \frac{1}{\beta} \log \mathcal{Z}. \quad (3.25)$$

We can relate the propagator to the ground state energy, which is an observable of interest.

3.2.1 Free Path Integrals

Free path integrals are some of the path integrals we can compute exactly. These come from quadratic action S in the variable q . For example,

$$K_E = \langle q_f | e^{-\beta \mathcal{H}} | q_i \rangle, \quad \mathcal{H} = p^2/2m. \quad (3.26)$$

This integral is easy (compared to oscillatory integrals) to compute, because the integral is just Gaussian. It turns out that

$$K_E = \sqrt{\frac{m}{2\pi\beta}} e^{-S[q_c]/\hbar} \quad (3.27)$$

where q_c is the classical trajectory. The way to derive this is via Gaussian integration techniques.

The key idea here is that in this case K_E turns out to have the form with N -coupled Gaussian integrals. To understand this, we only need to understand a single integral. In any case, we can learn a lot of about path integrals by just looking at the case $N = 1$. When $N = 1$, we can actually go through a lot of the details. When N is larger, we don't gain much more knowledge. We will discuss, of course, what happens when $N \rightarrow \infty$.

Next, we will look at perturbation theory, where the action is no longer quadratic. Again, we will look at the case where $N = 1$. We will also see where the path integral and perturbation theory fails. We will, of course, look at how we can address these failures.

3.3 Perturbation theory

Consider the action

$$S_E = \int d\tau \left[\frac{1}{2} m \dot{q}^2(\tau) + \frac{1}{2} m \omega^2 q^2(\tau) + \frac{\lambda}{4!} q^4(\tau) \right] \quad (3.28)$$

We want to compute the propagator:

$$K_E = \int \mathfrak{D}[q(t)] e^{-S_E[q(\tau)]s/\hbar} \quad (3.29)$$

where we're working in Euclidean time. We want to compare this integral to

$$I = \int_{\mathbb{R}} dx e^{-f(x)/\hbar}. \quad (3.30)$$

The analogy we will keep in mind the following

$$K_E \leftrightarrow I, \quad q(\tau) \leftrightarrow x, \quad S_E \leftrightarrow f. \quad (3.31)$$

Suppose that f is some function with a global minimum at zero. If we want to find I , we want to find where $f(x)$ is minimum – in this case it is near zero. Similarly, to approximate K_E , we want to find paths where the action $S[q]$ is smallest. This means we want places where $\delta S/\delta q = 0$. But of course this is just the equation of motion. This means we want to look at paths close the classical path q_C .

The strategy is to Taylor expand f , so that we can write the integral I as

$$I \approx \sum_{c_1, c_2, n} c_1 \int_{\mathbb{R}} dx x^n e^{-C_n x^2} \quad (3.32)$$

which we will show in the exercise.

The trick to evaluate these integrals is to add a source J :

$$\int \mathfrak{D}[q(\tau)] \exp \left[-\frac{1}{\hbar} \int d\tau (S[q] - J(\tau)q(\tau)) \right]. \quad (3.33)$$

The one-dimensional analog is

$$I = \int dx \exp [-x^2 + xJ]. \quad (3.34)$$

Note that

$$\frac{d}{dJ} I = \int dx x \exp [-x^2 + xJ] = \pi \frac{J}{2} e^{J^2/4} \Big|_{J=0} \implies \int dx x \exp (-x^2) = 0. \quad (3.35)$$

In general, this technique works when we evaluate path integrals. With this technique we can construct a systematic way to approximate path integrals. This systemic way is Feynman diagrams.

3.4 Non-perturbative physics and quantum tunneling

Recall that terms like $e^{-A/\lambda}$ are invisible to perturbation theory. In this section, we look at how we fix this problem in non-perturbative physics. In particular, we can look at phenomena like quantum tunneling. In this case, we are interested in the transmission coefficient:

$$T \propto e^{-2/\hbar\sqrt{2m(V-E)} dx} \quad (3.36)$$

This is invisible to perturbation theory. Another example is meta-stable potential. Another example is a potential with two classically degenerate ground states. In this case, the transmission coefficient also has a similar form.

Recall the propagator amplitude:

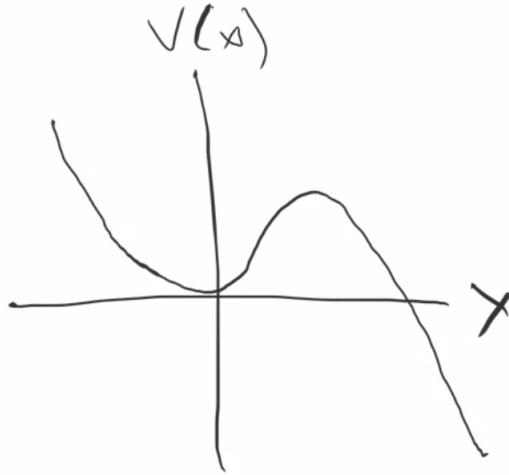
$$K_E = \int \mathfrak{D}[q(\tau)] e^{-S_E[q(\tau)]/\hbar} \quad (3.37)$$

with

$$E_0 = - \lim_{\beta \rightarrow 0} \frac{1}{\beta} \log K_E(q_0, \beta\hbar/2, q_0 - \beta\hbar/2) \quad (3.38)$$

where we have shifted the time, which doesn't change the expression, of course.

What we're interested in is the decay rate for the following metastable potential (with no ground state): While there's no notion of a ground state energy,



we still have **resonant states** which are those with quantized complex energies.

The decay rate is given by

$$\Gamma = 2|\text{Im } E_0| \quad (3.39)$$

which is why the decay rate is related to this setup. We want to compute K_E using path integrals. To do this, we think about the example

$$I = \int_{\mathbb{R}} e^{-f(x)/\hbar} dx. \quad (3.40)$$

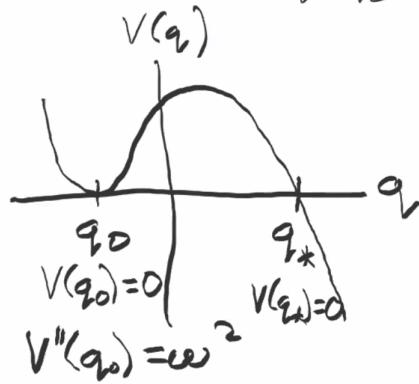
Last time we looked at perturbation theory, which says that $f(x)$ has a global minimum at zero, then we get an expansion. Now, we can think about a slightly more general case where f has several minima. Then expansions around all of the minima are roughly equally important. In this case, we have that

$$\begin{aligned} I &\approx \sum_{x_i} \int dx \exp \left[-\frac{1}{\hbar} (f(x_i) + (1/2)f''(x_i)(x - x_i)^2) \right] \times [1 + \dots] \\ &\approx \sum_{x_i} e^{-f(x_i)/\hbar} (1 + \mathcal{O}(\hbar^\alpha)) \\ &\approx \sum_{x_i} e^{-f(x_i)/\hbar} N[\bar{x}] \end{aligned} \quad (3.41)$$

as we have seen in the exercises. Now let's go back to the path integral. There might be multiple good trajectories, so we should sum over all of them:

$$K_E = \sum_{\bar{q}=q_c} e^{-S_E[q(\tau)]/\hbar} N[\bar{q}]. \quad (3.42)$$

The decay rate is given by



$$\frac{\Gamma}{2} = |\text{Im } E_0|. \quad (3.43)$$

Then

$$E_0 = - \lim_{\beta \rightarrow 0} \frac{1}{\beta} \log K_E(q_0, \beta \hbar/2, q_0 - \beta \hbar/2) \quad (3.44)$$

The equation of motion is

$$-m d_\tau^2 \bar{q}(\tau) + V'(\bar{q}) = 0 \quad (3.45)$$

with the boundary conditions $\bar{q}(\pm T_E/2) = q_0$ with $T_e = \beta \hbar$. We want to look at the limit where $\beta \rightarrow \infty$.

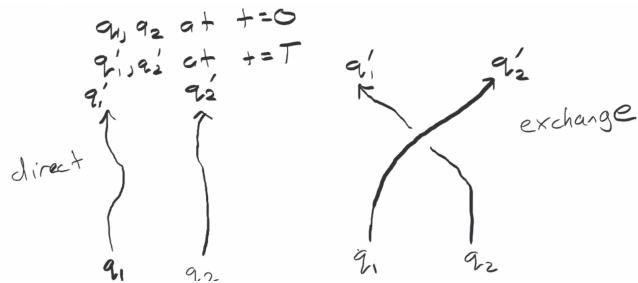
One other thing we care about is that the action should be finite. If the action is infinite than these trajectories don't contribute at all to the propagator.

3.5 Topology and Particle Statistics

Consider an amplitude

$$A = \sum_{\text{paths}} e^{iS[q_1(t), q_2(t)]}. \quad (3.46)$$

for two identical particles to go from q_1, q_2 at $t = 0$ to some positions q'_1, q'_2 at $t = T$. There are two possibilities of this happening: **direct** and **exchange** (since these are indistinguishable particles) Note we assume things happening

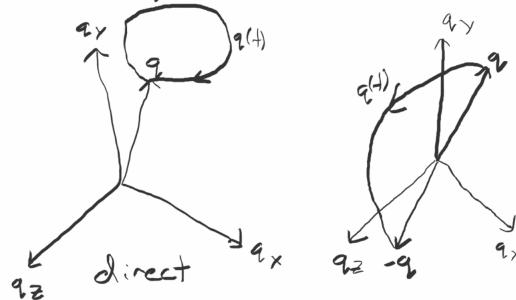


in three dimensions of space.

Now, what do the path integrals look like in both scenarios? Before we do this, let us introduce *relative coordinates*:

$$q(t) = q_2(t) - q_1(t) \quad (3.47)$$

and assume that $q(T) = q(0)$ for the direct path case and $q(T) = -q(0)$ for the exchange case. This space of $q(\cdot)$ is called the **covering space**. The covering



space is basically the three-dimensional Euclidean space. We also assume that particles cannot be at the exact same position at any given time due to some repulsion, so the covering space looks like $\mathbb{R}^3 \setminus \{0\}$. The **configuration space** is half of the covering space: $(\mathbb{R}^3 \setminus \{0\}) \setminus \mathcal{Z}_2$, in which we make the identification $q \leftarrow -q$.

With this, the amplitude in the configuration space $A(q, T, q, 0)$ is

$$\begin{aligned} A(q, T, q, 0) &= \sum_{\text{direct}} e^{iS[q_1, q_2]} + \sum_{\text{exchange}} e^{iS[q_1, q_2]} \\ &= \bar{A}(q, T, q, 0) + \bar{A}(-q, T, q, 0) \end{aligned} \quad (3.48)$$

where q is the relative position and \bar{A} are the amplitudes in the covering space.

We can also write, for $q' \neq q$,

$$A^\phi(q', T, q, 0) = \bar{A}(q', T, q, 0) + e^{i\phi} \bar{A}(q', T, q, 0). \quad (3.49)$$

These amplitudes are well-defined in the covering space. The phase can be there because the SE is linear. We can also write

$$A^\phi(q', T, q, 0) = e^{i\alpha} A^\phi(-q', T, q, 0) \quad (3.50)$$

where the overall phase α and the exchange $q' \leftrightarrow q$ does not change the physics. In the covering space, this becomes

$$\bar{A}(q', T, q, 0) + e^{i\phi} \bar{A}(-q', T, q, 0) = e^{i\alpha} (\bar{A}(-q', T, q, 0) + e^{i\phi} \bar{A}(q', T, q, 0)) \quad (3.51)$$

so we get that

$$1 = e^{i(\alpha+\phi)}, \quad e^{i\phi} = e^{i\alpha}. \quad (3.52)$$

This means that $\phi = \alpha + 2\pi k$ and thus $\phi = \{0, \pi\} + 2\pi k$. When $\phi = 0$, we get Bose-Einstein statistics (phase +1), and if $\phi = \pi$ then we get Fermi-Dirac statistics (phase -1).

3.6 Applications

- Bohm-Aharonov effect.
- Path integrals not only compute propagators but also anything other thing that can be computed in the other formulations of quantum mechanics.
- Using path integrals where the variables are fermionic, i.e., integration variables are grassmann variables. Grassmann variables and grassmann integration is an entirely different field of study.
- Generalizing path integrals over the history of a particle to over the history of a field. This is one way to combine special relativity with quantum mechanics. This is one way to understand quantum field theory.
-

3.7 Exercises

1. Composition.

- In lecture we argued that the propagator K can be written as a sum (or an integral) over paths γ of a functional $A[\gamma]$.

$$K = \sum_{\gamma} A[\gamma]. \quad (3.53)$$

Show that the propagator obeys the composition property

$$K(q_f, t_f, q_i, t_i) = \int dy K(q_f, t_f, y, t_{int}) K(y, t_{int}, q_i, t_i) \quad (3.54)$$

of A has the factorization property

$$A[\gamma_{12}] = A[\gamma_1] A[\gamma_2]. \quad (3.55)$$

Proof. Let A satisfy the factorization property. Then,

$$K = \sum_{\gamma_1, \gamma_2} A[\gamma_1] A[\gamma_2] = \int dy \sum_{\gamma_1} A[\gamma_1] \sum_{\gamma_2} A[\gamma_2] \quad (3.56)$$

where the path γ_1 ends at y and γ_2 starts at y . With this, it's straightforward:

$$\begin{aligned} K &= \int dy \langle q_f | e^{-i\mathcal{H}(t_f - t_{int})/\hbar} | y \rangle \langle y | e^{-i\mathcal{H}(t_{int} - t_i)/\hbar} | q_i \rangle \\ &= \langle q_f | e^{-i\mathcal{H}(t_f - t_{int} + t_{int} - t_i)/\hbar} | q_i \rangle \\ &= K(q_f, t_f, q_i, t_i). \end{aligned} \quad (3.57)$$

So, we're done, because from the first line we conclude

$$K(q_f, t_f, q_i, t_i) = \int dy K(q_f, t_f, y, t_{int}) K(y, t_{int}, q_i, t_i). \quad (3.58)$$

□

- Explain why the factorization property

$$A[\gamma_{12}] = A[\gamma_1] A[\gamma_2] \quad (3.59)$$

holds if $A[\gamma] = e^{iS[\gamma]/\hbar}$ where $S[\gamma]$ is the classical action of the trajectory γ .

Proof. We note that $S[\gamma_{12}] = S[\gamma_1] + S[\gamma_2]$ if γ_{12} is composed of γ_1, γ_2 . So this factorization property follows. □

3. Suppose $A[\gamma] = C_1 e^{C_2 S[\gamma]}$ for some constants C_1, C_2 . Which values of C_1, C_2 give rise to a composition property?

Proof. C_1 must be constantly 1, so that the composition property is satisfied. C_2 can be any real or complex number. \square

2. Classical Limit

1. Suppose $f(x)$ is a real function. Which values of x provide the most important contribution to

- $\int e^{-f(x)} dx$? Why? (Optional: what conditions should be placed on $f(x)$?)

Proof. The most important contribution to this integral is when f is small, so x 's where $f(x)$ is small. \square

- $\int e^{if(x)} dx$? Why? (Optional: what conditions should be placed on $f(x)$?)

Proof. The most important contribution to this integral is when f changes slowly, so most comes from x where $f'(x)$ small. \square

2. The classical equation of motion for a particle with classical action S is that the first-order variation vanishes:

$$\delta S = 0. \quad (3.60)$$

A trajectory with $\delta S = 0$ is a classical trajectory. No extra conditions on the second-order variation of the action is required. Why does $A[\gamma] = e^{iS[\gamma]/\hbar}$ give the correct classical limit?

Proof. Recall that $\delta S = 0$. So this sort of shows why $A[\gamma] = e^{iS[\gamma]/\hbar}$ works. We won't worry too much about the details here. \square

3. What other choices $A[\gamma]$ give rise to the correct classical limit? Do they have a composition property?

4. Do any of the other possibilities you identified in 1(c) have the correct classical limit?

3. **First-Order Term.** By inserting the resolution of identity

$$\mathbf{1} = \int \frac{dp}{2\pi\hbar} |p\rangle \langle p| \quad (3.61)$$

and using

$$\langle q|p\rangle = e^{ipq/\hbar} \quad (3.62)$$

show that the first-order term can be expressed as

$$-i\frac{\Delta t}{\hbar} \langle q_{j+1} | \mathcal{H} | q_j \rangle = -i\frac{\Delta t}{\hbar} \int \frac{dp_j}{2\pi\hbar} e^{ip_j(q_{j+1}-q_j)/\hbar} \left(\frac{p_j^2}{2m} + V(q_j) \right). \quad (3.63)$$

What changes at second order?

Proof. This looks hard, but isn't hard at all. I won't show the full solution here, but I'll just sketch things out here. By inserting the resolution of identity we get

$$\begin{aligned} & \langle q_{j+1} | \int \frac{d_i}{2\pi\hbar} \left(\frac{p_i^2}{2m} + V(q_i) |p_i\rangle \langle p_i| \right) |q_j \rangle \\ &= \int \frac{dp_i}{2\pi\hbar} e^{ip_i(q_{j+1}-q_j)/\hbar} \left(\frac{p_i^2}{2m} + V(q_i) \right). \end{aligned} \quad (3.64)$$

With this we're essentially done. \square

4. Statistical Physics and Imaginary Time. Show that the partition function

$$\mathcal{Z}[\beta] = \sum_j e^{-\beta E_j} \quad (3.65)$$

can be expressed as a sum over imaginary time propagators

$$\mathcal{Z}[\beta] = \int dq K_E(q, \beta\hbar, q, 0). \quad (3.66)$$

Proof. Again, the resolution of identity is helpful here. Consider the energy eigenbasis, then we have

$$\mathbf{1} = \sum_j |e_i\rangle \langle e_i|. \quad (3.67)$$

With this, and by identifying $-T_E \mathcal{H}/\hbar \equiv \beta \mathcal{H}$ we have

$$\begin{aligned}
\int dq K_E(q, 0, q, \beta \mathcal{H}) &= \int dq \langle q | e^{\beta \mathcal{H}} | q \rangle \\
&= \sum_i \int dq \langle q | e_i \rangle e^{\beta \mathcal{H}} | q \rangle \langle e_i | \\
&= \int dq \sum_i \langle e_i | q \rangle \langle q | e^{-\beta \mathcal{H}} | e_i \rangle \\
&= \sum_i \langle e_i | e^{-\beta \mathcal{H}} | e_i \rangle \\
&= \text{tr}(e^{-\beta \mathcal{H}}) \\
&= \sum_n e^{-\beta E_n}.
\end{aligned} \tag{3.68}$$

where we have used the completion relation in the position space:

$$\int dq |q\rangle \langle q| = \mathbf{1}.$$

Of course, we can also go the other way from \mathcal{Z} and use the completion relation in the energy eigenbasis. Either way, the solutions are equivalent. \square

5. Path Integral Derivation of the Propagator. In the path integral formalism the propagator in imaginary time is

$$K_E(q, \beta \hbar, q, 0) = \lim_{N \rightarrow \infty} \left(\frac{m}{2\pi\Delta\tau\hbar} \right)^{N/2} \int dq_1 \dots dq_{N-1} \exp \left[-\tau\hbar \sum_{j=1}^N \frac{m}{2} \left(\frac{q_j - q_{j-1}}{\Delta\tau\hbar} \right)^2 \right] \tag{3.69}$$

1. Show that the propagator can be rewritten as

$$\begin{aligned}
K_E(q, \beta \hbar, q, 0) &= \lim_{N \rightarrow \infty} \left(\frac{m}{2\pi\Delta\tau\hbar} \right)^{N/2} \left(\frac{2\Delta\tau\hbar}{m} \right)^{\frac{N-1}{2}} \\
&\quad \times \int dy_1 \dots dy_{N-1} \exp \left[- \sum_{j=1}^N (y_j - y_{j-1})^2 \right]
\end{aligned} \tag{3.70}$$

Proof. This is just change of variables, so I won't do that here. \square

2. Use induction to show that

$$\int dy_1 \dots dy_{N-1} \exp \left[- \sum_{j=1}^N (y_j - y_{j-1})^2 \right] = \left(\frac{\pi^{N-1}}{N} \right)^{1/2} e^{-(y_N - y_0)^2/N}. \tag{3.71}$$

Proof. Just use induction and crank. At one point we will have to evaluate a 1-dimensional Gaussian integral, but this shouldn't be a challenge. This one-dimensional Gaussian integral will introduce some factor π^α with other constants such that we get to the right hand side at iteration $N - 1 + 1 = N$. In any case, this problem also looks hard but is quite easy. \square

3. Take the continuum limit $N \rightarrow \infty$ to recover the result from lecture:

$$K_E(q, \beta\hbar, q, 0) = \sqrt{\frac{m}{2\pi\beta}} \exp\left[-\frac{m(q_f - q_i)^2}{2\beta}\right]. \quad (3.72)$$

Proof. Combining things in the previous parts, we get that in the $N \rightarrow \infty$ limit, we have

$$\begin{aligned} K_E &= \lim_{N \rightarrow \infty} \left(\frac{m}{2\pi\Delta\tau\hbar}\right)^{N/2} \left(\frac{2\Delta\tau\hbar}{m}\right)^{\frac{N-1}{2}} \left(\frac{\pi^{N-1}}{N}\right)^{1/2} e^{-(y_N - y_0)^2/N} \\ &= \dots \end{aligned} \quad (3.73)$$

hmm... There might be something wrong here... or might not. In any case, this is just a case of taking the limit correctly and using some change of variables. I won't go into the details. \square

6. Semi-Classical Limit.

1. Let $f(x)$ be a real valued function. By expanding $f(x)$ around its minimal values x_c show that we can approximate

$$I = \int_{\mathbb{R}} dx e^{-f(x)/\hbar} \quad (3.74)$$

by

$$I \approx \sum_{x_c} \sqrt{\frac{2\pi\hbar}{f''(x_c)}} e^{-f(x_c)/\hbar} (1 + \mathcal{O}(\hbar^a)) \quad (3.75)$$

Proof. First, we Taylor expand $f(x)$ around x_c . Since x_c is the minimum $f'(x_c)$ vanishes. After this, we can write the integrand as

$$e^{a+bx^2+cx^3} = e^{a+bx^2} (1 + cx^3 + \dots). \quad (3.76)$$

With this, we can write out the approximation as in the problem. With these, we have

$$\begin{aligned} I &\approx \sum_{x_c} \int_{\mathbb{R}} e^{-f(x_c)/\hbar} e^{\frac{-f''(x_c)(x-x_c)^2}{2\hbar}} \left(1 + \frac{f'''(x_c)(x-x_c)^3}{6\hbar} + \dots \right) \\ &\quad \times \left(1 + \frac{f^{(4)}(x_c)(x-x_c)^4}{24\hbar} + \dots \right) dx \\ &= \sum_{x_c} \sqrt{\frac{2\pi\hbar}{f''(x_c)}} e^{-f(x_c)/\hbar} (1 + \mathcal{O}(\hbar^a)). \end{aligned} \quad (3.77)$$

□

2. Express the most important corrections to equation (2) in the form

$$\int_{\mathbb{R}} dx C_1 x^n e^{-C_2 x^2} \quad (3.78)$$

where C_1, C_2 are constants.

Proof. Note that when n is odd, the corrections vanish. So, we only pay attention to the even n , the smallest of which is $n = 4$. So the largest correction is

$$\int_{\mathbb{R}} \frac{f^{(4)}(x_c)(x-x_c)^4}{24\hbar} e^{\frac{-f''(x_c)(x-x_c)^2}{2\hbar}} dx. \quad (3.79)$$

□

3. Determine the value of a in equation (2) by making a change of variables.

Proof. Now, making the substitution $y = (x - x_c)/\sqrt{\hbar}$. with this, we get $a = 1$. Without loss of generality, let us set $x_c = 0$. This makes things a bit easier. Then $dx = \sqrt{\hbar} dy$ and the correction integral becomes

$$\int_{\mathbb{R}} dy \sqrt{\hbar} \hbar^2 \frac{C_1}{\hbar} e^{-C_2 y^2} \sim \mathcal{O}(\hbar^{\exists/\epsilon}). \quad (3.80)$$

Now because we have a leading factor of $\sim \sqrt{\hbar}$ for the overall integral, the correction as written in the problem follows $\mathcal{O}(\hbar^1)$. □

7. Perturbation Theory. In this problem we will show where perturbation theory breaks down.

1. Show that

$$\exp \left[-\lambda \left(\frac{d}{dJ} \right)^4 \right] \int_{\mathbb{R}} \exp [-x^2 + xJ] = \int_{\mathbb{R}} dx \exp [-x^2 - \lambda x^4 + xJ]. \quad (3.81)$$

Path integrals with interactions can be computed by taking derivatives of free path integrals with a source J using generalizations of this formula.

Proof. This is actually not hard. One questionable step we have to make is the interchanging of the sum and the integral. However, if we assume that the integral of a sum is equal to the sum of integrals then we can Taylor expand the leading factor in monomials of the operator d/dJ and let it act on the integrand. Once this is done, we can rewrite the sum in the integrand as $e^{-\lambda x^4}$, as we should be able to do. I won't write out the math here, because it's already been done in my [Q-CFT notes](#). \square

2. Use induction to show that

$$\int_{\mathbb{R}} dx x^{2n} \exp \left[-\frac{1}{2} ax^2 \right] = \left(\frac{2\pi}{a} \right)^{1/2} \frac{1}{a^n} (2n-1)!! \quad (3.82)$$

where $(2n-1)!! = (2n-1)(2n-3)\dots 5 \cdot 3 \cdot 1$. Hint: differentiate with respect to a .

Proof. This boils down to doing induction correctly. This requires evaluating a quite simple Gaussian integral. Most of this will be algebra, so I won't write a solution here. \square

3. Show that

$$\int_{\mathbb{R}} dx \exp [-x^2 - \lambda x^4] = \sqrt{\pi} \sum_{j=0}^{\infty} \frac{(-\lambda)^j (4j-1)!!}{2^{2j} j!} \quad (3.83)$$

Proof. The key is to expand the factor $e^{-\lambda x^4}$ into a power series. This power series will contain powers of x :

$$\begin{aligned} \int_{\mathbb{R}} dx \exp [-x^2 - \lambda x^4] &= \sum_{j=0}^{\infty} \int_{\mathbb{R}} \frac{(-\lambda)^j x^{4j}}{j!} e^{-x^2} \\ &= \sqrt{\pi} \sum_{j=0}^{\infty} \frac{(-\lambda)^j (2 \times 2j-1)!!}{2^{2j} j!} \\ &= \sqrt{\pi} \sum_{j=0}^{\infty} \frac{(-\lambda)^j (4j-1)!!}{2^{2j} j!} \end{aligned} \quad (3.84)$$

\square

4. Show that the series in part (c) diverges.

Proof. The series diverges because of the factorial terms – they don’t die out quickly enough, or not at all! \square

5. Which step in the derivation was not justified?

Proof. The problem occurs when we swap the integral and the sum. It’s not always a valid operation. So why do we still do this? We still do this because most problems can be solved this way. The series diverges sometimes, but the operation can give very good predictions. The series, while diverges, is asymptotic to the series

$$\lim_{\lambda \rightarrow 0+} \frac{I(\lambda) - \sum_{j=0}^N a_j \lambda^j}{\lambda^N} = 0. \quad (3.85)$$

This means that we should be careful when using perturbation theory. Also, this points out a problem with perturbation theory. There are effects that can’t be accounted for by perturbation theory. There are effects such that

$$\lim_{\lambda \rightarrow 0+} \frac{e^{-1/\lambda}}{\lambda^N} = 0 \quad (3.86)$$

that are invisible to perturbation theory. These are phenomena such as tunneling rates or energy splittings. \square

8. Instantons

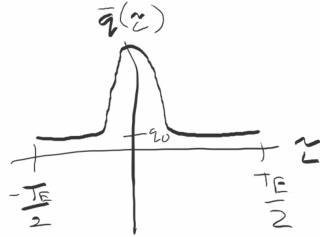
1. We want to solve the equation

$$-\frac{d^2 \bar{q}}{d\tau^2} \quad (3.87)$$

with boundary conditions $\bar{q}(\pm T_E/2) = q_0$. These equations are the equations of motion of a particle with mass $m = 1$ moving in a potential. Make the analogy with a particle moving in a potential explicit, paying attention to the signs.

Proof. The EOM is the same as the EOM for a particle moving in an inverted potential. \square

2. There is a trivial solution to equations of motion and boundary conditions. What is it? What is the action of this trivial solution?



Proof. The trivial solution is that the particle stays at q_0 at all times. Non-trivial solutions look like

This solution has finite action:

$$S_0 = S_E = \int d\tau \left(\frac{1}{2} \left(\frac{d\bar{q}}{d\tau} \right)^2 + V(\bar{q}) \right) \quad (3.88)$$

This is not the only solution, however. Another solution is where the particle bounces back and forth between q_0, q_* . The action in this case is just going to be nS_0 where n is the number of times the particle bounces back and forth. In this case,

$$K_E = \sum_{n=0}^{\infty} \frac{T_E^n}{n!} e^{-nS_0/\hbar} = \exp \left[cT_E e^{-S_0/\hbar} \right]. \quad (3.89)$$

□

3. Using conservation of energy, show that a solution to the equation of motion with zero energy has

$$\bar{q}(\tau) - q_0 \approx e^{-\omega|\tau|} \quad (3.90)$$

for large $|\tau|$ (that is $|\tau| \approx T_E/2$).

Proof. Differentiate the differential equation, we will get $V'' = \omega^2$. Then integrate again. This is very hand-wavy, so it is understandable to be extremely confused. □

4. We are interested in the $T_E \rightarrow \infty$ limit. Is the action of the solution you found finite and non-zero in this limit?

Proof. This is more or less straightforward. □

5. Change integration variables and use conservation of energy to show that the action

$$S_0 = \lim_{T_E \rightarrow \infty} \int_{-T_E/2}^{T_E/2} d\tau \left(\frac{1}{2} \left(\frac{d\bar{q}}{d\tau} \right)^2 + V(\bar{q}) \right) \quad (3.91)$$

of the solution you found can be expressed as

$$S_0 = 2 \int_{q_0}^{q_*} d\bar{q} \sqrt{2V(\bar{q})} \quad (3.92)$$

where q_* is the other point where the potential vanishes $V(q_*) = 0$.

Proof. □

6. Are there any solutions with non-zero energy with finite action in the $T_E \rightarrow \infty$ limit?

Proof. □

7. Are there any other solutions with zero energy and finite action?

Proof. □

9. Anyons.

1. In two spatial dimensions we can also write the amplitude for two particles with initial relative position \mathbf{q} to propagate to the relative position \mathbf{q}' in a time T as

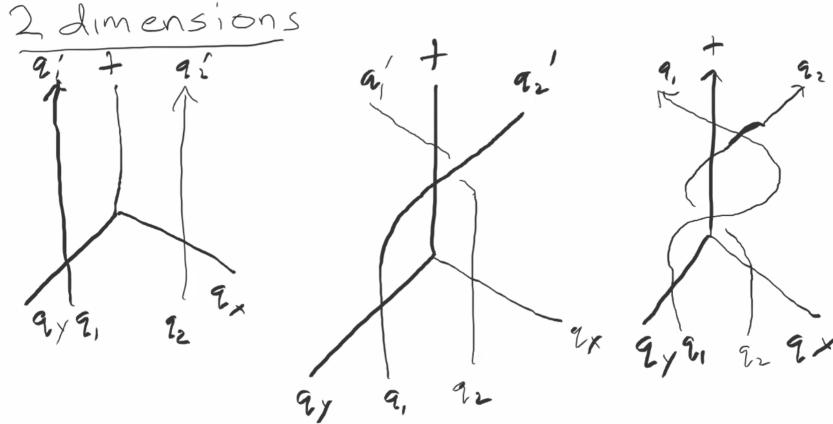
$$A(\mathbf{q}', T, \mathbf{q}, 0) = \sum_{\text{paths}} e^{iS[\mathbf{q}_1(t), \mathbf{q}_2(t)]} \quad (3.93)$$

As in lecture we will assume the particles cannot be at the same point due to some strong repulsive short range interaction. If the particles are identical, what is the configuration space? What is a natural covering space?

Proof. $(\mathbb{R}^2 \setminus \{0\}) \setminus \mathcal{Z}_2$ and $\mathbb{R}^2 \setminus \{0\}$. □

2. As in three spatial dimensions the paths where the initial and final relative positions are the same $\mathbf{q} = \mathbf{q}'$ can be classified into topological classes. These paths can be characterized by the angular displacement $n\pi$ of the relative coordinate where n is an integer. Sketch some of these paths. Argue that two paths with angular displacements $n\pi$ and $m\pi$ are topologically distinct if $m \neq n$.

Proof. Two paths are topologically equivalent if they can be smoothly deformed into another. We want to show that there are infinitely many classes of topologically distinct paths. Here are the pictures: There's no way to "untangle" the paths in two dimensions (which is something we can do in 3 dimensions), so these are topologically distinct paths. In the first case, the net angular displacement is 0. The subsequent ones are $\pi, 2\pi, \dots$. So, the n here is the number of exchanges, which corresponds exactly its topological class. □



3. Argue that in polar coordinates $\mathbf{q} = (q, \theta)$ the path integral in configuration space $A(\mathbf{q}, T, \mathbf{q}, 0) = A(q, \theta, T, q, \theta, 0)$ satisfies

$$A(q', \theta' + \pi, T, q, \theta, 0) = e^{i\phi} A(q', \theta', T, q, \theta, 0). \quad (3.94)$$

Proof. From the exercise above, we get that A is a sum over all topological classes $n : -\infty \rightarrow \infty$:

$$A = \sum_{n=-\infty}^{\infty} C_n \bar{A}_n. \quad (3.95)$$

Now, if we add π to the final relative displacement, we get one more exchange. In the covering space, n one bigger, but in the configuration space it is the same, so we get a phase. So we get

$$A(\theta' + \pi) = e^{-i\phi} A(\theta'). \quad (3.96)$$

□

4. The path integral in configuration space $A(\mathbf{q}, T, \mathbf{q}, 0)$ can be written as a sum of path integrals in the covering space $\bar{A}_n(\mathbf{q}, T, \mathbf{q}, 0)$

$$A(\mathbf{q}, T, \mathbf{q}, 0) = \sum_{n=-\infty}^{\infty} \bar{A}_n(\mathbf{q}, T, \mathbf{q}, 0) \quad (3.97)$$

where only paths with angular displacement $n\pi$ contribute to the path integral $\bar{A}_n(\mathbf{q}, T, \mathbf{q}, 0)$. If we add phases C_n as in lecture, the amplitude in polar coordinates $\mathbf{q} = (q, \theta)$ becomes

$$A(q, \theta, T, q, \theta, 0) = \sum_{n=-\infty}^{\infty} C_n \bar{A}_n(q, \theta, T, q, \theta, 0). \quad (3.98)$$

Use equation

$$A(q', \theta' + \pi, T, q, \theta, 0) = e^{i\phi} A(q', \theta', T, q, \theta, 0) \quad (3.99)$$

to show that

$$C_n = e^{i\phi} C_{n-1}. \quad (3.100)$$

Proof. From the item above, we basically have a shifted sum. The index shift by 1 when we add π . So

$$C_n = e^{i\phi} C_{n-1}. \quad (3.101)$$

□

5. Are there any constraints on the phase ϕ ?

Proof. $\phi \in \mathbb{R}$, i.e., there are no constraints. This means we not only get BE and FD statistics but also fractional statistics such as anyons. They haven't been observed, but they are used to study other phenomena (e.g. fractional quantum Hall effect). □

Part 4

Symmetries

Instructor: Giuseppe Sellaroli

The aim of this course is to explore some of the many ways in which symmetries play a role in physics. We'll start with an overview of the concept of symmetries and their description in the language of group theory. We will then discuss continuous symmetries and infinitesimal symmetries, their fundamental role in Noether's theorem, and their formalization in terms of Lie groups and Lie algebras. In the last part of the course we will focus on symmetries in quantum theory and introduce representations of (Lie) groups and Lie algebras.

4.1 Review: Equivalence Classes

4.1.1 Equivalence relations

An equivalence relation is a special kind of binary relation. A (homogeneous) binary relation on a set X is a way of encoding a relationship between two elements of X . Given a relation R we use the notation xRy to mean “ x is related to y .”

Examples of relations are $=, <, >, \leq, \geq$ on \mathbb{R} .

Definition 4.1.1 (Homogeneous binary relation). A homogeneous binary relation on a set X is a set $R \subset X \times X$ of pairs of elements of X . When $(x, y) \in R$ we say that x is relate to y , which we also indicate with the less cumbersome notation xRy .

We can now define equivalence relations.

Definition 4.1.2 (Equivalence relation). An equivalence relation on a set X is a homogeneous binary relation R that satisfies the following properties:

- Reflexive: $xRx \forall x \in X$.
- Symmetric: $xRy \implies yRx$.
- Transitive: $xRy, yRz \implies xRz$

These properties are chosen as they encode the fundamental features of the notion of things being “equivalent.”

Equivalence relations are often denoted by the symbol \sim . In this case we read $x \sim y$ as “ x is equivalent to y .”

4.1.2 Equivalence Classes

Now that we know what an equivalence relation is, we can define the extremely useful concept of equivalence class.

Definition 4.1.3. Let \sim be an equivalence relation on X . The equivalence class of $x \in X$ is the set

$$[x]_\sim = \{y \in X | y \sim x\} \quad (4.1)$$

of all the elements in X equivalent to x . We often drop the subscript and just write $[x]$ if there is no risk of confusion. Note that

$$[x]_\sim = [y]_\sim, \quad \forall y \in [x]_\sim, \quad (4.2)$$

that is we can label the equivalence class using any of its elements. The element that we choose as the label is called the representative of the equivalence class.

Equivalence classes are essentially buckets in which we put elements of X that are equivalent to each other. Inequivalent elements end up in different equivalence classes. One way of thinking about this is that an equivalence class is a way to encode as a single object the properties common to its elements, which are otherwise indistinguishable.

Example 4.1.1. An example of equivalence relation appears in modular arithmetic. Let's consider the simplest case of the integers modulo 2. We define an equivalence relation \sim on \mathbb{Z} by

$$x \sim y \implies x - y = 2k, k \in \mathbb{Z}, \quad (4.3)$$

that is x, y are equivalent if they differ by a multiple of two. If you want, you can check that this is indeed an equivalence relation as an exercise. There are only two distinct equivalence classes, which are the respectively the sets of even and odd integers. The sets of integers modulo 2 is defined as

$$\mathbb{Z}_2 = \{[0], [1]\}. \quad (4.4)$$

Although this is a set of sets, we really want to pretend that it's a set of numbers. To do that, we define a way to add elements in \mathbb{Z}_2 . The obvious choice is

$$[a] + [b] = [a + b], \quad (4.5)$$

but at this stage this expression is purely formal: since we are specifying this operation through the sum of the representatives of each class, we need to make sure that the definition is actually independent of the choice of representatives (since they should be equivalent). This happens all the time when dealing with equivalence classes! In this case we are fine, because of the rules of adding odd and even numbers.

Example 4.1.2. Here's an example from quantum physics. Let \mathcal{H} be the Hilbert space of a quantum theory. While extremely useful, the states $|\psi\rangle \in \mathcal{H}$ are not *physical*, where physical is meant as something that can be measured. The physical quantities in a quantum theory are the expectation values of the observables in a given state, that is

$$\langle A \rangle_\psi = \frac{\langle \psi | A | \psi \rangle}{\langle \psi | \psi \rangle}. \quad (4.6)$$

Now consider another state $|\phi\rangle = \lambda|\psi\rangle$, with $\lambda \in \mathbb{C} \setminus \{0\}$. For any observable A we have

$$\langle A \rangle_\phi = \frac{|\lambda|^2}{|\lambda|^2} \frac{\langle \psi | A | \psi \rangle}{\langle \psi | \psi \rangle} = \langle A \rangle_\psi. \quad (4.7)$$

This means that the states $|\psi\rangle$ and $|\phi\rangle$ are physically indistinguishable. It is then natural to consider them equivalent, and define the equivalence relation

$$|\psi\rangle \sim |\phi\rangle \iff |\phi\rangle = \lambda|\psi\rangle, \quad \lambda \in \mathbb{C} \setminus \{0\}. \quad (4.8)$$

The set

$$P(\mathcal{H}) = \{[|\psi\rangle] \mid |\psi\rangle \in \mathcal{H} \setminus \{0\}\} \quad (4.9)$$

is called the *projective Hilbert space* associated to \mathcal{H} , and can be interpreted as the set of all distinguishable quantum states. The zero vector was removed as it does not describe any physical state (expectation values don't make sense for the zero vector).

In a sense $P(\mathcal{H})$ is the true set of states of a quantum theory. The reason why we work with \mathcal{H} instead is that the latter is a vector space, which comes with a lot of nice properties and allows for an easy description of the superposition principle.

4.2 Definition of symmetry

We can't really give a mathematical definition of symmetry, because it depends on the context. But in any case we can define "symmetry" like this.

Definition 4.2.1 (Symmetries). Loosely speaking, a symmetry is a transformation of A that leaves B invariant.

The idea is that A and B are not necessarily the same thing.

Example 4.2.1. If B is the golden spiral and $A = \mathbb{R}^2$. When A is scaled, B remains invariant.

Example 4.2.2. Consider a Lagrangian,

$$S = \int d^4x \mathcal{L} \quad (4.10)$$

and $B = S$ and A is time, which we translate. We can see that S is invariant under time translation.

Some properties:

- (Identity) "Doing nothing" is a symmetry. It is trivial, but is a symmetry nonetheless. This is also called the "identity" transformation.
- (Inverses) Symmetries are invertible.
- (Composition) Compositions of symmetries are also symmetries.

4.3 Elements of group theory

Definition 4.3.1. A group is a set G together with an operation $* : G \times G \rightarrow G$ satisfying the following properties:

- There is an identity element: $g * e = e * g = g, \forall g \in G$
- Each element of G has an inverse, i.e., $\forall g \in G, \exists g^{-1} \in G$ s.t. $g^{-1} * g = g * g^{-1} = e \in G$.
- The operation $*$ is associative, that is $a * (b * c) = (a * b) * c, \forall a, b, c \in G$.

Additionally, we say that the group G is abelian/commutative if $a * b = b * a, \forall a, b \in G$.

Definition 4.3.2 (Subgroups). Let G be a group with operation $*$. A subgroup of G is a subset $H \subset G$ that contains the identity element and is closed under the operation $*$ and under inversion, that is

- $e \in H$
- $a * b \in H \forall a, b \in H$

- $a \in H \implies a^{-1} \in H$.

The notation $H \leq G$ is commonly used to indicate that H is a subgroup of G .

Definition 4.3.3 (Group homomorphism). A group homomorphism is a map $\varphi : G \rightarrow H$ between two groups G and H such that

$$\varphi(a *_G b) = \varphi(a) *_H \varphi(b), \forall a, b \in G. \quad (4.11)$$

Definition 4.3.4 (Kernel of a group homomorphism). The kernel of a group homomorphism $\varphi : G \rightarrow H$ is the set

$$\ker \varphi = \{g \in G | \varphi(g) = e_H\} \quad (4.12)$$

of all the elements of G that are sent to the identity in H .

Proposition 4.3.1. A group homomorphism $\varphi : G \rightarrow H$ is injective if and only if the kernel is trivial: $\ker \varphi = \{e_G\}$.

Proof.

- (\rightarrow), if φ is injective, then at most one thing can be sent to e_H . This means that $\ker \varphi = \{e_G\}$.
- (\leftarrow) Let $a, b \in G$ such that $\varphi(a) = \varphi(b) \implies e_H = \varphi(a)^{-1}\varphi(b) = \varphi(a^{-1})\varphi(b) = \varphi(a^{-1}b)$. But this means $a^{-1}b \in \ker \varphi \implies a^{-1}b = e_G \implies a = b$. So φ is injective.

□

Definition 4.3.5 (Isomorphisms). Two groups G and H are isomorphic (denoted by $G \sim H$) if there exists an invertible group homomorphism $\varphi : G \rightarrow H$. Such a map is called an isomorphism between G and H .

Definition 4.3.6 (Normal Subgroup). Let G be a group. A subgroup $N \leq G$ is called *normal* if

$$gng^{-1} \in N, \forall g \in G, \forall n \in N. \quad (4.13)$$

The notation $N \trianglelefteq G$ is commonly used to indicated that N is a normal subgroup of G .

Definition 4.3.7 (Quotient Group). Let $N \trianglelefteq G$. We can define an equivalence relation on G as

$$g \sim h \iff h^{-1}g \in N \quad (4.14)$$

with equivalence classes

$$[g] = \{h \in G | h^{-1}g \in N\}. \quad (4.15)$$

The quotient group $G/N \equiv G \bmod N$ is the set of equivalence classes

$$G/N = \{[g] | g \in G\} \quad (4.16)$$

which is made into a group by defining

$$[g][h] = [gh], \quad [g]^{-1} = [g^{-1}], \quad e_{G/N} = [e_G]. \quad (4.17)$$

We note that the equivalence class does not require N be normal. However, to turn the equivalence classes into a group, we require that N is normal.

Theorem 4.3.1 (First Isomorphism Theorem). Let $\varphi : G \rightarrow H$ be a group homomorphism. Then:

- $\text{Im } \varphi$ is a subgroup of H .
- $\ker \varphi$ is a normal subgroup of G .
- $\text{Im } \varphi$ is isomorphism to the quotient group $G/\ker \varphi$.

The notion that $\text{Im } \varphi \sim G/\ker \varphi$ is important. This will appear again when we see Lie groups.

4.4 Examples: Symmetry & Group Theory

Example 4.4.1. $(\mathbb{Z}, +), (\mathbb{R}, +), (\mathbb{C}, +), \dots$ are abelian groups.

Example 4.4.2. $(\mathbb{R} \setminus \{0\}, \times)$

Example 4.4.3. Invertible $n \times n$ matrices, $GL(n, \mathbb{R})$ under usual matrix multiplication.

Example 4.4.4. Circle group $S = \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 = 1\}$. Identity $(x, y) \in S$ with a complex number $z = x + iy$. Then use complex number multiplication. This works because $|z| = 1 \implies z = e^{i\theta}$.

Example 4.4.5 (Isomorphism). Consider the two groups $(\mathbb{R}, +)$ and $(\mathbb{R}^+ \setminus \{0\}, \times)$. The exponential map: $x \in \mathbb{R} \mapsto e^x$. This map is obviously invertible. It is also a group homomorphism because it relates multiplication in the target group to addition in the reals. So, the two groups are isomorphic.

This is particularly useful when we want to move from additive groups to multiplicative groups.

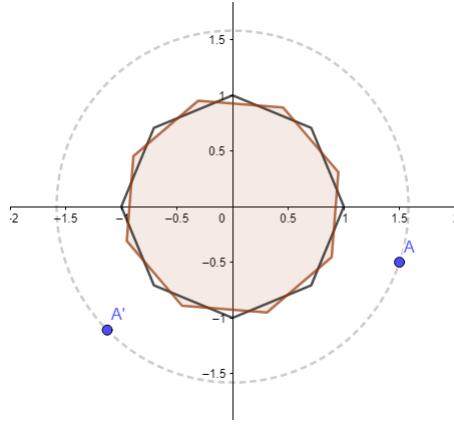
This map is related to Lie groups, etc which we will see later.

4.5 Continuous and discrete symmetries

We will be quite vague with these definitions.

4.5.1 Discrete Symmetries

Think of rotating polygons. Rotations that leave the polygon invariant are in integer steps. This is a prototype example of what a discrete symmetry is.

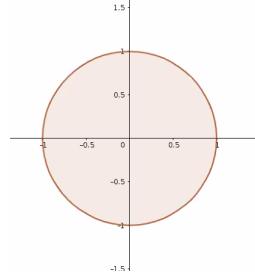


We can also easily define equivalence classes etc and see how the first isomorphism comes up, but we won't worry about that.

The idea of discrete symmetries is this: there is a discrete parameter for the group elements.

4.5.2 Continuous Symmetries

In view of the previous example, we can consider rotating a circle.



Obviously we can rotate by any angle and keep the circle invariant.

The idea of continuous symmetries is this: there is a continuous parameter for the group elements. Consider $\epsilon \in \mathbb{R}^+$ and take $\epsilon \mapsto \psi(\epsilon) = g_\epsilon \in G$. ψ is a

group homomorphism. Now, if there is a **topology** on G such that ψ is continuous, then we say $\{g_\epsilon | \epsilon \in \mathbb{R}^+\} = \text{Im}(\psi)$ is continuous. This is quite vague, but is essentially the idea.

The group $\{g_\epsilon | \epsilon \in \mathbb{R}^+\}$ is called a **one-parameter subgroup**. One-parameter subgroups satisfy the following properties:

- $g_0 = e_G$
- $g_{-\epsilon} = g_\epsilon^{-1}$
- $g_{\epsilon+\epsilon'} = g_\epsilon g_{\epsilon'}$. This means the one-parameter subgroup is abelian.

For example consider g_θ define by 2×2 rotation matrix R_θ . We can check that g_θ satisfies all of the properties above.

4.6 Infinitesimal symmetries

Rotations are a particular example of infinitesimal symmetry. Just like calculus, we want to talk about “derivatives.” Assume that $\epsilon \in \mathbb{R} \mapsto g_\epsilon \in G$ is a “smooth” map. Here, “smooth” means infinitely differentiable, or “as nice as we want.” This only makes sense if we have a “differentiable structure” on G .

Now, if the map above is smooth, we can linearize it. (Now we’re doing things in the physicist’s way). To linearize:

$$g_\epsilon = g_0 + \epsilon \frac{d}{d\epsilon} g_\epsilon \Big|_{\epsilon=0} + \mathcal{O}(\epsilon^2). \quad (4.18)$$

The infinitesimal symmetry is then

$$g_\epsilon = g_0 + \epsilon \frac{d}{d\epsilon} g_\epsilon \Big|_{\epsilon=0}. \quad (4.19)$$

where g_0 is the identity.

More formally, infinitesimal symmetry has the form $e + \epsilon X$ where

$$X = \frac{d}{d\epsilon} \Big|_{\epsilon=0} g_\epsilon. \quad (4.20)$$

Example 4.6.1. Consider

$$R_\theta = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \quad (4.21)$$

With this,

$$\frac{d}{d\theta} \Big|_{\theta=0} R_\theta = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad (4.22)$$

This means that the Taylor expansion of R_θ around 0 is given by

$$R_\theta = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \theta \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + \mathcal{O}(\theta^2) \quad (4.23)$$

The infinitesimal rotation is thus

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \theta \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad (4.24)$$

Note that because of the group structure, we can actually reconstruct R_θ from the linearization, unlike in calculus where we don't necessarily have the function back knowing its derivative.

4.7 Noether's theorem

We first want to look at the symmetry of an action. Consider the Lagrangian

$$\mathcal{L}(q, \dot{q}, t) \quad (4.25)$$

for a particle in one dimension. It's easy to generalize to higher dimensions, so we'll just worry about the 1-d case. The action is given by

$$S[q] = \int_a^b \mathcal{L}(q, \dot{q}, t) dt. \quad (4.26)$$

Recall that symmetry means some transformation (in t, q) that leaves the Lagrangian invariant.

Consider continuous (smooth) symmetries in t, q given by $\tilde{t} = T_\epsilon(t)$ and $\tilde{q}(\tilde{t}) = Q_\epsilon(q(t)) = Q_\epsilon(q(T_{-\epsilon}(\tilde{t})))$. We assume that $T_0 = \mathcal{I}$ and other characteristics of a one-parameter subgroup.

We want to look at how S changes. So we look at

$$\tilde{S}[\tilde{q}] = \int_{T_\epsilon(a)}^{T_\epsilon(b)} d\tilde{t} \mathcal{L}(\tilde{q}, \dot{\tilde{q}}, \tilde{t}). \quad (4.27)$$

Now, S is invariant under the transformations above if $S[q] = \tilde{S}[\tilde{q}]$. If this happens, we say that we have a symmetry.

Note: There is also a notion of a quasi-symmetry, but we won't worry about that.

Definition 4.7.1 (Noether's Theorem). Every differentiable symmetry of the action of a physical system has a corresponding conservation law.

Proof. Consider $\tilde{t} = T_\epsilon(t)$ and $\tilde{q}(\tilde{t}) = Q_\epsilon(q(t)) = Q_\epsilon(q(T_{-\epsilon}(\tilde{t})))$ with T_ϵ, Q_ϵ smooth one-parameter subgroups ($T_0 = \mathcal{I}, T_{-\epsilon} = T_\epsilon^{-1}, T_{\epsilon+\epsilon'} = T_\epsilon \circ T_{\epsilon'}^{-1}$), i.e., $(d/d\epsilon)_{\epsilon=0} T_\epsilon$ makes sense. Consider the action

$$S[q] = \int_a^b \mathcal{L}(q(t), \dot{q}(t), t). \quad (4.28)$$

Under the transformation, we get

$$\tilde{S}[\tilde{q}] = \int_{T_\epsilon(a)}^{T_\epsilon(b)} \mathcal{L}(\tilde{q}(\tilde{t}), \dot{\tilde{q}}(\tilde{t}), \tilde{t}) d\tilde{t} = \int_a^b \frac{dT_\epsilon}{dt} \mathcal{L}(\tilde{q}(T_\epsilon(t)), \dot{\tilde{q}}(T_\epsilon(t)), T_\epsilon(t)) dt. \quad (4.29)$$

Suppose that $\tilde{S}[\tilde{q}] = S[q]$ for all paths (even those that are not solutions of the Euler-Lagrange equations) and for all choices of a, b . Then

$$\delta S \equiv \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} \tilde{S}[\tilde{q}] = 0. \quad (4.30)$$

Some notation:

$$\begin{aligned} \delta t &= \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} T_\epsilon \\ \delta q &= \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} Q_\epsilon \end{aligned} \quad (4.31)$$

These are generators of T_ϵ and Q_ϵ , respectively (see Lie algebras).

Let's apply $D := (d/d\epsilon)|_{\epsilon=0}$ everywhere we can, then

$$\begin{aligned} D \frac{dT_\epsilon}{dt} &= \frac{d}{dt} DT_\epsilon = \frac{d}{dt} \delta t \\ D\tilde{q}(\tilde{t}) &= DQ_\epsilon(q(t)) = \delta q(q(t)) \\ \frac{d\tilde{q}}{d\tilde{t}}(\tilde{t}) &= \frac{d}{d\tilde{t}} Q_\epsilon(q(t)) = \frac{\partial Q_\epsilon}{\partial q}(q(t)) \dot{q}(t) \frac{dt}{d\tilde{t}}(\tilde{t}) \\ \implies \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} \dot{\tilde{q}}(\tilde{t}) &= \underbrace{\frac{\partial}{\partial q} DQ_\epsilon(q(t))}_{1} \dot{q}(t) \underbrace{\frac{1}{\frac{dT_0}{dt}}}_{-(d/dt)\delta t} + \underbrace{\frac{\partial}{\partial q} Q_0(q(t))}_{-(d/dt)\delta t} \dot{q}(t) \quad D \frac{1}{\frac{dT_\epsilon}{dt}} \\ \implies D\dot{\tilde{q}}(\tilde{t}) &= \frac{\partial}{\partial q} \delta q \dot{q}(t) - \dot{q}(t) \frac{d}{dt} \delta t = \frac{d}{dt} \delta q - \dot{q}(t) \frac{d}{dt} \delta t. \end{aligned} \quad (4.32)$$

Now suppose that q satisfies the Euler-Lagrange equation:

$$\frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}} = 0. \quad (4.33)$$

Then because

$$D\tilde{S}[\tilde{q}] = 0, \quad \tilde{S}[\tilde{q}] = S[q] \quad (4.34)$$

which we should check without using the Euler-Lagrange equations, we have that

$$\begin{aligned} 0 = D\tilde{S}[\tilde{q}] &= \int_a^b D \left[\frac{dT_\epsilon}{dt} \mathcal{L}(\tilde{q}(\tilde{t}), \dot{\tilde{q}}(\tilde{t}), \tilde{t}) \right] \\ &= \int_a^b dt \left\{ \frac{d\delta t}{dt} \mathcal{L}(q, \dot{q}, t) + \underbrace{\frac{dT_0}{dt}}_1 D\mathcal{L}(\tilde{q}(\tilde{t}), \dot{\tilde{q}}(\tilde{t}), \tilde{t}) \right\}. \end{aligned} \quad (4.35)$$

Now, what is the term on the RHS of the integral? Well,

$$\begin{aligned} D\mathcal{L}(\tilde{q}(\tilde{t}), \dot{\tilde{q}}(\tilde{t}), \tilde{t}) &= \frac{\partial \mathcal{L}}{\partial t} D\tilde{t} + \frac{\partial \mathcal{L}}{\partial q} D\tilde{q}(\tilde{t}) + \frac{\partial \mathcal{L}}{\partial \dot{q}} D\dot{\tilde{q}}(\tilde{t}) \\ &= \frac{\partial \mathcal{L}}{\partial t} \delta t + \frac{\partial \mathcal{L}}{\partial q} \delta q + \frac{\partial \mathcal{L}}{\partial \dot{q}} \left(\frac{d}{dt} \delta q - \dot{q} \frac{d}{dt} \delta t \right) \\ &= \frac{\partial \mathcal{L}}{\partial t} \delta t + \frac{\partial \mathcal{L}}{\partial q} \delta q + \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}} \delta q \right) - \delta q \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}} \dot{q} \delta t \right) + \delta t \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}} \dot{q} \right) \\ &= \frac{d}{dt} \left[\frac{\partial \mathcal{L}}{\partial \dot{q}} (\delta q - \dot{q} \delta t) \right] + \frac{\partial \mathcal{L}}{\partial t} \delta t + \delta t \left(\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}} \right) \dot{q} + \delta t \frac{\partial \mathcal{L}}{\partial \dot{q}} \ddot{q} \\ &= \frac{d}{dt} \left[\frac{\partial \mathcal{L}}{\partial \dot{q}} (\delta q - \dot{q} \delta t) \right] + \delta t \left[\frac{\partial \mathcal{L}}{\partial t} + \frac{\partial \mathcal{L}}{\partial q} \dot{q} + \frac{\partial \mathcal{L}}{\partial \dot{q}} \ddot{q} \right] \\ &= \frac{d}{dt} \left[\frac{\partial \mathcal{L}}{\partial \dot{q}} (\delta q - \dot{q} \delta t) \right] + \delta t \frac{d}{dt} \mathcal{L}(q(t), \dot{q}(t), t). \end{aligned} \quad (4.36)$$

From here we have that

$$\begin{aligned} 0 &= \int_a^b dt \left\{ \mathcal{L}(q, \dot{q}, t) \frac{d}{dt} \delta t + \delta t \frac{d}{dt} \mathcal{L}(q, \dot{q}, t) + \frac{d}{dt} \left[\frac{\partial \mathcal{L}}{\partial \dot{q}} (\delta q - \dot{q} \delta t) \right] \right\} \\ &= \int_a^b dt \frac{d}{dt} \left[\mathcal{L} \delta t + \frac{\partial \mathcal{L}}{\partial \dot{q}} [\delta q - \dot{q} \delta t] \right] \\ \implies &\boxed{\mathcal{L} \delta t + \frac{\partial \mathcal{L}}{\partial \dot{q}} (\delta q - \dot{q} \delta t) \text{ is conserved!}} \end{aligned} \quad (4.37)$$

□

Example 4.7.1. Consider a time-independent Lagrangian $\mathcal{L} = \mathcal{L}(q, \dot{q})$. Let $\tilde{t} = t + \epsilon$ and $\tilde{q}(\tilde{t}) = q(t)$, i.e., $Q_\epsilon = \mathcal{I}$. Indeed, these two transformations are one-parameter subgroups. Now,

$$\begin{aligned} \tilde{S} &= \int_{a+\epsilon}^{b+\epsilon} \mathcal{L}(\tilde{q}, \dot{\tilde{q}}) d\tilde{t} \\ &= \int_a^b \mathcal{L}(q, \dot{q}) dt \\ &= S[q]. \end{aligned} \quad (4.38)$$

This does not depend on a, b , so we have a symmetry. What is the conserved quantity in this case? Well, we know that $\delta t = (d/d\epsilon)|_{\epsilon=0} T_\epsilon(t) = 1$ and $\delta q = 0$. So the conserve quantity is given by

$$\mathcal{L}\delta t + \frac{\partial \mathcal{L}}{\partial \dot{q}} (\delta q - \dot{q}\delta t) = \mathcal{L}(q, \dot{q}) - \frac{\partial \mathcal{L}}{\partial \dot{q}} \dot{q} = \text{Energy}. \quad (4.39)$$

4.7.1 Noether's Theorem for fields in flat space (quais-symmetry)

4.7.2 Noether's Theorem: Fields in Flat Space

4.8 Lie groups and Lie algebras

4.8.1 Matrix Exponential

Let's review of facts about matrix exponential:

- The matrix exponential is the map $\exp : M_n(\mathbb{C}) \rightarrow GL(n, \mathbb{C})$ defined by $X \mapsto \sum_{k=0}^{\infty} X^k/k!$.
- The map \exp is surjective.
- $e^{AXA^{-1}} = Ae^XA^{-1}$, for $A \in GL(n, \mathbb{C})$.
- $\det(e^E) = e^{\text{tr } E}$.
- $e^O = \mathcal{I}$.
- $(e^X)^* = e^{X^*}$.
- $(e^X)^\top = e^{X^\top}$
- $(e^X)^{-1} = e^{-X}$.
- $e^{(\alpha+\beta)X} = e^{\alpha X}e^{\beta X}$.
- If $XY = YX$ then $e^{X+Y} = e^Xe^Y = e^Ye^X$.
- In general, we have the Lie product formula: $e^{(X+Y)} = \lim_{k \rightarrow \infty} (e^{X/k}e^{Y/k})^k$.
- $\frac{d}{d\epsilon}e^{\epsilon X} = Xe^{\epsilon X} = e^{\epsilon X}X, \epsilon \in \mathbb{R}$.
- $\left. \frac{d}{d\epsilon} \right|_{\epsilon=0} e^{\epsilon X} = X$.
- $e^{\epsilon X} = e^{\epsilon Y} \forall \epsilon \in \mathbb{R} \iff X = Y$.

4.8.2 Matrix Lie groups and associated Lie algebras

Definition 4.8.1 (Matrix Lie Groups). A matrix Lie group is a closed subgroup $G \leq GL(n, \mathbb{C})$ for some $n \in \mathbb{N}$ (closed wrt the topology induced from $M_n(\mathbb{C})$) using the operator norm.

Recall the $GL(n, \mathbb{C})$ is the group of $n \times n$ invertible matrices over \mathbb{C} . Let's look at some examples:

Example 4.8.1.

- $GL(n, \mathbb{C})$, the general linear group over \mathbb{C} .
- $SL(n, \mathbb{C}) = \{A \in GL(n, \mathbb{C}) | \det\{A\} = 1\}$, special linear group over \mathbb{C} .
- $GL(n, \mathbb{R}) = \{A \in GL(n, \mathbb{C}) | \bar{A} - A = 0\}$, general linear group over \mathbb{R} .

- $SL(n, \mathbb{R}) = \{A \in GL(n, \mathbb{R}) \mid \det\{A\} = 1\}$, special linear group over \mathbb{R} .
- $O(n) = \{A \in GL(n, \mathbb{R}) \mid A^\top A = \mathcal{I}\}$, the orthogonal group.
- $SO(n) = \{A \in O(n) \mid \det\{A\} = 1\}$, the special orthogonal group.
- $U(n) = \{A \in GL(n, \mathbb{C}) \mid A^\dagger A = \mathcal{I}\}$, the unitary group.
- $SU(n) = \{A \in U(n) \mid \det\{A\} = 1\}$, the special unitary group.

Note that the last two groups are real Lie groups despite having complex matrices.

Example 4.8.2. Some familiar examples. First we have the circle group:

$$U(1) = \{z \in \mathbb{C} \setminus \{0\} \mid |z|^2 = 1\} \quad (4.40)$$

Second we have the 2D rotational group:

$$SO(2) = \left\{ \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \mid \theta \in [0, 2\pi] \right\}. \quad (4.41)$$

Definition 4.8.2 (Lie Algebra). A real or complex Lie algebra is a vector space V over \mathbb{R} or \mathbb{C} with an operation $[\cdot, \cdot] : V \times V \rightarrow V$ (Lie bracket) satisfying

- $[\alpha x + \beta y, z] = \alpha[x, z] + \beta[y, z]$.
- $[z, \alpha x + \beta y] = \alpha[z, x] + \beta[z, y]$.
- $[y, x] = -[x, y]$.
- $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$.

Definition 4.8.3 (Associated Lie algebra of MLG). Let $G \leq GL(n, \mathbb{C})$ be a matrix Lie group. The associated Lie algebra is the set $\text{Lie}(G) = \{X \in M_n(\mathbb{C}) \mid e^{\epsilon X} \in G \forall \epsilon \in \mathbb{R}\}$ with Lie bracket $[x, y] = xy - yx$.

Note: $\text{Lie}(G)$ is usually denoted \mathfrak{g} .

We need to prove that $\text{Lie}(G)$ is really a Lie algebra:

Proof. We need to show $\text{Lie}(G)$ satisfies all the defining characteristics of a Lie algebra:

- $[\cdot, \cdot]$ is anti-commutative, bilinear, and satisfies Jacobi identity: this is not hard to show.
- The zero matrix belongs to $\text{Lie}(G)$ since $e^{\epsilon O} = \mathcal{I}$, for all $\epsilon \in \mathbb{R}$.
- If $X \in \text{Lie}(G)$ and $\alpha \in \mathbb{R}$ then $e^{\epsilon \alpha X} \in G, \forall \epsilon \in \mathbb{R}$ since $\epsilon \alpha \in \mathbb{R}$. This means $\alpha X \in \text{Lie}(G)$.

- If $X, Y \in \text{Lie}(G)$ then $e^{\epsilon(X+Y)} = \lim_{k \rightarrow \infty} (e^{\epsilon X/k} e^{\epsilon Y/k})^k$ which is a limit of products of elements of G . Since G is closed, the limit converges to something in G . This means $X + Y \in \text{Lie}(G)$.
- This means that $\text{Lie}(G)$ is a subspace of $M_n(\mathbb{C})$ (as a real vector space). Thus $\text{Lie}(G)$ is a Lie algebra.

□

Definition 4.8.4 (Some more notions about MLGs).

- $G \leq GL(n, \mathbb{C})$ is called compact if it is closed and bounded with respect to $M_n(\mathbb{C})$. This means $O(n), SO(n), U(n), SU(n)$ are compact.
- G is connected if for every $g \in G$ there is a continuous path connecting it to the identity $\gamma : [0, 1] \rightarrow G$ with $\gamma(0) = e, \gamma(1) = g$. Note that $O(n)$ is not connected because there's no way to connect something to determinant -1 to the identity which has determinant 1 .
- G is simply connected if it is connected and every loop can be shrunk to a single point (no holes). Example: $U(1)$ is not simply connected.
- If $X \in \text{Lie}(G)$ then $\{e^{\epsilon X} | \epsilon \in \mathbb{R}\}$ is a one-parameter subgroup of G . In fact, every (smooth) one-parameter subgroup of G is of this form. We say that X generates the subgroup $\{e^{\epsilon X} | \epsilon \in \mathbb{R}\}$.
- If G is connected, every group element $g \in G$ can be written as $g = e^{X_1} \dots e^{X_n}$ with $X_i \in G$. We say that G is generated by $\text{Lie}(G)$.

Definition 4.8.5 (MLG homomorphism). A MLG homomorphism is a continuous group homomorphism $\phi : G \rightarrow H$.

Definition 4.8.6 (MLG isomorphism). A MLG isomorphism is an invertible continuous homomorphism.

Example 4.8.3. $\phi : z \in U(1) \mapsto \begin{bmatrix} \text{Re}(z) & \text{Im}(z) \\ -\text{Im}(z) & \text{Re}(z) \end{bmatrix} \in SO(2)$ is an isomorphism.

4.9 Representation theory

Definition 4.9.1 (Representations). Let G be a MLG. A complex vector space V is a G -module if there is a continuous action of G on V , i.e., there exists a map $G \times V \rightarrow V$ that is continuous, such that

- $g(\alpha |\psi\rangle + \beta |\phi\rangle) = \alpha g |\psi\rangle + \beta g |\phi\rangle.$
- $g(h |\psi\rangle) = (gh) |\psi\rangle.$

Notes:

- If V is infinite dimensional we have to be careful about its topology.
- The action of G on V is also called a representation of G on V (same thing, different point of view).

Example 4.9.1. Let $V = \text{span}\{|n\rangle\}$ with the action of $U(1)$. Then $e^{i\theta} \triangleleft |n\rangle = e^{in\theta} |n\rangle, n \in \mathbb{N}$.

The action is extended to the other vectors by linearity.

Definition 4.9.2 (Unitarity). A G -module V is unitary if it has an inner product and $\overline{\langle \phi | g \triangleleft | \phi \rangle} = \langle \phi | g^{-1} \triangleleft | \phi \rangle$.

Example 4.9.2. Define an inner product on V_n as $\langle n | n \rangle = 1$. For the action of $U(1)$ we get $\overline{\langle n | e^{i\theta} \triangleleft | n \rangle} = \overline{\langle n | e^{in\theta} | n \rangle} = e^{-in\theta}$ and $\langle n | (e^{i\theta})^{-1} \triangleleft | n \rangle = e^{-in\theta}$ so V_n is unitary.

Example 4.9.3. For the action of $GL(1, \mathbb{C})$ we get $\langle n | z^{-1} \triangleleft | n \rangle = z^{-n}$, but $\langle n | z \triangleleft | n \rangle = \bar{z}^n \neq z^{-n}$ in general. So V_n here is not unitary.

Definition 4.9.3 (Irreducibility). V is irreducible if the only invariant subspaces ($W \subseteq V$ closed such that $g|\psi\rangle \in W, \forall g \in G, \forall |\psi\rangle \in W$) are V and $\{0\}$.

Definition 4.9.4 (Lie algebras). Let \mathfrak{g} be a Lie algebra. A complex vector space V is a \mathfrak{g} -module if there is an action of \mathfrak{g} on V such that

- $X \triangleleft (\alpha |\psi\rangle + \beta |\phi\rangle) = \alpha X \triangleleft |\psi\rangle + \beta X \triangleleft |\phi\rangle.$
- $(\alpha X + \beta Y) \triangleleft |\psi\rangle = \alpha X \triangleleft |\psi\rangle + \beta Y \triangleleft |\psi\rangle$
- $[X, Y] \triangleleft |\psi\rangle = X \triangleleft (Y \triangleleft |\psi\rangle) - Y \triangleleft (X \triangleleft |\psi\rangle).$

In this case, V is unitary if $\overline{\langle \psi | X \triangleleft | \phi \rangle} = -\langle \phi | X \triangleleft | \psi \rangle$.

If V is a G -module, we can make it into a $\text{Lie}(G)$ -module by defining $X |\psi\rangle = (d/d\epsilon) e^{\epsilon X} |\psi\rangle \Big|_{\epsilon=0}$.

Example 4.9.4.

example: $U(1) = \{ i\theta \mid \theta \in \mathbb{R} \} = \text{span}\{X\}$, $X = i$

using $V = \text{span}\{|n\rangle\}$ with $e^{i\theta}|n\rangle = e^{in\theta}|n\rangle$

we get $|X|n\rangle = \frac{d}{d\theta} e^{in\theta}|n\rangle \Big|_{\theta=0} = in|n\rangle$

Figure 4.1: Giuseppe's notes

example: $\underline{\text{su}(2)} = \{ X \in M_n(\mathbb{C}) \mid X^* = -X, \text{tr}(X) = 0 \} = \text{span}\{X_1, X_2, X_3\}$

$$\text{where } X_1 = \frac{1}{2} \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \quad X_2 = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad X_3 = \frac{1}{2} \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

are a basis. Since $[\cdot, \cdot]$ is bi-linear, we only need to know

Figure 4.2: Giuseppe's notes

Example 4.9.5.

4.10 Symmetries in quantum mechanics

Theorem 4.10.1 (Wigner's Theorem). If $T : H \rightarrow H$ is an invertible transformation of an Hilbert space into itself that preserves transition amplitudes

$$\frac{|\langle T(\psi), T(\phi) \rangle|^2}{\|T\psi\|^2 \|T\phi\|^2} = \frac{|\langle \psi, \phi \rangle|^2}{\|\phi\|^2 \|\psi\|^2} \quad (4.42)$$

for all $\phi, \psi \in H$ then one of the following happens:

- T is linear and unitary (up to a multiplicative constant)
- T is anti-linear and anti-unitary (up to a multiplicative constant).

Since symmetries, at the very least, preserve transition amplitudes, they should act as unitary or anti-unitary operators.

Since the identity is unitary, if G is a connected group of symmetries, by continuity G must act unitarily.

Note: Anti-unitary ones are used for time reversal, and all these sound like unitary representations!

Now, back to representations, if V is a G -module, we can make it into a $\text{Lie}(G)$ -module by defining

$$X \triangleleft |\psi\rangle = \frac{d}{d\epsilon} (e^{\epsilon X} \triangleleft |\psi\rangle) \Big|_{\epsilon=0} = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} (e^{\epsilon X} \triangleleft |\psi\rangle - e^{0X} \triangleleft |\psi\rangle). \quad (4.43)$$

Example 4.10.1. If $U(1)$ acts on $V_n = \text{span}\{|n\rangle\}$ as $e^{i\theta} \triangleleft |n\rangle = e^{in\theta} |n\rangle$ and $i = (d/d\epsilon)|_{\epsilon=0} e^{i\epsilon} \in \mathfrak{u}(1)$, then

$$i \triangleleft |n\rangle = \dots = i n |n\rangle \quad (4.44)$$

which is indeed an action of $\mathfrak{u}(1)$. Note that the action of $\mathfrak{u}(1)$ is anti-hermitian. So we have a $\mathfrak{u}(1)$ -module.

Now, we ask, Does the converse work? In other words, do all the representations of $\text{Lie}(G)$ come from representations of G ? Well, suppose we have a new class of representations $V_a = \text{span}\{|a\rangle\}$, $a \in \mathbb{R}$ with the action $U(1) : i \triangleleft |a\rangle = ia |a\rangle$. This is a unitary $\mathfrak{u}(1)$ -module, since

$$\overline{\langle a | i \triangleleft |a\rangle} = \dots = -\langle a | i \triangleleft |a\rangle, \quad (4.45)$$

which we see is anti-hermitian. Can we say that

$$e^{i\theta} \triangleleft |a\rangle = \text{"exp } [i\theta \triangleleft |a\rangle]''? \quad (4.46)$$

which is the reverse of $i\theta \triangleleft |n\rangle = (d/d\epsilon)|_{\epsilon=0}e^{\epsilon X} \triangleleft |a\rangle$. Suppose we say that

$$e^{i\theta} \triangleleft |a\rangle = \sum_{n=0}^{\infty} \frac{(i\theta)^n}{n!} \triangleleft |a\rangle. \quad (4.47)$$

With this, we can see that

$$\sum_{n=0}^{\infty} \frac{(i\theta)^n}{n!} \triangleleft |a\rangle = \left(\sum_{n=0}^{\infty} \frac{(ia\theta)^n}{n!} \right) |a\rangle = e^{i\theta a} |a\rangle. \quad (4.48)$$

So can we say

$$e^{i\theta} \triangleleft |a\rangle = e^{i\theta a} |a\rangle? \quad (4.49)$$

The answer is NO. One thing we did that doesn't make sense is this; The mistake is the ambiguity of

$$e^{i\theta} \triangleleft |a\rangle = \exp[i\theta \triangleleft |a\rangle] \quad (4.50)$$

because the choice of θ is ambiguous $e^{i\theta} = e^{i\theta + i2\pi k}$. To make this clear, we need to make a choice, say $\theta \in (-\pi, \pi]$. So, we must say that if $z \in U(1), z \triangleleft |a\rangle = e^{ia \arg(z)} |a\rangle$.

Now, all seems well. Is this a $U(1)$ -module?

- It is linear.
- The problem arises in this next item because $\arg(z) + \arg(w) = \arg(z + w) + 2\pi k$:

$$e^{ia(\arg(z)+\arg(w))} = e^{ia(\arg(z)+\arg(w))} e^{i2\pi ak}. \quad (4.51)$$

This last term is not one in general. If $a = 1/2$, then we end up with

$$e^{i\pi} \triangleleft e^{i\pi} s \triangleleft |1/2\rangle = -|1/2\rangle \neq e^{i\pi} e^{i\pi} |a\rangle = |1/2\rangle. \quad (4.52)$$

Now, is everything broken? We see that whether things work depends on the value of a . It turns out that NOTHING has been broken! Remember that physical states are only defined up to a non-zero scalar ($|\psi\rangle \sim \lambda |\psi\rangle$) if $\lambda \neq 0$. We have equivalence classes!

$$[e^{i\pi} \triangleleft e^{i\pi} \triangleleft |1/2\rangle] = [-|1/2\rangle] = [|1/2\rangle] = [e^{i2\pi} \triangleleft |1/2\rangle]. \quad (4.53)$$

And thus we have in general, $z \triangleleft z \triangleleft |a\rangle = (zw) \triangleleft |a\rangle$. We call these equivalence classes *projective representations*. In general, we will have that

$$g \triangleleft h \triangleleft |\psi\rangle = e^{i\omega(g,h)} (gh) \triangleleft |\psi\rangle \quad (4.54)$$

still describes a symmetry.

Proposition 4.10.1 (Nice things that happen). Suppose V is finite-dimensional.

- The projective representations of G are all obtained by “exponentiating” pure representations of $\text{Lie}(G)$ (G connected) – pure in the sense that they are not projective.
 - Pure representations of $\text{Lie}(G)$ instead of projective representations of G .
 - Representations of $\text{Lie}(G)$ are easier to find than representations of G .
- Projective representations of G are in “one-to-one” correspondence with pure representations of \tilde{G} where (\tilde{G} is a simply connected cover of G).

Example 4.10.2. $\tilde{SO}(3) = SU(2)$.

Note that simply connected covers are unique. With this, instead of looking at projective representations of $SO(3)$, we can look at representations of $SU(2)$. The pure representations of $SO(3)$ are labeled by integers, while the projective representations of $SO(3)$ are labeled by half-integers. These concepts are related to angular momentum.

The idea is that, in QM, in most of what we do is related to representation theory.

4.11 Exercises

1. Prove the following consequences of the definition of a group:

1. the identity element is unique (if two elements satisfy the identity property, they are necessarily equal).

Proof. If e, e' are identities of G , then $e = e'e = e'e = e'$. \square

2. For each $g \in G$ the inverse g^{-1} is unique.

Proof. Multiplying both sides on the right by any inverse to get $g^{-1} * g = g'^{-1} * g = e \implies g^{-1} = g'^{-1}$. \square

3. $(g^{-1})^{-1} = g$.

Proof. $g^{-1} * (g^{-1})^{-1} = e = g^{-1} * g \implies g = (g^{-1})^{-1}$, by multiplying by g^{-1} on the right on both sides. \square

4. $(gh)^{-1} = h^{-1}g^{-1}$.

Proof. Multiplying both sides by gh we get $(gh)(gh)^{-1} = e = ghh^{-1}g^{-1}$, so we're done, by the previous items. \square

2. Prove that if $\varphi : G \rightarrow H$ is a group homomorphism then

1. $\varphi(e_G) = e_H$. (Hint: look at $\varphi(e_G e_G)$).

Proof. $\varphi(e_G) = \varphi(e_G e_G) = \varphi(e_G)\varphi(e_G) \implies \varphi(e_G) = e_H$ (we just showed that $\varphi(e_G)$ is an identity element in H). \square

2. $\varphi(g^{-1}) = \varphi(g)^{-1}, \forall g \in G$.

Proof. $\varphi(e_G) = \varphi(g^{-1}g) = \varphi(g)\varphi(g^{-1}) = \varphi(g)\varphi(g)^{-1}$, so we have that $\varphi(g^{-1}) = \varphi(g)^{-1}$. \square

3. Prove that \mathbb{Z}_2 is isomorphic to the subgroup $\{-\mathbb{I}_n, \mathbb{I}_n\} \leq GL(n, \mathbb{C})$. While you are at it, prove that the latter is indeed a subgroup.

Proof. Showing the latter set is a subgroup is easy so I won't do it. Consider the group homomorphism $\varphi : \mathbb{Z}_2 \rightarrow \{-\mathbb{I}_n, \mathbb{I}_n\}$ defined by $\varphi(0) = \mathbb{I}_n$, and $\varphi(1) = -\mathbb{I}_n$. The map is bijective by definition. By inspection, the map is also a group homomorphism. So, the map is an isomorphism. This means the two groups are isomorphic to each other. \square

4. I'll do you one better: prove that any group with only two elements is isomorphic to \mathbb{Z}_2 .

Proof. Call the two-element group $G = \{g, g^{-1}\}$. Create a map just like that in the previous exercise and we're done. \square

5. Show that $2\mathbb{Z} = \{2n | n \in \mathbb{Z}\}$ is a normal subgroup of $(\mathbb{Z}, +)$ and that $\mathbb{Z}_2 = \mathbb{Z}/2\mathbb{Z}$.

Proof. This follows almost directly from the definitions. \square

6. Prove the first isomorphism theorem.

Proof. Refer to the Abstract Algebra [notes](#) for the proof. The first two items are easy. The last item is also easy but requires a trick. $\ker \varphi$ is normal, so

$$h^{-1}g \in \ker \varphi \iff \varphi(h^{-1}g) = \varphi(h)^{-1}\varphi(g) = e_H \implies \varphi(g) = \varphi(h). \quad (4.55)$$

This means that

$$[g] = \{h \in G | \varphi(h) = \varphi(g)\}. \quad (4.56)$$

With these we can prove the third item. \square

7. **Noether's Theorem.** Consider the Lagrangian

$$\mathcal{L}(q, \dot{q}, t) = \frac{1}{2} (m\dot{q}^2 - kq^2) e^{\alpha t}, \quad \alpha \in \mathbb{R} \quad (4.57)$$

with the transformations

$$\tilde{t} = T_\epsilon(t) = t + \epsilon, \quad \tilde{q} = Q_\epsilon(q) = q e^{-\epsilon\alpha/2}. \quad (4.58)$$

- Show that T_ϵ, Q_ϵ are one-parameter subgroups.

Proof. Let's not worry about this. We only need to verify that these transformations satisfy the defining characteristics of one-parameter subgroups. Remember that one-parameter subgroups satisfy the following properties:

- $g_0 = e_G$
- $g_{-\epsilon} = g_\epsilon^{-1}$
- $g_{\epsilon+\epsilon'} = g_\epsilon g_{\epsilon'}$. This means the the one-parameter subgroup is abelian.

\square

- Show that $\delta S = (d/d\epsilon)|_{\epsilon=0} \tilde{S}[\tilde{q}] = 0$ without using the Euler-Lagrange equations, using the fact that

$$\delta S = \int_a^b dt \left[\mathcal{L}(q, \dot{q}, t) \frac{d}{dt} \delta t + \frac{\partial \mathcal{L}}{\partial t} \delta t + \frac{\partial \mathcal{L}}{\partial q} \delta q + \frac{\partial \mathcal{L}}{\partial \dot{q}} \left(\frac{d}{dt} \delta q - \dot{q} \frac{d}{dt} \delta t \right) \right] \quad (4.59)$$

and substitute $\mathcal{L}, \delta t, \delta q$.

Proof. By definition we get $\delta t = 1$ and $\delta q = -\alpha q/2$. With these we get

$$\delta S = \int_a^b dt \left[\frac{\partial \mathcal{L}}{\partial t} + \frac{-\alpha q}{2} \frac{\partial \mathcal{L}}{\partial q} + \frac{\partial \mathcal{L}}{\partial \dot{q}} \left[\dot{q} \frac{-\alpha}{2} - 0 \right] \right] = 0. \quad (4.60)$$

□

- Find the associated conserved quantity.

Proof. The associated conserved quantity is given, by definition, as

$$\mathcal{L}\delta t + \frac{\partial \mathcal{L}}{\partial \dot{q}}(\delta q - \dot{q}\delta t) \quad (4.61)$$

A simple computation gives the result. □

- Find the Euler-Lagrange equations

Proof. Remember that the Euler-Lagrange equations are given by

$$\frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}} = 0. \quad (4.62)$$

Simple computations give us the EOM. □

- Show that the conserved quantity is indeed conserved if the Euler-Lagrange equations hold.

Proof. This is also a simple verification, so I won't worry about it. □

8. Find out the explicit form and dimension of the Lie algebras of the following groups: $GL(n, \mathbb{R})$, $SL(n, \mathbb{R})$, $SL(n, \mathbb{C})$, $O(n)$, $SO(n)$, $SU(n)$.

Proof. For example, $U(n) = \{A \in GL(n, \mathbb{C}) | A^\dagger A = 1\}$. Then $\mathfrak{U}(n) = \{X \in M_n(\mathbb{C}) | e^{\epsilon X} \in U(n) \forall \epsilon \in \mathbb{R}\}$. Using properties of matrix exponentials, we find that $\mathfrak{U}(n) = \{X \in M_n(\mathbb{C}) | X^\dagger = -X\}$.

Example: $U(n) = \{A \in GL(n, \mathbb{C}) | A^\dagger A = 1\}$

$$\Rightarrow \mathfrak{U}(n) = \{X \in M_n(\mathbb{C}) | e^{\epsilon X} \in U(n) \forall \epsilon \in \mathbb{R}\}$$

$$e^{\epsilon X} \in U(n) \forall \epsilon \in \mathbb{R} \Leftrightarrow (e^{\epsilon X})^\dagger e^{\epsilon X} = 1 \quad \forall \epsilon \in \mathbb{R} \Leftrightarrow e^{\epsilon X^\dagger} e^{\epsilon X} = 1 \quad \forall \epsilon \in \mathbb{R}$$

$$\Leftrightarrow e^{\epsilon X^\dagger} = e^{-\epsilon X} \quad \forall \epsilon \in \mathbb{R} \Leftrightarrow X^\dagger = -X$$

$$\Rightarrow \mathfrak{U}(n) = \{X \in M_n(\mathbb{C}) | X^\dagger = -X\}$$

Figure 4.3: From Giuseppe's Notes

As a real vector space, Let $X = A + iB$, then $X^\dagger = -X$ if and only if $A^\top = -A$, $B^\top = B$. This means the dimension is $n(n+1)/2 + n(n-1)/2 = n^2$.

Some useful properties of the matrix exponential:

$\cdot (e^x)^{-1} = e^{-x}$	$\cdot \overline{(e^x)} = e^{\bar{x}}$
$\cdot (e^x)^T = e^{x^T}$	$\cdot \det(e^x) = e^{\text{tr}(x)}$
$\cdot (e^x)^* = e^{x^*}$	$\cdot e^{\epsilon x} = e^{\epsilon y} \quad \forall \epsilon \in \mathbb{R} \Leftrightarrow x = y$

✓ proof: $\begin{cases} \text{trivial} \\ \text{take derivative on both sides and evaluate at } \epsilon = 0 \end{cases}$

Figure 4.4: From Giuseppe's notes

With this we can proceed to the other examples. It's useful to remember the following facts in the figure.

□

9. Do the same thing for

- The complex special orthogonal group: $SO(n, \mathbb{C})$.
- Indefinite special orthogonal group: $SO(p, q)$.
- Real symplectic group: $Sp(2n, \mathbb{R})$.
- Complex symplectic group: $Sp(2n, \mathbb{C})$.
- Compact symplectic group: $Sp(n) = Sp(2n, \mathbb{C}) \cap U(2n)$.

$$\begin{aligned} SO(n, \mathbb{C}) &= \left\{ A \in SL(n, \mathbb{C}) \mid A^T A = \mathbb{1} \right\} \text{ complex special orthogonal group} \\ SO(p, q) &= \left\{ A \in SL(p+q, \mathbb{R}) \mid A^T \begin{pmatrix} \mathbb{1}_p & 0 \\ 0 & -\mathbb{1}_q \end{pmatrix} A = \begin{pmatrix} \mathbb{1}_p & 0 \\ 0 & -\mathbb{1}_q \end{pmatrix} \right\} \text{ indefinite special orthogonal group} \\ Sp(2n, \mathbb{R}) &= \left\{ A \in SL(2n, \mathbb{R}) \mid A^T \begin{pmatrix} 0 & \mathbb{1}_n \\ -\mathbb{1}_n & 0 \end{pmatrix} A = \begin{pmatrix} 0 & \mathbb{1}_n \\ -\mathbb{1}_n & 0 \end{pmatrix} \right\} \text{ real and complex} \\ Sp(2n, \mathbb{C}) &= \left\{ A \in SL(2n, \mathbb{C}) \mid A^T \begin{pmatrix} 0 & \mathbb{1}_n \\ -\mathbb{1}_n & 0 \end{pmatrix} A = \begin{pmatrix} 0 & \mathbb{1}_n \\ -\mathbb{1}_n & 0 \end{pmatrix} \right\} \text{ symplectic group} \\ Sp(n) &= Sp(2n, \mathbb{C}) \cap U(2n) \quad \text{compact symplectic group} \end{aligned}$$

Figure 4.5: From Giuseppe's notes

Proof. For example,

□

example: $U(p, q) = \{A \in GL(p+q, \mathbb{C}) \mid A^* \begin{pmatrix} 1_p & 0 \\ 0 & -1_q \end{pmatrix} A = \begin{pmatrix} 1_p & 0 \\ 0 & -1_q \end{pmatrix}\}$

$$\cdot e^{\varepsilon A^*} \begin{pmatrix} 1_p & 0 \\ 0 & -1_q \end{pmatrix} e^{\varepsilon A} = \begin{pmatrix} 1_p & 0 \\ 0 & -1_q \end{pmatrix} \quad \forall \varepsilon \in \mathbb{R} \quad \textcolor{blue}{\checkmark Ce^{A^*} = e^{CAC^*}}$$

$$\Leftrightarrow e^{\varepsilon A^*} = \begin{pmatrix} 1_p & 0 \\ 0 & -1_q \end{pmatrix} e^{-\varepsilon A} = \exp \left[-\varepsilon \begin{pmatrix} 1_p & 0 \\ 0 & -1_q \end{pmatrix} A \begin{pmatrix} 1_p & 0 \\ 0 & -1_q \end{pmatrix} \right] \quad \forall \varepsilon \in \mathbb{R}$$

$$\Leftrightarrow A^* = - \begin{pmatrix} 1_p & 0 \\ 0 & -1_q \end{pmatrix} A \begin{pmatrix} 1_p & 0 \\ 0 & -1_q \end{pmatrix} \quad \text{or} \quad A^* \begin{pmatrix} 1_p & 0 \\ 0 & -1_q \end{pmatrix} = - \begin{pmatrix} 1_p & 0 \\ 0 & -1_q \end{pmatrix} A$$

$\xrightarrow{\text{pre}}$ $\xrightarrow{\text{pre}}$

if we write A in block form $A = \begin{pmatrix} X & Y \\ Z & W \end{pmatrix}$ we get

$$\begin{pmatrix} X^* & Z^* \\ Y^* & W^* \end{pmatrix} \begin{pmatrix} 1_p & 0 \\ 0 & -1_q \end{pmatrix} = - \begin{pmatrix} 1_p & 0 \\ 0 & -1_q \end{pmatrix} \begin{pmatrix} X & Y \\ Z & W \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} X^* & -Z^* \\ Y^* & -W^* \end{pmatrix} = \begin{pmatrix} -X & -Y \\ Z & W \end{pmatrix} \quad \Rightarrow \quad \begin{array}{l} X^* = -X \rightarrow p^2 \text{ independent real entries} \\ Y^* = -W \rightarrow q^2 \text{ independent real entries} \\ Z^* = Y \rightarrow 2pq \text{ independent real entries} \end{array}$$

$$\Rightarrow U(p, q) = \left\{ \begin{pmatrix} X & Y \\ Y^* & W \end{pmatrix} \mid X \in U(p), W \in U(q), Y \in M_{p \times q}(\mathbb{C}) \right\}$$

$$\dim(U(p, q)) = (p+q)^2$$

Figure 4.6: From Giuseppe's notes

Part 5

Research Project: Quantum Simulation

Quantum many-body physics on quantum hardware

Project description: Recently, there have been significant advances in several quantum computing platforms, including superconducting qubits and trapped ions. At this point, nontrivial quantum operations can be implemented on tens of qubits, and the frontier continues to expand. This project will explore how these emerging technologies can assist the realization and understanding of complex many-body quantum systems, regarding both static aspects such as ground state properties and dynamical aspects such as quantum chaos and thermalization. What new physics can we learn from near-term quantum computers? The ideal outcome is an interesting and realistic proposal that can be carried out in one of the quantum platforms (see for example the interplay between theoretical proposals and experiment).

5.1 Review: The density operator

We often use the language of state vectors in quantum mechanics. An alternative language is that of *density matrices/density operators*. We will use this language extensively in quantum information/quantum computation. First, we will introduce the formulation. Second, we look at some properties of the density operator. Finally, we look at an application where the density operator really shines – as a tool for describing *individual subsystems* of a composite quantum system.

5.1.1 Ensembles of Quantum States

The density operator language provides a convenient means for describing quantum systems whose state is not completely known. Suppose a quantum system is in one of the states $|\psi_i\rangle$ where i is an index, with respective probabilities p_i . We call $\{p_i, |\psi_i\rangle\}$ an *ensemble of pure states*.

Definition 5.1.1 (Density Operator).

$$\boxed{\rho \equiv \sum_i p_i |\psi_i\rangle \langle \psi_i|} \quad (5.1)$$

Suppose we let a unitary \mathcal{U} act on a closed system. If the system was initially in the state $|\psi_i\rangle$ with probability p_i then we get $\mathcal{U}|\psi_i\rangle$ with probability p_i . The density operator evolves as

$$\rho = \sum_i p_i |\psi_i\rangle \langle \psi_i| \xrightarrow{\mathcal{U}} \sum_i p_i \mathcal{U} |\psi_i\rangle \langle \psi_i| \mathcal{U}^\dagger = \mathcal{U} \rho \mathcal{U}^\dagger. \quad (5.2)$$

Suppose we have a measurement gate \mathcal{M}_m . If the initial state is $|\psi_i\rangle$ then the probability of getting result m is

$$p(m|i) = \langle \psi_i | \mathcal{M}_m^\dagger \mathcal{M}_m |\psi_i\rangle = \text{tr}(\mathcal{M}_m^\dagger \mathcal{M}_m \langle \psi_i | \psi_i \rangle) \quad (5.3)$$

where we have used the identity:

$$\text{tr}(A |\psi\rangle \langle \psi|) = \sum_i \langle i | A |\psi\rangle \langle \psi | i \rangle = \langle \psi | A |\psi\rangle. \quad (5.4)$$

With this, the total probability of measuring m is

$$p(m) = \sum_i p(m|i)p_i = \sum_i p_i \text{tr}(\mathcal{M}_m^\dagger \mathcal{M}_m |\psi_i\rangle \langle \psi_i|) = \text{tr}(\mathcal{M}_m^\dagger \mathcal{M}_m \rho). \quad (5.5)$$

The state after the measurement is thus

$$|\psi_i^m\rangle = \frac{\mathcal{M}_m |\psi_i\rangle}{\sqrt{\langle \psi_i | \mathcal{M}_m^\dagger \mathcal{M}_m |\psi_i\rangle}}. \quad (5.6)$$

The corresponding density operator for this state is thus

$$\rho_m = \sum_i p(i|m) |\psi_i^m\rangle \langle \psi_i^m| = \sum_i p(i|m) \frac{\mathcal{M}_m |\psi_i\rangle \langle \psi_i| \mathcal{M}_m^\dagger}{\langle \psi_i | \mathcal{M}_m^\dagger \mathcal{M}_m | \psi_i \rangle}. \quad (5.7)$$

From probability theory we have $p(i|m) = p(m|i)p_i/p(m)$, so:

$$\rho_m = \sum_i p_i \frac{\mathcal{M}_m |\psi_i\rangle \langle \psi_i| \mathcal{M}_m^\dagger}{\text{tr}(\mathcal{M}_m^\dagger \mathcal{M}_m \rho)} = \frac{\mathcal{M}_m \rho \mathcal{M}_m^\dagger}{\text{tr}(\mathcal{M}_m^\dagger \mathcal{M}_m \rho)}. \quad (5.8)$$

Finally, suppose we have a quantum system in state ρ_i with probability p_i . The system might be described by the density matrix

$$\rho = \sum_i p_i \rho_i. \quad (5.9)$$

Here's why:

$$\rho = \sum_{i,j} p_i p_{ij} |\psi_{ij}\rangle \langle \psi_{ij}| = \sum_i p_i \rho_i \quad (5.10)$$

where we have used the definition

$$\rho_i = \sum_j p_{ij} |\psi_{ij}\rangle \langle \psi_{ij}|. \quad (5.11)$$

We call ρ the *mixture* of the states ρ_i with probabilities p_i . This concept of mixture comes up repeatedly in the analysis of problems like quantum noise, where the effect of the noise is to introduce ignorance into our knowledge of the quantum state. A simple example is provided by the measurement scenario. Imagine that for some reason out record of the result m of the measurement was lost. We would have a quantum system in the state ρ_m with probability $p(m)$, but would no longer know the actual value of m . The state of such a quantum system would therefore be described by the density operator

$$\rho = \sum_m p(m) \rho_m = \sum_m \text{tr}(\mathcal{M}_m^\dagger \mathcal{M}_m \rho) \frac{\mathcal{M}_m \rho \mathcal{M}_m^\dagger}{\text{tr}(\mathcal{M}_m^\dagger \mathcal{M}_m \rho)} = \sum_m \mathcal{M}_m \rho \mathcal{M}_m^\dagger. \quad (5.12)$$

5.1.2 General Properties of the Density Operator

Theorem 5.1.1 (Characterization of density operators). An operator ρ is the density operator to some ensemble $\{p_i, |\psi_i\rangle\}$ if and only if it satisfies the conditions

- **Trace condition:** $\text{tr}(\rho) = 1$.
- **Positivity:** ρ is a positive operator, i.e., $\langle \varphi | \rho | \varphi \rangle \geq 0 \forall \varphi$.

With this definition we can reformulate the postulates of quantum mechanics in the language of density operators as

- **Postulate 1:** Associated to any isolated physical system is a Hilbert space known as the *state space* of the system. The system is completely described by its *density operator*, which is a positive operator ρ with trace one, acting on the state space of the system. If a quantum system is in the state ρ_i with probability p_i then the density operator for the system is $\sum_i p_i \rho_i$.
- **Postulate 2:** The evolution of a *closed* quantum system is described by a *unitary transformation*:

$$\rho' = \mathcal{U} \rho \mathcal{U}^\dagger. \quad (5.13)$$

- **Postulate 3:** Measurements are described by a collection $\{\mathcal{M}_m\}$ of *measurement operators*. The index m refers to the measurement outcomes that may occur. If the state of the quantum system is ρ immediately before the measurement then the probability that result m occurs is given by

$$p(m) = \text{tr} (\mathcal{M}_m^\dagger \mathcal{M}_m \rho_m) \quad (5.14)$$

and the state of the system after the measurement is

$$\frac{\mathcal{M}_m \rho \mathcal{M}_m^\dagger}{\text{tr} (\mathcal{M}_m^\dagger \mathcal{M}_m \rho)}. \quad (5.15)$$

The measurement operators satisfy the completeness equation

$$\sum_m \mathcal{M}_m^\dagger \mathcal{M}_m = \mathcal{I}. \quad (5.16)$$

- **Postulate 4:** The state space of a composite physical system is the tensor product of the state spaces of the component physical systems. E.g. the joint state of the total system can be written as $\rho_1 \otimes \cdots \otimes \rho_n$.

Theorem 5.1.2 (Pure states). $\text{tr}(\rho^2) \leq 1$. Equality occurs if and only if ρ is a pure state.

Theorem 5.1.3 (Unitary freedom in the ensemble for density matrices). The sets $|\tilde{\varphi}_i\rangle$ and $|\tilde{\psi}_j\rangle$ generate the same density matrix if and only if they are unitarily similar, i.e.,

$$|\tilde{\varphi}_i\rangle = \sum_j u_{ij} |\tilde{\psi}_j\rangle \quad (5.17)$$

where the u_{ij} are entries of a unitary \mathcal{U} .

5.1.3 The Reduced Density Operator

The deepest application of the density operator is as a descriptive tool for *subsystems* of a composite quantum system. Such a description is given by the *reduced density operator*.

Suppose we have physical systems A and B whose state is described by the density operator ρ^{AB} . The reduced density operator for A is then given by

$$\rho^A = \text{tr}_B (\rho^{AB}) \quad (5.18)$$

where tr_B is the *partial trace* over system B , defined by

$$\text{tr}_B (|a_1\rangle\langle a_2| \otimes |b_1\rangle\langle b_2|) \equiv |a_1\rangle\langle a_2| \text{tr}(|b_1\rangle\langle b_2|) = |a_1\rangle\langle a_2| \langle b_2|b_1\rangle \quad (5.19)$$

where $|a_1\rangle, |a_2\rangle$ are any two vectors in the state space of A , and $|b_1\rangle, |b_2\rangle$ are any two vectors in the state space of B .

Example 5.1.1. To understand this construction, consider an example where $\rho^{AB} = \rho \otimes \sigma$ where ρ, σ describe A, B respectively. Then

$$\rho^A = \text{tr}_B (\rho \otimes \sigma) = \rho \text{tr}(\sigma) = \rho \quad (5.20)$$

as expected.

5.2 Quantum Computational Complexity: An Introduction

In this section we look at quantum computational complexity, as introduced in this [paper](#) by John Watrous at the IQC, the University of Waterloo.

5.2.1 Introduction

- P
- NP
- BPP
- PP
- MA
- AM
- SZK
- PSPACE
- EXP
- NEXP
- PL
- NC

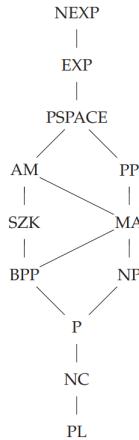


Figure 1: A diagram illustrating known inclusions among most of the classical complexity classes discussed in this paper. Lines indicate containments going upward; for example, AM is contained in PSPACE.

5.3 Quantum Simulation: A Primer

Here's what Feynman said in 1982:

“Can physics be simulated by a universal computer? [...] the physical world is quantum mechanical, and therefore the proper problem is the simulation of quantum physics [...] the full description of quantum mechanics for a large system with R particles [...] has too many variables, it cannot be simulated with a normal computer with a number of elements proportional to R [...] but it can be simulated with] quantum computer elements. [...] Can a quantum system be probabilistically simulated by a classical (probabilistic, I'd assume) universal computer? [...] If you take the computer to be the classical kind I've described so far [...] the answer is certainly, No!”

5.3.1 Simulation in action

The heart of simulation is the solution of differential equations which capture the physical laws governing the dynamical behavior of a system. The goal is generally: given an initial state of the system, what is the state at some other time and/or position? Solutions are usually obtained by *approximating* the state with a digital representation, then *discretizing* the differential equation in space and time such that an iterative application of a procedure carries the state from the initial to the final conditions.

The error in this procedure is *bounded*, and known not to grow faster than some small power of the number of iterations. Furthermore, not all dynamical systems can be simulated efficiently: generally, only those systems which can be

described efficiently can be simulated efficiently.

Simulating quantum systems using classical computers is usually inefficient. The dynamical behavior of many simple quantum systems is governed by the SE:

$$i\hbar\partial_t |\psi\rangle = \mathcal{H} |\psi\rangle. \quad (5.21)$$

We will set $\hbar = 1$. For real particles in space, we often have

$$i\partial_t \psi(x) = \left[-\frac{1}{2m}\partial_x^2 + V(x) \right] \psi(x). \quad (5.22)$$

The key challenge in simulating quantum systems is the *exponential* number of differential equations which must be solved. For 1 qubit evolving according to the SE, a system of 2 differential equations must be solved; for 2 qubits, 4 equations; and for n qubits, 2^n equations. Sometimes, insightful approximations can be made which reduce the effective number of equations involved, but there are many physically interesting quantum systems for which no such approximations are known.

There are many important quantum systems for which classical simulation is intractable. For example:

$$\text{Hubbard model: } \mathcal{H} = \sum_{k=1}^n V_0 n_{k\uparrow} n_{k\downarrow} + \sum_{k,j \text{ neighbors}, \sigma} t_0 c_{k\sigma}^* c_{j\sigma}. \quad (5.23)$$

$$\text{Ising mode: } \mathcal{H} = \sum_{k=1}^n \vec{\sigma}_k \cdot \vec{\sigma}_{k+1}. \quad (5.24)$$

Quantum computers can efficiently simulate quantum systems for which there is no known efficient classical simulation.

5.3.2 The quantum simulation algorithm

The quantum simulation is concerned with the solution of $i\partial_t |\psi\rangle = \mathcal{H} |\psi\rangle$, which, for a time-independent \mathcal{H} , is just

$$|\psi(t)\rangle = e^{-i\mathcal{H}t} |\psi(0)\rangle. \quad (5.25)$$

\mathcal{H} is in general difficult to exponentiate, but a good beginning is the first order solution

$$|(\psi(t + \Delta t))\rangle \approx (\mathcal{I} - i\mathcal{H}\Delta t) |\psi(t)\rangle. \quad (5.26)$$

This is tractable, but often is not very satisfactory.

Efficient approximations to higher order is possible for many *classes* of Hamiltonians. For example, in most physical systems, the Hamiltonian can be written as a sum over many *local* interactions. For a system of n particles:

$$\mathcal{H} = \sum_{k=1}^L \mathcal{H}_k \sim \bigoplus_{k=1}^L \mathcal{H}_k \quad (5.27)$$

where each \mathcal{H}_k acts on at most a constant c number of systems, and L is a polynomial in n . For instance, the terms \mathcal{H}_k are often just 2-body interactions such as $X_i X_j$ and 1-body Hamiltonians such as X_i . Note that both the Hubbard and Ising models have Hamiltonians of this form. There are sometimes additional global symmetry constraints such as particle statistics.

Note that although $e^{-i\mathcal{H}t}$ is difficult to compute, $e^{-i\mathcal{H}_k t}$ acts on a much smaller subsystem, and is straightforward to approximate using quantum circuits. However, because $[\mathcal{H}_j, \mathcal{H}_k] \neq 0$ in general (obviously, they don't commute in general), $e^{-i\mathcal{H}t} \neq \prod_k e^{-i\mathcal{H}_k t}$. So, we might want to ask how $e^{-i\mathcal{H}_k t}$ be useful in constructing $e^{-i\mathcal{H}t}$.

Remarks:

- If the \mathcal{H}_k 's commute, then we do have $e^{-i\mathcal{H}t} = \prod_k e^{-i\mathcal{H}_k t}$.
- The restriction of \mathcal{H}_k to involve at most c particles *implies* that in the sum of the Hamiltonians, L is upper bounded by a polynomial in n .

The heart of quantum simulation algorithms is the following asymptotic approximation theorem:

Theorem 5.3.1 (Trotter formula). Let A and B be Hermitian operators. Then for any real t ,

$$\lim_{n \rightarrow \infty} \left(e^{-iAt/n} e^{iBt/n} \right)^n = e^{i(A+B)t}. \quad (5.28)$$

□

Of course, the theorem holds even if A and B don't commute. For now, we only consider the cases where A, B are Hermitian. But A, B can actually be generators of certain kinds of semigroups, which correspond to general quantum operations.

Using the theorem (and the proof of the theorem, which I won't show here), we can get

$$e^{i(A+B)\Delta t} = e^{iA\Delta t} e^{iB\Delta t} + \mathcal{O}(\Delta t^2). \quad (5.29)$$

Similarly,

$$e^{i(A+B)\Delta t} = e^{iA\Delta t/2} e^{iB\Delta t} e^{iA\Delta t/2} + \mathcal{O}(\Delta t^3). \quad (5.30)$$

Example 5.3.1 (Hamiltonian simulation & Trotterization, taken from [here](#)). The evolution of quantum system is governed by the SE $i\partial_t |\psi\rangle = \mathcal{H}|\psi\rangle$. The solution to the SE:

$$|\psi(t)\rangle = e^{-i\mathcal{H}t} |0\rangle \quad (5.31)$$

is a description on what state a system will be in after a Hamiltonian has been applied to it for a certain period of time. The problem of Hamiltonian simulation is thus states as: given a Hamiltonian \mathcal{H} and an evolution time t , **output a sequence of computational gates** that implement $\mathcal{U} = e^{-i\mathcal{H}t}$.

Hamiltonians are Hermitian operators that are usually a sum of a large number of individual Hamiltonians \mathcal{H}_j . For example, $\mathcal{H} = H_1 + H_2$. This sum of 2 Hamiltonians (which don't necessarily commute) can be described by the Trotter formula or the Lie product formula:

$$e^{-i(H_1+H_2)t} = \lim_{N \rightarrow \infty} \left(e^{-iH_1 t/N} e^{-iH_2 t/N} \right)^N. \quad (5.32)$$

Since the limit of this formula is infinite, we have to truncate the series when implementing this formula on a quantum compute. The truncation introduces error in the simulation that we can bound by ϵ such that $\|e^{-i\mathcal{H}t} - \mathcal{U}\| \leq \epsilon$, where $\|\cdot\|$ is some operator norm. This truncation is known as **Trotterization** and it's widely used to simulate non-commuting Hamiltonians on quantum computers. The **Trotterization formula** is then

$$e^{-i\mathcal{H}t} = \left(e^{-iH_0 t/r} e^{-iH_1 t/r} \dots e^{-iH_{d-1} t/r} \right)^r + \mathcal{O}(\text{polynomial factors}). \quad (5.33)$$

Consider for example the Hamiltonian

$$\mathcal{H} = X_0 + Y_1 + Z_2 \quad (5.34)$$

where X, Y, Z are Pauli matrices and the subscripts label the qubits that the Hamiltonians act on. We can't simulate each Hamiltonian separately because they don't commute. This is why we use Trotterization where we evolve the entire Hamiltonian via repeatedly switching between evolving X_0, Y_1, Z_2 each for a small period of time.

The first step in deriving the circuit that will simulate this Hamiltonian is finding the quantum gates that implement each of its individual term (so that we can make $|\psi(t)\rangle = e^{-i\mathcal{H}t} |0\rangle$ from $|0\rangle$ – in a sense we want to find the correct form for $e^{f(X_0)}$). It turns out that this is easy since the quantum gates

$$R_x(\theta) = e^{-i\theta X/2}, \quad R_y(\theta) = e^{-i\theta Y/2}, \quad R_z(\theta) = e^{-i\theta Z/2} \quad (5.35)$$

will implies the individual terms perfectly. θ here denotes an **angle**, but we can also think of it as **time** because θ specifies the angle by which to rotate the state in a specified axis, which is synonymous with time in the sense that it specifies

how “long” the rotation is applied on the state.

Since we don’t care too much about the simulation error in this example we will use $r = 2, t = 1$ in the Trotter formula to get

$$e^{-i(X_0+Y_1+Z_2)} = \left(e^{-iX_0/2} e^{-iY_1/2} e^{-iZ_2/2} \right) \left(e^{-iX_0/2} e^{-iY_1/2} e^{-iZ_2/2} \right) \quad (5.36)$$

5.3.3 The algorithm

- **Inputs:** (1) A Hamiltonian $\mathcal{H} = \sum_k \mathcal{H}_k$ acting on an N -dimensional system, where each \mathcal{H}_k acts on a small subsystem of size independent of N , (2) an initial state $|\psi_0\rangle$, of the system at $t = 0$, (3) a positive, non-zero accuracy δ , and (3) a time t_f at which the evolved state is desired.
- A state $|\tilde{\psi}(t_f)\rangle$ such that

$$\left| \langle \tilde{\psi}(t_f) | e^{-i\mathcal{H}t_f} |\psi_0 \rangle \right|^2 \geq 1 - \delta. \quad (5.37)$$

- Choose a representation such that the state $|\tilde{\psi}\rangle$ of $n = \text{poly}(\log N)$ qubits approximates the system and the operators $e^{-\mathcal{H}_k \Delta t}$ have efficient quantum circuit approximations. Selection an approximation method and Δt such that the expected error is acceptable, construct the corresponding quantum circuit $\mathcal{U}_{\Delta t}$ for the iterative step, and do the following:

- Initialize state: $|\tilde{\psi}_0\rangle \leftarrow |\psi_0\rangle; j = 0$.
- Iterative update: $\rightarrow |\tilde{\psi}_{j+1}\rangle = \mathcal{U}_{\Delta t} |\tilde{\psi}_j\rangle$.
- Loop: $\rightarrow j = j + 1$; go to the second step until $j\Delta t \geq t_f$.
- Final result: $\rightarrow |\tilde{\psi}(t_f)\rangle = |\tilde{\psi}\rangle$.

Theorem 5.3.2 (Baker-Campbell-Hausdorff formula).

$$e^{(A+B)\Delta t} = e^{A\Delta t} e^{B\Delta t} e^{-(1/2)[A,B]\Delta t^2} + \mathcal{O}(\Delta t^3). \quad (5.38)$$

Proof. The proof is quite easy. Just, by definition, expand things out in power series. \square

\square

5.3.4 An illustrative example

We have seen the case where the Hamiltonian is a sum of local interactions. However, this is not a fundamental requirement. Efficient quantum simulations are possible even for Hamiltonians which act non-trivially on all or nearly all parts of a large system.

Consider the Hamiltonian

$$\mathcal{H} = Z_1 \otimes Z_2 \otimes \cdots \otimes Z_n \quad (5.39)$$

which acts on an n qubit system. This can be simulated efficiently. What we desire is a simple quantum circuit which implements $e^{-i\mathcal{H}\Delta t}$, for arbitrary values of Δt . A circuit that does this is given by

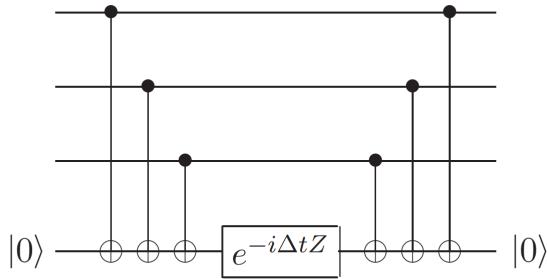


Figure 4.19. Quantum circuit for simulating the Hamiltonian $H = Z_1 \otimes Z_2 \otimes Z_3$ for time Δt .

Even though the Hamiltonian involves all the qubits in the system, it does so in a *classical* manner: the phase shift applied to the system is $e^{-i\Delta t}$ if the *parity* of the n qubits in the computational basis is even. else the phase shift would be $e^{i\Delta t}$. Thus, a simple simulation of \mathcal{H} is possible by first classical computing the parity, then applying the appropriate phase shift conditioned on the parity, then uncomputing the parity.

Extending the same procedure allows us to simulate more complicate extended Hamiltonians. Specifically, we can efficiently simulate any Hamiltonian of the form

$$\mathcal{H} = \bigotimes_{k=1}^n \sigma_{c(k)}^k \quad (5.40)$$

where $\sigma_{c(k)}^k$ is a Pauli matrix (or identity) acting on the k th qubit, with $c(k) \in \{0, 1, 2, 3\}$ specifying one of $\{I, X, Y, Z\}$. The qubits upon which the identity acts can be disregarded, and X or Y terms can be transformed by single qubit gates to Z operations. This leaves a Hamiltonian of the form given above, which can be efficiently simulated.

Using this procedure allows us to simulate a wide class of Hamiltonians containing terms which are not local. It is possible to simulate a Hamiltonian of the form $\mathcal{H} = \sum_{k=1}^L \mathcal{H}_k$ where we only require that the individual \mathcal{H}_k have a tensor product structure, and that L is polynomial in the total number of particles n .

5.3.5 Perspectives on quantum simulation

Quantum simulation is similar to classical methods, but differs fundamentally in many aspects. Each iteration of the quantum algorithm must completely replace the old state with a new one and there is no way to obtain any nontrivial information from an intermediate step without changing the algorithm (due to the no-cloning theorem). Furthermore, the final measurement must be chosen to provide the desired result because it collapses the entire state.

Another difficult problem is the simulation of equilibration processes. A system with Hamiltonian \mathcal{H} in contact with an environment at temperature T will come to thermal equilibrium in a state known as the *Gibbs* state,

$$\rho_{\text{therm}} = \frac{e^{-\mathcal{H}/k_B T}}{\mathcal{Z}} \equiv \frac{e^{-\mathcal{H}/k_B T}}{\text{tr}(e^{-\mathcal{H}/k_B T})} \quad (5.41)$$

where of course \mathcal{Z} is the partition function. We can check that $\text{tr}(\rho) = 1$. For now, we don't know how a quantum computer can simulate this.

Indistinguishability of particles also places a constraint on the state vector of a system which manifests itself in two ways: bosons and fermions. The state vector of a system of bosons remains unchanged under permutation of any two constituents. System of fermions experience a sign change in their state vector under interchange of any two constituents. Both kinds of systems can be simulated efficiently on a quantum computer.

5.4 Problems in Quantum Simulations and Quantum Algorithms

In this section, we will be looking at the following [review](#) on quantum algorithms for quantum chemistry and materials science.

5.4.1 Quantum Simulation: An Overview

Introduction

The interest in quantum computing for quantum simulations of molecules and materials stems from the fact that in many cases, the chemistry and physics of molecules and materials is best described using quantum mechanics. In the worst case, quantum simulation is exponentially hard on classical computers. Quantum computers have advantages over classical computers, but these advantages are problem-specific.

As we have seen before, the natural problem to solve on a quantum computer is the time evolution of a quantum system given some initial state:

$$i\partial_t |\psi(t)\rangle = \mathcal{H} |\psi(t)\rangle. \quad (5.42)$$

This problem, as we have pointed out, is of polynomial cost on a quantum computer, which can offer exponential speedup over classical computers. However, it is necessary to prepare the initial state, which may be difficult. In particular, preparing a low-energy state may be challenging, which naturally leads to considering other important problems:

$$\text{Ground state: } \mathcal{H} |\psi_0\rangle = E_0 |\psi_0\rangle; \quad E_0 = \min_{|\psi\rangle} \langle \psi | \mathcal{H} | \psi \rangle \quad (5.43)$$

$$\text{Thermal averages: } \langle \mathcal{A} \rangle = \frac{\text{tr}(\mathcal{A} e^{-\beta \mathcal{H}})}{\text{tr}(e^{-\beta \mathcal{H}})} \quad (5.44)$$

Ground state determination lies in complexity class QMA, a class of problems *not known to be efficiently solvable in general on a quantum computer*. This means that thermal averages cannot in general be computed efficiently on a quantum computer, since in the limit of zero temperature, this problem reduces to ground-state determination.

To understand quantum advantage in chemistry, condensed matter physics, and quantum materials science, we must be guided by actual empirical data in the form of numerical and theoretical experiments with quantum algorithms and quantum devices on simulation problems of interest. This requires progress, of course.

Current quantum architectures

There's *analog quantum computation* and there's *digital quantum computation*. Analog quantum computation was proposed by Feynman. The idea of analog

quantum computation is to build a lattice of spins with tunable interactions and let it simulate physical systems. The limitation, however, is that this is quite limited because it can only simulate a certain set of physical systems. Furthermore, its accuracy is limited.

Digital quantum computation is more general. Digital quantum computation involves qubits and gates. One can show that with a finite set of gates, one can in principle generate arbitrary quantum evolutions to arbitrary precision. Every problem in the complexity class BQP can be mapped into a quantum program.

The third model of quantum computation is *adiabatic quantum computation*. This model is considered equivalent to the digital version, but is less practical. We won't worry about this.

Building a circuit-based digital quantum computer

The natural enemy of quantum computation is decoherence – the tendency of quantum systems to decay into their classical states. We are current in the *noisy intermediate-scale quantum* (NISQ) era, where qubit technologies (superconducting qubits, ion traps, etc) have reached the point where small devices of a few dozen qubits can be sufficiently isolated from decoherence to execute non-trivial quantum algorithms of a few tens to hundreds of gates. With these devices, an artificial but well-defined problem that is intractable classically is solved on a quantum computer. However, these problems are often not practical. The more important question is when a more relevant problem can be solved on a quantum computer.

To address large-scale problems, *quantum error-correction* (QEC) is necessary. QEC leads to a distinction between *logical* and *physical* qubits. The former are error-corrected and encoded in the state of many physical qubits. Quantum algorithms are performed on the logical qubits, and the error-correction scheme translates the operations on logical qubits into physical operations.

5.4.2 Simulation challenges in molecular and materials science

In this section we look at some scientific problems of relevance to quantum simulation.

Quantum chemistry

Quantum chemistry is concerned with determining the low-lying eigenstates of the electronic Hamiltonian of a molecule. The eigenstates are determined for fixed sets of nuclear positions within the Born-Oppenheimer approximation. Determining the electronic energy as a function of nuclear position is key to understanding chemical reactivity, product distribution, and reaction rates.

Some examples of problems in quantum chemistry include

- The chemistry of enzyme active sites
- Transition metal nanocatalysts and surface catalysts
- Light harvesting and the vision process

The basic metric is whether ground-states or low-energy eigenstate quantum algorithms yield more accurate energies than the best classical algorithms, for the problem sizes of interest.

Quantum molecular spectroscopy

The theoretical goal is to compute the eigenstates of the nuclear SE. There are challenges, of course. First, the nuclear Hamiltonian is often not known, because interactions are mediated by electrons. Second, nuclear rotational and vibrational motion is often far from harmonic and not well approximated by simple mean-field theories. Third, even when the nuclear SE has been properly formulated, one faces the challenge of representing the eigenstates.

Some famous examples include:

- Spectra of floppy molecules
- Hydrogen bonded clusters

There are similarities with the quantum chemistry problems, but the differences are more significant. The Hamiltonian is often more complicated because interactions are now many-body. Also, the number of states is often orders-of-magnitude greater.

Chemical quantum dynamics

Chemical quantum dynamics is concerned with modeling time-dependent electronic and nuclear quantum effects in molecules. Some problems include:

- Proton coupled electron transfer
- Vibrational dynamics in complex environments
- Plasmonic chemistry

There are multiple practical challenges. Some of this may be viewed as an issue of representation. In addition, the dynamical quantum state involves near continuum degrees of freedom, which makes discretizing the Hilbert space a challenging problem.

Correlated electronic structure in materials

The goal here is to determine low-energy properties of materials such as Mott insulators. Some famous problems include:

- High-temperature superconductivity
- Non-Fermi liquid behavior in heavy fermion compounds and fermionic systems near criticality
- Quantum fluctuations in two-dimensional systems
- Frustrated spin systems

5.4.3 Challenges for quantum algorithms in quantum simulation

Overview of algorithms

In this section, we look at the current status and theoretical and practical challenges to implement quantum algorithms for quantum problems.

The first step in a quantum simulation is to choose a *representation* for the Hamiltonian and the states. We will examine the possibilities for different qubit representations, and open questions.

Quantum ground-state algorithms fall into different classes. QPE is a direct route to nearly exact eigenstate determination, but has been challenging to implement in the near-term era. On the other hand, *variational quantum algorithms* provide the possibility to introduce approximations with adjustable circuit depth, which are determined via optimization, usually implemented in a hybrid-quantum classical loop.

While polynomial time algorithms for *quantum time evolution* have been known for some time, algorithms with optimal asymptotic complexity as well as favorable scaling with error and low prefactors, remain an area of active research.

Qubit representation of many-body systems

To simulate many-body systems on a digital computer (either quantum or classical), the infinite-dimensional Hilbert space of a many-body system has to be truncated. The most direct route then is to define a finite set of *basis functions* and then try to project the exact many-body Hamiltonian onto the chosen basis. The resulting discretized system is then expressed in terms of qubits. A choice of a good representation is important as it may affect the simulation cost dramatically.

1. *Ab initio* electronic structure qubit representations:

This concerns with the electronic structure. Consider the Hamiltonian that appears regularly in chemistry and physics

$$\mathcal{H} = \sum_{i=1}^K -\frac{1}{2} \nabla^2 + V(r_i) + \sum_{1 \leq i < j \leq K} \frac{1}{|r_i - r_j|} \quad (5.45)$$

where K is the number of electrons, V is the potential created by atomic nuclei at a point r , and the last term represents the Coulomb interaction. Each electronic has $r_i \in \mathbb{R}^3$ and spin $\omega_i \in \{\uparrow, \downarrow\}$. So, the total wavefunction (for K electrons) has the form $\psi = \psi(\vec{x}_1, \dots, \vec{x}_K)$ where $\vec{x}_i = (r_i, \omega_i)$. From Fermi statistic, we require that ψ is anti-symmetric under a position

swap.

The first step is to approximate the electronic Hamiltonian with a simpler simulator Hamiltonian. This is usually achieved by truncate the Hilbert space of a single electron to a finite set of basis functions ψ_1, \dots, ψ_N .

Electronic structure simulation algorithms are based on the *first quantization method* ([link](#)). We won't worry about this too much.

The *second quantization method* often results in simpler simulator Hamiltonian and requires few qubits. This methods is particularly well suited for quantum simulation algorithms and has been experimentally demonstrated for small molecules. Given a set of N basis functions ψ_1, \dots, ψ_N , the second quantized simulator Hamiltonian is given by

$$\mathcal{H} = \sum_{p,q=1}^N t_{pq} \hat{c}_p^\dagger \hat{c}_q + \frac{1}{2} \sum_{p,q,r,s=1}^N u_{pqrs} \hat{c}_p^\dagger \hat{c}_q^\dagger \hat{c}_r \hat{c}_s. \quad (5.46)$$

Again, we're not going to worry too much about these.

2. Electronic basis functions:

The discretization of the electronic Hilbert space for a quantum simulation requires balancing two concerns. We need to represent the state with the smallest number of qubits, but also retain maximal Hamiltonian sparsity.

There are two families of basis functions in wide use in quantum chemistry and quantum materials science: atomic orbital *Gaussian bases* and *Plane waves*. Gaussian bases are commonly used in molecular simulations due to their compactness, while plane waves are used in crystalline materials simulations.

In Gaussian bases, linear combinations of Gaussian functions are placed at the nuclear positions. When they are place where the ground-state electron density is highest, they give a compact representation of the wavefunction for bound states. However, the Hamiltonian is not sparse, which leads to high gate counts.

Plane waves offer greater simplicity as the accuracy of the basis is controlled by a single parameter, namely the kinetic energy cutoff. The number of plane waves needed to reach a desired accuracy is larger than the number of Gaussian states, but the Hamiltonian contains fewer terms due to momentum conservation.

The need to expose more sparsity in the Hamiltonian while retaining a reasonable compact wavefunction is an active area of research in both classical and quantum algorithms.

3. Fermions-to-qubit mappings:

Since the basic units of a quantum computer are qubits rather than fermions, any quantum simulation algorithm of fermions employs a suitable encoding of fermionic degrees of freedom into qubits. The standard *Jordan-Wigner* mapping identifies each Fermi mode with a qubit such that the empty and the occupied states are mapped to the qubit basis states $|0\rangle$ and $|1\rangle$ respectively. The Jordan-Wigner mapping/transformation is a transformation that maps spin operators onto fermionic creation and annihilation operators. A natural question is whether the number of

$$\begin{array}{cccccccccc} | \cdot \cdot \cdot \rangle & | \cdot \cdot \bullet \rangle & | \cdot \bullet \cdot \rangle & | \bullet \cdot \cdot \rangle & | \cdot \bullet \bullet \rangle & | \bullet \cdot \bullet \rangle & | \bullet \bullet \cdot \rangle & | \bullet \bullet \bullet \rangle \\ \downarrow & \downarrow \\ | 0,0,0 \rangle & | 0,0,1 \rangle & | 0,1,0 \rangle & | 1,0,0 \rangle & | 0,1,1 \rangle & | 1,0,1 \rangle & | 1,1,0 \rangle & | 1,1,1 \rangle \end{array}$$

FIG. 10. Schematic of the Jordan-Wigner encoding for 3 spin-orbitals. Fermions are represented by blue spheres. Each Fock basis vectors (from the vacuum state, left, to the completely filled state, right) is mapped onto a 3-qubit state, with empty (filled) spin-orbitals corresponding to qubits in 0 (1).

qubits required to express a Fermi system can be reduced by exploiting symmetries such as the particle number conservation or the point group symmetries of molecules. For example, zero temperature simulations often target only one symmetry sector containing the ground state. This motivates the study of *symmetry-adapted fermion-to-qubit mappings*.

Quantum algorithms for ground and excited states

There are many approaches to obtaining ground states or excited states on a quantum computer. State preparation procedures attempt to construct a circuit to prepare a state with as large as possible overlap with the desired eigenstate. One set of such procedures, which includes adiabatic state preparation and quantum imaginary time evolution, uses a prescribe evolution path. An alternative strategy is based on *variational methods*, which are often called variational quantum eigensolvers. Here, the preparation circuit itself is defined via the optimization of the energy with respect to parameters of the circuit.

Given some state with sufficiently large overlap with the desired state, one can perform *quantum phase estimation* (QPE) which simultaneously projects the state onto an eigenstate of the Hamiltonian and obtains an estimate for the energy of this eigenstate. The error, as we have seen, is inversely proportional to the simulation time. The probability of successfully projecting onto the desired state is given by the square overlap of the input state and the desired state, and

it is thus necessary to use some other method (such as the state preparation procedures above) to prepare an input state with sufficient overlap with the desired state.

There are weaknesses to this, of course. While phase estimation allows the deviation of the final state from an exact eigenstate to be systematically reduced, it can require deep circuits with many controlled gates that are challenging for devices with limited coherence and without error correction. Variational methods replace such circuits by a large number of potentially short simulations. However, if one does not measure the energy by phase estimation, but instead by expressing the Hamiltonian as a sum of multi-qubit Pauli operators and measuring the terms individually, the state preparation and measurements must be repeated many times, with the error converging only as the square root of the number of repetitions. Variational methods are also limited by the variational form and ability to solve the associated optimization problem, which may by itself represent a difficult classical optimization.

Preparing ground states along a prescribed path

1. Adiabatic state preparation:

This state preparation relies on the *adiabatic theorem*

Theorem 5.4.1 (Adiabatic Theorem). A physical system remains in its instantaneous eigenstate if a given perturbation is acting on it slowly enough and if there is a gap between the eigenvalue and the rest of the Hamiltonian's spectrum.

In simpler terms, this says that gradually changing conditions allow the system to adapt its configuration, hence the probability density is modified by the process. If the system started in an eigenstate of the initial Hamiltonian, it will end in the *corresponding* eigenstate of the final Hamiltonian.

To make use of this theorem, one chooses a Hamiltonian path $\mathcal{H}(\lambda)$ where $\lambda \in [0, 1]$, such that the ground state of $\mathcal{H}(0)$ is easily prepared, while $\mathcal{H}(1)$ is the Hamiltonian whose ground state one wants to obtain. The system is then evolved under the time-dependent SE

$$i\partial_t |\psi_t\rangle = \mathcal{H}(t/T) |\psi_t\rangle, \quad t \in [0, T]. \quad (5.47)$$

When $T \rightarrow \infty$ and if the spectrum of $\mathcal{H}_{t/T}$ is gapped for all t , the final state is the exact ground state of \mathcal{H} . Away from the adiabatic limit $T \rightarrow \infty$, corrections are necessary.

The main disadvantage of this approach is that it is limited by the $T \rightarrow \infty$ requirement, and the fact that a path without degeneracies must be chosen.

2. *Quantum imaginary-time evolution:*

In classical simulations, one popular approach to prepare (nearly exact) ground-states is imaginary-time evolution, which expresses the ground-state as the long-time limit of the imaginary-time SE $-\partial_t |\Phi(\beta)\rangle = \mathcal{H} |\Phi(\beta)\rangle$:

$$|\psi\rangle = \lim_{\beta \rightarrow \infty} \frac{e^{-\beta\mathcal{H}} |\Phi(\beta)\rangle}{\|e^{-\beta\mathcal{H}} \Phi(\beta)\|}. \quad (5.48)$$

To perform imaginary time evolution on a quantum computer, it is necessary to implement the (repeated) action of the short-imaginary-time propagator $e^{-\Delta\tau\mathcal{H}}$ on a state. Given a Hamiltonian that can be decomposed into geometrically local terms,

$$\mathcal{H} = \sum_m \hat{h}_m \quad (5.49)$$

and a state $|\psi\rangle$ with finite **correlation length** C , the action $e^{-\Delta\tau\hat{h}_i}$ can be generated by a unitary $\mathcal{U} = e^{i\hat{A}}$ acting on $\mathcal{O}(C)$ qubits surrounding those acted on by \hat{h}_i , i.e.,

$$\frac{e^{-\beta\mathcal{H}} |\psi\rangle}{\|e^{-\beta\mathcal{H}} \psi\|} = \mathcal{U} |\psi\rangle = e^{i\hat{A}} |\psi\rangle \quad (5.50)$$

where the coefficients of the Pauli strings in \hat{A} can be determined from local measurements of the quits around \hat{h}_i .

Quantum imaginary time evolution becomes inefficient in terms of the number of measurements and complexity of the operator \hat{A} if the domain C grows to be large along the imaginary time evolution path.

Remark: Correlation length is the distance beyond which the correlation $\langle \psi | \mathcal{O}_i \mathcal{O}_j | \psi \rangle$ decays beyond some ϵ . We note that every gapped Hamiltonian has an exponentially decaying correlation length. Gapped Hamiltonians are those with a nontrivial $\Delta E = E_1 - E_0$ even in the thermodynamic limit.

Variational state preparation and variational quantum eigensolver

States can also be prepared through variational methods. Here, similar to classical variational approaches, one chooses a class of ansatz states for the ground state of the Hamiltonian of interest. Generally speaking, such an ansatz consists of some initial state and a unitary circuit parametrized by some set of classical variational parameters. Applying this circuit to the initial state yields a guess for the ground state, whose energy is then evaluated. This yields an upper bound to the true ground state energy. One then varies the variational

parameters to lower the energy of the ansatz state.

Choosing the class of ansatz states is also a balancing act. On the one hand, the class should contain an accurate approximation to the true ground state of the system. On the other hand, we want a class of circuits that are easily executed on the available quantum computer. Finally, it is important for the classical optimization over the variational parameters to be well-behaved, so as to be able to find low-energy minima.

1. Unitary coupled cluster:

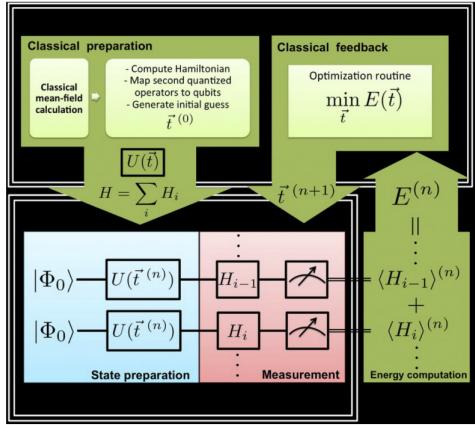


FIG. 13. Workflow of the variational quantum eigensolver algorithm. The classical optimization routine adds expectation values of the Hamiltonian Pauli terms to calculate the energy and estimates new values for the unitary parameters. The process is repeated until convergence. From [172].

We won't worry about the details of how this works at this point. For more information, consider this [paper](#).

2. Adapt-VQE ansatz:

In the adapt-VQE scheme, a collection of operators \hat{A}_i is chosen in advance, and the ground state is approximated by

$$|\psi_{\text{adapt-VQE}}\rangle = e^{\theta_n \hat{A}_n} \dots e^{\theta_1 \hat{A}_1} |\psi_{\text{HF}}\rangle, \quad (5.51)$$

Given a current parameter configuration $\vec{\theta}$, the commutator of the Hamiltonian with each operator in the pool is measured to obtain the gradient of the energy

$$E = \langle \psi_{\text{adapt-VQE}} | \mathcal{H} | \psi_{\text{adapt-VQE}} \rangle. \quad (5.52)$$

wrt the parameters $\vec{\theta}$. Repeating this multiple times and averaging over the obtained samples gives the gradient of the expectation value of the Hamiltonian wrt to the coefficient of each operator. The ansatz is improved by adding the operator \hat{A}_i with the largest gradient to the left end of the ansatz with a new variational parameter, thereby increasing n . The operation is repeated until convergence.

3. Other considerations:

There are other considerations with regards to variational methods, but we won't worry about them too much for now.

Excited states

While much of the above discussion of state preparation and variational algorithms has focused on ground-states, most of the same methods can also be used with minor extensions for excited states. For example, adiabatic state preparation can be used to prepare an excited state, so long as it is connected to the initial state without a vanishing gap.

Phase estimation

We have a section on this, so let's just be brief and gloss over the details.

Quantum Phase Estimation (QPE) is a crucial step in many quantum algorithms. QPE enables high-precision measurements of the ground and excited energy levels. This is done by preparing a trial initial state $|\psi(0)\rangle$ that has a non-negligible overlap with the relevant eigenvector of the target Hamiltonian \mathcal{H} and applying a circuit that creates a superposition of time evolved states $|\psi(t)\rangle = e^{-i\mathcal{H}t} |\psi(0)\rangle$ over a suitable range of evolution times t .

Since QPE is used ubiquitously in a variety of quantum applications, it is crucial to optimize its performance. Below we list some open problems that are being actively investigated:

- Given limitations of near-term quantum devices, of particular interest are tradeoffs between the depth of the QPE circuit and its spectral-resolution power as well as its sensitivity to noise.
- Several methods have been proposed for mitigating experimental errors for VQE-type simulations. Such methods enable reliable estimation of expected values of observables on a given trial state without introducing any overhead in terms of extra qubits or quantum gates.
- Classical post-processing methods that enable simultaneous estimation of multiple eigenvalues are highly desirable.

- Finally, a natural question is whether the time evolution operator $e^{-i\mathcal{H}t}$ in QPE can be replaced by some other functions of \mathcal{H} that are easier to implement

Quantum algorithms for time evolution

1. Hamiltonian simulation problem:

A quantum computer can be programmed to efficiently simulate the unitary time evolution of almost any physically realistic quantum system. Consider any evolution which is of course given by the SE

$$i\partial_t |\psi(t)\rangle = \mathcal{H} |\psi(t)\rangle, \quad t \geq 0. \quad (5.53)$$

We can just assume that \mathcal{H} describes a system of n qubits, since any fermionic or spin system can be mapped to qubits. Integrating we get

$$|\psi(t)\rangle = e^{-it\hat{\mathcal{H}}} |\psi(0)\rangle. \quad (5.54)$$

A quantum algorithm for Hamiltonian simulation takes as input a description of \mathcal{H} , the evolution time t , and outputs a quantum circuit \mathcal{U} that approximates the time evolution operator $e^{-it\mathcal{H}}$ within a specified precision ϵ :

$$\|\mathcal{U} - e^{-it\mathcal{H}}\| \leq \epsilon. \quad (5.55)$$

\mathcal{U} may use ancillary qubits initialized in the $|0\rangle$ state. The simulation cost is usually quantified by the runtime of the algorithm (gate count) and the total number of qubits. Applying \mathcal{U} to $|\psi(0)\rangle$ gives the approximation state $|\psi(t)\rangle$.

Quantum simulation algorithms apply to general classes of Hamiltonians satisfying mild technical conditions that enable a quantum algorithm to access the Hamiltonian efficiently.

5.5 Quantum Phase Estimation (QPE)

One of the most important applications of the Quantum Fourier Transform is the Quantum Phase Estimation. Phase estimation is the key for many quantum algorithms. Suppose a unitary operator \mathcal{U} has an eigenvector $|u\rangle$ with eigenvalue $2^{2\pi i\varphi}$, where the value of φ is unknown. The goal of the phase estimation algorithm is to estimate $0 \leq \varphi \leq 1$ with high probability within additive error ϵ . The algorithm uses $\mathcal{O}(\log(1/\epsilon))$ qubits and $\mathcal{O}(1/\epsilon)$ controlled- \mathcal{U} operations.

To perform the estimation we assume that we have available *oracles* capable of preparing the state $|u\rangle$ (the eigenstate) and performing the controlled- \mathcal{U}^{2^j} operation, for suitable nonnegative integers j . Since preparing eigenstates is not necessarily easy, the phase estimation algorithm is not a complete algorithm by itself. Rather, we should view it as a “subroutine” or a “module.”

The QPE uses two registers. The first register contains t qubits initially in $|0\rangle$. The choice of t depends on two things: the number of digits of accuracy we wish to have in our estimate of φ , and with what probability we wish the phase estimation procedure to be successful. The second register begins in the eigenstate $|u\rangle$, and contains as many qubits as necessary to store $|u\rangle$.

QPE occurs in two main stages. First, the circuit applies an \mathcal{H} to each of the t qubits. Then, the circuit applies controlled- \mathcal{U} operators on the second register, with \mathcal{U} raised to successive powers of two. This stage looks like the following: Recall that immediately after $\mathcal{H}^{\otimes t}$ the state of the upper register looks like

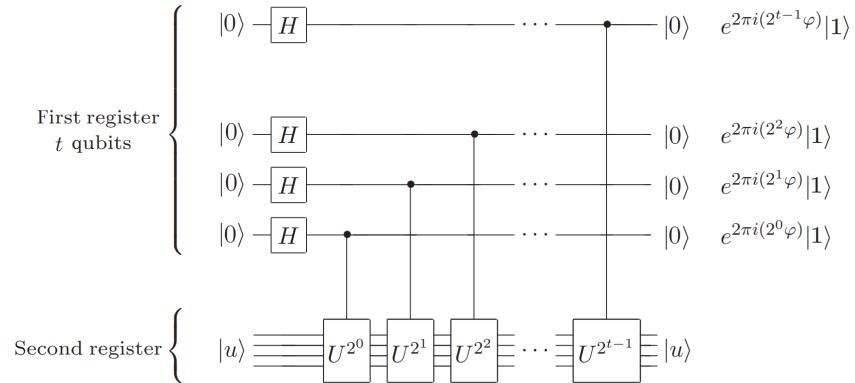


Figure 5.2. The first stage of the phase estimation procedure. Normalization factors of $1/\sqrt{2}$ have been omitted, on the right.

Figure 5.1: Source: Mike and Ike

$$\mathcal{H}|0\rangle = \frac{1}{2^{1/2}}(|0\rangle + |1\rangle) \quad (5.56)$$

The controlled- \mathcal{U} acts on $|u\rangle$ only if the control qubit is in $|1\rangle$. With this, we see that the final state of the first register is given by

$$\begin{aligned} & \frac{1}{2^{t/2}} \left(|0\rangle + e^{2\pi i 2^{t-1}\varphi} |1\rangle \right) \otimes \left(|0\rangle + e^{2\pi i 2^{t-2}\varphi} |1\rangle \right) \otimes \cdots \otimes \left(|0\rangle + e^{2\pi i 2^0\varphi} |1\rangle \right) \\ &= \dots \\ &= \frac{1}{2^{t/2}} \sum_{0 \leq k < 2^t} e^{2\pi i \varphi k} |k\rangle_t. \end{aligned} \quad (5.57)$$

Example 5.5.1. For $t = 2$, we have

$$\begin{aligned} & \frac{1}{2} \left(|0\rangle + e^{4\pi i \varphi} |1\rangle \right) \otimes \left(|0\rangle + e^{2\pi i \varphi} |1\rangle \right) \\ &= \frac{1}{2} \left(|0\rangle_2 + e^{1 \cdot 2\pi i \varphi} |1\rangle_2 + e^{2 \cdot 2\pi i \varphi} |2\rangle_2 + e^{3 \cdot 2\pi i \varphi} |3\rangle_2 \right) \\ &= \frac{1}{2^{2/2}} \sum_{0 \leq k < 2^2} e^{2\pi i \varphi \cdot 2} |k\rangle_2 \end{aligned} \quad (5.58)$$

□

Note that we're leaving out the qubits carrying the $|u\rangle$ state because we don't do anything to it (because $|u\rangle$ is an eigenstate of \mathcal{U}). We should also note that the effect of the sequence of controlled- \mathcal{U} operations is to take $|j\rangle|u\rangle$ to $|j\rangle\mathcal{U}^j|u\rangle$. We should also note that this does not depend on the fact that $|u\rangle$ is an eigenstate of \mathcal{U} .

In the second state, the algorithm applies the inverse quantum Fourier transform on the first register (which is just the QFT, applied in reverse order of operations). This can be done efficiently in $\mathcal{O}(t^2)$ steps, as we have seen the Quantum Information [notes](#). Since the input in this case is the maximal superposition in the t -vector space, the output of the inverse QFT is going to be a single state: $|\tilde{\varphi}\rangle$, which we hope is approximately $|\varphi\rangle$. We will see how close $|\tilde{\varphi}\rangle$ can get to $|\varphi\rangle$.

The third and final stage of the algorithm is to read out the state of the first register by doing a measurement in the computational basis.

Example 5.5.2. To make sure we see what is going on, suppose φ is expressed exactly in t qubits, as $\varphi = 0, \varphi_1\varphi - 2\dots\varphi_t$ (note that $\varphi \in [0, 1]$). Then the state of the input register after the first stage is given by

$$\frac{1}{2^{t/2}} \left(|0\rangle + e^{2\pi i 0 \cdot \varphi_t} |1\rangle \right) \otimes \left(|0\rangle + e^{2\pi i 0 \cdot \varphi_{t-1} \varphi_t} |1\rangle \right) \otimes \cdots \otimes \left(|0\rangle + e^{2\pi i 0 \cdot \varphi_1 \varphi_2 \dots \varphi_t} |1\rangle \right). \quad (5.59)$$

Applying the inverse QFT to this, we see that output state from the second state is the product state $|\varphi_1 \dots \varphi_t\rangle$. A measurement in the computational basis gives φ exactly. □

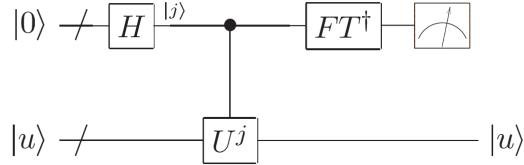


Figure 5.3. Schematic of the overall phase estimation procedure. The top t qubits (the ‘/’ denotes a bundle of wires, as usual) are the first register, and the bottom qubits are the second register, numbering as many as required to perform U . $|u\rangle$ is an eigenstate of U with eigenvalue $e^{2\pi i \varphi}$. The output of the measurement is an approximation to φ accurate to $t - \lceil \log(2 + \frac{1}{2\epsilon}) \rceil$ bits, with probability of success at least $1 - \epsilon$.

Figure 5.2: Source: Mike and Ike

Schematically, the algorithm for QPE looks like the following:
In summary, what we basically did was using the inverse QFT to get

$$\frac{1}{2^{t/2}} \sum_{0 \leq k < 2^t} e^{2\pi i \varphi j} |j\rangle |u\rangle \rightarrow |\tilde{\varphi}\rangle |u\rangle \quad (5.60)$$

where $|\tilde{\varphi}\rangle$ is a good estimator for φ when measured.

5.5.1 Performance and Requirements

What happens when φ cannot be written exactly with a t -bit expansion? We will now show that the procedure we described will produce a good approximation to φ with high probability.

Let $b \in [0, 2^t - 1]$ an integer be given such that $b/2^t = 0.b_1b_2\cdots_t$ is the best t bit approximation to φ which is less than φ , i.e.,

$$0 \leq \delta \equiv \varphi - b/2^t \leq 2^{-t}. \quad (5.61)$$

(This is getting pretty close to an analysis proof – because it *is* an analysis proof.) Applying the inverse QFT to the state immediately after the first stage gives

$$\begin{aligned} \mathcal{F}^{-1} \left[\frac{1}{2^{t/2}} \sum_{0 \leq k < 2^t} e^{2\pi i \varphi k} |k\rangle \right] &= \frac{1}{2^t} \sum_{0 \leq l < 2^t} e^{\frac{-2\pi i k l}{2^t}} e^{2\pi i \varphi k} |l\rangle \\ &= \frac{1}{2^t} \sum_{0 \leq l < 2^t} \left(e^{2\pi i (\varphi - l/2^t)} \right)^k |l\rangle \end{aligned} \quad (5.62)$$

by definition. Note that when φ can be expressed using exactly t bits we are reduced to the simpler version of the problem described above.

Let α_l be the amplitude of $| (b + l) \bmod 2^t \rangle$, i.e.,

$$\alpha_l \equiv \frac{1}{2^t} \sum_{0 \leq k < 2^t} \left(e^{2\pi i (\varphi - (b+l)/2^t)} \right)^k. \quad (5.63)$$

This is a partial sum of a geometric series. It is straightforward to show that

$$\alpha_l = \frac{1}{2^t} \left(\frac{1 - e^{2\pi i (2^t \varphi - (b+l))}}{1 - e^{2\pi i (\varphi - (b+l)/2^t)}} \right) = \frac{1}{2^t} \left(\frac{1 - e^{2\pi i (2^t \delta - l)}}{1 - e^{2\pi i (\delta - l/2^t)}} \right). \quad (5.64)$$

Suppose the outcome of the final measurement is m . We want to bound the probability that $|m - b| > \epsilon$ (where ϵ is the tolerance), i.e., we don't want the probability that m is far from b to be large. Well, the probability of observing such an m is given by

$$P(|m - b| > \epsilon) = \sum_{-2^{t-1} < l \leq -(\epsilon+1)} |\alpha_l|^2 + \sum_{\epsilon+1 \leq l \leq 2^{t-1}} |\alpha_l|^2. \quad (5.65)$$

For any real θ , we have $|1 - \exp(i\theta)| \leq 2$, so

$$|\alpha_l| \leq \frac{2}{2^t |1 - e^{2\pi i (\delta - l/2^t)}|}. \quad (5.66)$$

Further, provided $-\pi \leq \theta \leq \pi$, we have $|1 - \exp(i\theta)| \geq 2|\theta|/\pi$. Also, when $-2^{t-1} < l \leq 2^{t-1}$ we have $-\pi \leq 2\pi(\delta - l/2^t) \leq \pi$, so we have

$$|\alpha_l| \leq \frac{1}{2^{t+1}(\delta - l/2^t)}. \quad (5.67)$$

With this we have

$$\begin{aligned} p(|m - b| > \epsilon) &\leq \frac{1}{4} \left(\sum_{l=-2^{t-1}+1}^{-(\epsilon+1)} \frac{1}{(l - 2^t \delta)^2} + \sum_{l=\epsilon+1}^{2^{t-1}} \frac{1}{(l - 2^t \delta)^2} \right) \\ &\leq \frac{1}{4} \left(\sum_{l=-2^{t-1}+1}^{-(\epsilon+1)} \frac{1}{(l - 1)^2} + \sum_{l=\epsilon+1}^{2^{t-1}} \frac{1}{(l - 1)^2} \right), \quad 0 \leq 2^t \delta < 1 \\ &\leq \frac{1}{2} \sum_{l=\epsilon}^{2^{t-1}-1} \frac{1}{l^2} \\ &\leq \frac{1}{2} \int_{\epsilon-1}^{2^{t-1}-1} \frac{1}{l^2} dl \\ &= \frac{1}{2(\epsilon - 1)}. \end{aligned} \quad (5.68)$$

Suppose we wish to approximate φ to an accuracy 2^{-n} , i.e., we choose $\epsilon = 2^{t-n} - 1$. By using $t = n + p$ qubits, we see that the probability of obtaining an

approximation correct to this accuracy is at least

$$1 - \frac{1}{2(\epsilon - 1)} = 1 - \frac{1}{2(2^p - 2)}. \quad (5.69)$$

So, to successfully obtain φ accurate to n bits with probability of success at least $1 - \epsilon$ we choose

$$t = n + \left\lceil \log \left(2 + \frac{1}{2\epsilon} \right) \right\rceil. \quad (5.70)$$

Now here's a catch: to use QPE, we need the eigenstate $|u\rangle$. What if we don't know how to prepare such a state? Suppose that we prepare some other state $|\psi\rangle$ in place of $|u\rangle$. We can write

$$|\psi\rangle = \sum_u c_u |u\rangle \quad (5.71)$$

where $|u\rangle$ are the eigenstates of U . Suppose the eigenstate $|u\rangle$ has the eigenvalue $e^{2\pi i \varphi_u}$. Running the QPE on $|\psi\rangle$ will give

$$|\psi\rangle \xrightarrow{\text{QPE}} \sum_u c_u |\tilde{\varphi}_u\rangle |u\rangle, \quad (5.72)$$

where $\tilde{\varphi}_u$ is an estimation for φ_u . Reading out the first register gives us a good approximation for φ_u , where u is chosen at random with probability $|c_u|^2$. This procedure allows us to avoid preparing a possibly unknown eigenstate, but at the cost of introducing some additional randomness into the algorithm. It turns out that if t is chosen according to the minimization procedure above, then the probability for measuring φ_u accurate to n bits at the conclusion of the phase estimation algorithm is at least $|c_u|^2(1 - \epsilon)$.

5.5.2 The algorithm

QPE solves a problem which is both non-trivial and interesting from a physical point of view. It allows us to estimate the eigenvalue associated to a given eigenvector of a unitary operator. It turns out that other interesting problems can be reduced to phase estimation, which makes QPE particularly useful.

Here's the algorithm:

- **Inputs:** (1) An oracle which performs a controlled- \mathcal{U}^j operation, (2) an eigenstate $|u\rangle$ of \mathcal{U} with eigenvalue $e^{2\pi i \varphi_u}$, and (3) $t = n + \lceil \log(2 + 1/2\epsilon) \rceil$ qubits initialized to $|0\rangle$.
- **Outputs:** An n -bit approximation $\tilde{\varphi}_u$ of φ_u .
- **Runtime:** $\mathcal{O}(t^2)$ operations (due to the inverse QFT) and one call to controlled- \mathcal{U}^j oracle. Succeeds with probability at least $1 - \epsilon$.

• **Procedure:**

- Initial state: $|0\rangle|u\rangle$
- Create superposition: $\rightarrow (1/2^{t/2}) \sum_{0 \leq j < 2^t} |j\rangle|u\rangle$.
- Apply the oracle:
 $\rightarrow (1/2^{t/2}) \sum_{0 \leq j < 2^t} |j\rangle \mathcal{U}^j |u\rangle = (1/2^{t/2}) \sum_{0 \leq j < 2^t} e^{2\pi i j \varphi_u} |j\rangle|u\rangle$
- Apply inverse QFT: $\rightarrow |\tilde{\varphi}_u\rangle|u\rangle$.
- Measure the first register: $\rightarrow \tilde{\varphi}_u$.

The algorithm has a wide range of applications, including order-finding and factorization.

5.6 Quantum Variational Eigensolver (QVE)

In this section we're looking this [paper](#) on the quantum variational eigensolver (QVE).

For quantum systems, where the physical dimension grows exponentially, finding the eigenvalues of certain operators is one of many intractable problem and remains a fundamental challenge. QPE efficiently finds the eigenvalue of a given eigenvector but requires fully coherent evolution. This paper presents an alternative approach that greatly reduces the requirements for coherent evolution and combine this method with a new approach to state preparation based on ansatz and classical optimization.

As we have seen, quantum approaches to finding eigenvalues have previously relied on the quantum phase estimation (QPE) algorithm. The QPE algorithm offers an exponential speedup over classical methods and requires a number of quantum operations $\mathcal{O}(p^{-1})$ to obtain an estimate with precision p . Recall that in the standard formulation of QPE, one assumes the eigenvector $|\psi\rangle$ of a Hermitian operator \mathcal{H} is given as input and the problem is to determine the corresponding eigenvalue λ . The time the quantum computer must remain coherent is determined by the necessity of $\mathcal{O}(p^{-1})$ successive applications of $e^{-i\mathcal{H}t}$, each of which can require on the order of millions or billions of quantum gates for practical applications, as compared to the tens to hundreds of gates achievable in the short term.

This paper presents an alternative to QPE that significantly reduces the requirements for coherent evolution. This alternative method is given by a quantum processing unit which efficiently calculates the expectation value of a Hamiltonian, providing an exponential speedup over exact diagonalization, the only known exact solution to the problem on a traditional computer. In conjunction with a classical optimization algorithm, this approach reduces the requirement for coherent evolution of the quantum state, making more efficient use of quantum resources, and may offer an alternative route to practical quantum-enhanced computation.

5.6.1 Quantum Expectation Estimation (QEE)

QEE computes the expectation value of a given Hamiltonian \mathcal{H} for an input state $|\psi\rangle$. This is done via the decomposition:

$$\mathcal{H} = \sum_{i\alpha} h_\alpha^i \sigma_\alpha^i + \sum_{ij\alpha\beta} h_{\alpha\beta}^{ij} \sigma_\alpha^i \sigma_\beta^j + \dots \quad (5.73)$$

for real h , where Roman indices identify the subsystem on which the operator acts, and Greek indices identify the Pauli operator. For example, $\alpha = x$. This is possible because *products of Pauli matrices form a basis for the set of Hermitian*

matrices of dimensions that are powers of 2. In more generality, this is possible because we have the *Pauli group* G_n , which is the group generated by the operators applied to each of n qubits in the tensor product Hilbert space $(\mathbb{C}^2)^{\otimes n}$.

In any case, I digress. By linearity, we have

$$\langle \mathcal{H} \rangle = \sum_{i\alpha} h_\alpha^i \langle \sigma_\alpha^i \rangle + \sum_{ij\alpha\beta} h_{\alpha\beta}^{ij} \langle \sigma_\alpha^i \sigma_\beta^j \rangle \quad (5.74)$$

Just as in QPE, we only consider Hamiltonians that can be written as a polynomial number of terms, wrt the system size. This class of Hamiltonians encompasses a wide range of physical systems, so this requirement is not too strict. In fact, Hamiltonians of this kind include the electronic structure Hamiltonian of quantum chemistry, the quantum Ising Model, the Heisenberg Model, matrices that are well approximate as a sum of n -fold tensor products, and more generally any k -sparse Hamiltonian without evident tensor product structure.

Remark: Sparse Hamiltonians are important because it is possible to decompose \mathcal{H} into $\sum_k \mathcal{H}_k$ where \mathcal{H}_k 's all commute (making diagonalization straightforward). If the matrix is sparse, then we shouldn't need too many distinct \mathcal{H}_k 's. Then we can simulate the Hamiltonian evolution

$$e^{-i\mathcal{H}t} = \prod e^{-i\mathcal{H}_m \delta t} \quad (5.75)$$

where $t = N\delta t$. Now, what do we mean by “sparse?” Sparse means that there are at most d nonzero entries per row, $d = \text{poly}(\log N)$. In any given row, the location of the j th nonzero entry and its value can be computed efficiently (or is given by a black box). Sparse Hamiltonians are different from “Local” Hamiltonians, where $\mathcal{H} = \sum_j \mathcal{H}_j$ where each \mathcal{H}_j acts on $\mathcal{O}(1)$ qubits. \square

In any case, I digress again. The evaluation of $\langle \mathcal{H} \rangle$ reduces to the sum of a polynomial number of expectation values of simple Pauli operators for a quantum state $|\psi\rangle$, multiplied by some real constants. A quantum device can efficiently evaluate the expectation value of a tensor product of arbitrary number of simple Pauli operators. Therefore, with an n -qubit state we can efficiently evaluate the expectation value of this $2^n \times 2^n$ Hamiltonian.

This is possible with a classical computer, but it suffers from the N -representability problem. The power of the quantum approach derives from the fact that quantum hardware can store a global quantum state with exponentially fewer resources than required by classical hardware, and as a result the N -representability problem does not arise.

The expectation value of a tensor product of an arbitrary number of Pauli operators can be estimated by local measurement of each qubit (this can be found somewhere in Mike and Ike). Such independent measurements can be performed in parallel, incurring a constant cost in time. Furthermore, since

these operators are normalized and finite-dimensional, their spectra are bounded. As a result, each

$$\langle H_i^m \rangle = h_{\alpha\beta\dots}^{ij\dots} \langle \sigma_\alpha^i \otimes \sigma_\beta^j \otimes \dots \rangle \quad (5.76)$$

can be estimated to a precision p of an individual element with coefficient h , which is an arbitrary element from the set of constants $\{h_{\alpha\beta\dots}^{ij\dots}\}$ at a cost of $\mathcal{O}(|h_{\max}|^2 M p^{-2})$ repetitions. M is the number of terms in the decomposition of the Hamiltonian and h_{\max} is the coefficient with maximum norm in the decomposition of the Hamiltonian.

The advantage of this approach is that the coherence time to make a single measurement after preparing the state is $\mathcal{O}(1)$. Conversely, the disadvantage of this approach with respect to QPE is the scaling in the total number of operations, as a function of the desired precision is quadratically worse $\mathcal{O}(p^{-2})$ versus $\mathcal{O}(p^{-1})$. Moreover, this scaling will also reflect the number of state preparation repetitions required, whereas in QPE the number of state preparation steps is constant. In essence, we dramatically reduce the coherence time requirement while maintaining an exponential advantage over the classical case, by adding a polynomial number of repetitions with respect to QPE.

5.6.2 Quantum Variational Eigensolver (QVE)

The procedure outlined above replaces the long coherent evolution required by QPE by many short coherent evolutions. In both QPE and QEE we require a good approximation to the ground-state wavefunction to compute the ground-state eigenvalue, and we now consider this problem. The quantum variational eigensolver (QVE) algorithm is a variational method to prepare the eigenstate and, by exploiting QEE, requires short coherent evolution.

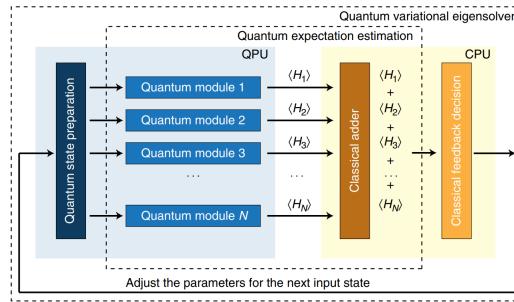


Figure 1 | Architecture of the quantum-variational eigensolver. In QEE, quantum states that have been previously prepared are fed into the quantum modules, which compute $\langle \mathcal{H}_i \rangle$, where \mathcal{H}_i is any given term in the sum defining \mathcal{H} . The results are passed to the CPU, which computes $\langle \mathcal{H} \rangle$. In the quantum variational eigensolver, the classical minimization algorithm, run on the CPU, takes $\langle \mathcal{H} \rangle$ and determines the new state parameters, which are then fed back to the QPU.

It is well known that the eigenvalue problem for an observable represented by an operator H can be restated as a variational problem on the RayleighRitz quotient such that the eigenvector $|\psi\rangle$ corresponding to the lowest eigenvalue is the one that minimizes

$$\frac{\langle \psi | \mathcal{H} | \psi \rangle}{\langle \psi | \psi \rangle} \quad (5.77)$$

By varying the experimental parameters in the preparation of $|\psi\rangle$ and computing the RayleighRitz quotient using QEE as a subroutine in a classical minimization, one may prepare unknown eigenvectors. At the termination of the algorithm, a simple prescription for the reconstruction of the eigenvector is stored in the final set of experimental parameters that define $|\psi\rangle$.

Note that if a quantum state is characterized by an exponentially large number of parameters, it cannot be prepared with a polynomial number of operations. The set of efficiently preparable states are therefore characterized by polynomially many parameters, and we choose a particular set of ansatz states of this type. Under these conditions, a classical search algorithm on the experimental parameters that define $|\psi\rangle$ needs only explore a polynomial number of dimensions – requirement for the search to be efficient.

One example of a quantum state parameterized by a polynomial number of parameters for which there is no known efficient classical implementation is the

unitary coupled cluster ansatz

$$|\Psi\rangle = e^{T-T^\dagger} |\Phi\rangle_{\text{ref}} \quad (5.78)$$

where $|\Phi\rangle_{\text{ref}}$ is some reference state, usually the Hartree Fock ground state, and T is the cluster operator for an N electron system, defined by

$$T = T_1 + T_2 + \cdots + T_N, \quad (5.79)$$

where

$$T_1 = \sum_{pr} t_p^r \hat{a}_p^\dagger \hat{a}_r, \quad T_2 = \sum_{pqrs} t_{pq}^{rs} \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_r \hat{a}_s, \quad \dots \quad (5.80)$$

By construction, $(T - T^\dagger)$ is anti-Hermitian, and exponentiation maps it to a unitary operator $e^{(T - T^\dagger)}$.

For any fixed excitation level k , the reduced cluster operator is written as

$$T^{(k)} = \sum_{i=1}^k T_i. \quad (5.81)$$

This state may be prepared efficiently on a quantum device. The reduced anti-hermitian cluster operator $(T^{(k)} - T^{(k)\dagger})$ is the sum of a polynomial number of termsnamely, it contains a number of terms $\mathcal{O}(N^k(M - N)^k)$, where M is the number of single-particle orbitals. By defining an effective Hermitian Hamiltonian $\mathcal{H} = i(T^{(k)} - T^{(k)\dagger})$ and performing the Jordan Wigner transformation to reach a Hamiltonian that acts on the space of qubits $\hat{\mathcal{H}}$, we are left with a Hamiltonian that is a sum of polynomially many products of Pauli operators. The problem then reduces to the quantum simulation of this effective Hamiltonian, $\hat{\mathcal{H}}$ which can be one in polynomial time using known procedures.

5.7 Some Models in Quantum Statistical & Condensed-Matter Physics

In this section we will be looking at some of the most well-known and most studied models in statistical and quantum many-body mechanics. I'll be sourcing the material from various places, so I'll try my best to cite all the sources.

5.7.1 The Ising Model

In this section we'll look at the Ising model from the mathematical physics perspective. I'll try to add connections to physics to the best of my ability.

Anyway, the Ising model is the simplest and most famous spin system model to study phase transition. It was introduced in 1925 by Ernst Ising in his PhD thesis. He completely solved the model in one dimension, but mistakenly inferred behaviors in higher dimensions.

Basic definitions

Let a graph $G = (V, E)$ be given, with state space $\Omega = \{-1, +1\}^V$ and elements $\sigma = \{\sigma_x\}_{x \in V}$. The variable $\sigma_x \in \{-1, +1\}$ is called the spin at vertex x . This is thus a spin system. Under some re-definitions we can think of this as a lattice gas model, but let's not worry about that.

The energy functional, or the Hamiltonian, of the system is given by

$$\mathcal{H}(\sigma) = -J \sum_{(x,y) \in E} \sigma_x \sigma_y \quad (5.82)$$

where J is a real constant, called the interaction strength. In this formulation, \mathcal{H} is built up from the local energy terms $-J\sigma_x\sigma_y$, i.e., the model is defined in terms of nearest-neighbor interactions. Next, we introduce the inverse temperature parameter $\beta \sim 1/T$ and consider the probability measure on Ω :

$$\mu(\sigma) = \frac{e^{-\beta\mathcal{H}(\sigma)}}{\sum_{\sigma \in \Omega} e^{-\beta\mathcal{H}}} \equiv \frac{e^{-\beta\mathcal{H}(\sigma)}}{Z_\beta} \quad (5.83)$$

where of course Z_β is called the partition function, which contains all information about the system. $\mu = \mu_\beta$ is called the Gibbs measure.

When $J > 0$, we have the ferromagnetic case in which the minimum of $\mathcal{H}(\sigma)$ and so the maximum of $\mu(\sigma)$ are attained at $\sigma \equiv \text{const.}$ More generally, spins like to agree along edges, and for $\beta \rightarrow \infty$ the measure concentrates on the two constant configurations.

When $J < 0$, we have antiferromagnetism where neighboring spins like to differ. This case is hard, and we're not going to worry about it.

For any function $f : \Omega \rightarrow \mathbb{R}$, we have the expectation and variance:

$$\mu(f) = \langle f \rangle = \sum_{\sigma \in \Omega} \mu(\sigma) f(\sigma), \quad \text{Var}(f) = \mu[f^2] - \mu[f]^2. \quad (5.84)$$

So in our case, “ μ ” is like $E[\cdot]$, denoting the expectation value.

Thermodynamic objects

Recall the following definitions

Definition 5.7.1 (Free Energy). The free energy of a system with inverse temperature β and partition function Z_β is given by

$$F(\beta) = -\frac{1}{\beta} \log Z_\beta. \quad (5.85)$$

We note that

$$Z_\beta = e^{-\beta F(\beta)} \quad (5.86)$$

and

$$\partial_\beta F(\beta) = \mu(\mathcal{H}) = \langle \mathcal{H} \rangle. \quad (5.87)$$

Definition 5.7.2 (Specific Heat).

$$C(\beta) = \frac{d}{d(1/\beta)} \mu(\mathcal{H}) = -\beta^2 \frac{d}{d\beta^2} F(\beta). \quad (5.88)$$

We can show that the specific heat is related to the fluctuation of the system via

$$C(\beta) = \beta^2 \text{Var}(\mathcal{H}). \quad (5.89)$$

Next we modify the Hamiltonian by introducing an external magnetic field:

$$\mathcal{H}(\sigma, h) = -J \sum_{(x,y) \in E} \sigma_x \sigma_y - \sum_{x \in V} h_x \sigma_x. \quad (5.90)$$

Here h_x are real numbers and sometimes h_x is identically constant. If $h > 0$, \mathcal{H} has a unique minimum at $\sigma = +1$. A very difficult famous problem is the case when the h_x 's are iid r.v.

Now,

$$F(\beta, h) = -\frac{1}{\beta} \log Z(\beta, h). \quad (5.91)$$

This is a nice smooth function for any finite graph G . With this we consider the magnetization.

Definition 5.7.3 (Magnetization).

$$M(\beta, h) = \frac{1}{|V|} \sum_{x \in V} \mu(\sigma_x) = -\frac{1}{|V|} \partial_h F(\beta, h). \quad (5.92)$$

Both the magnetization and energy will grow as $|V|$ as V grows to infinity, and we would like to take the limit $L \rightarrow \infty$ of quantities like magnetization or any expectation value $\langle f(\sigma) \rangle_L$ where f is any function of finitely many spins. These limits are called the “thermodynamic limit.”

Another quantity of physical interest is the correlation strength. A natural question is how much the spins σ_i and σ_j at two sites are correlated, especially if the spins are far apart. The simplest measurement of this is

$$\langle \sigma_i \sigma_j \rangle - \langle \sigma_i \rangle \langle \sigma_j \rangle. \quad (5.93)$$

This quantity is zero if the spins are independent. What typically happens is that this quantity **decays** exponentially as $|i - j| \rightarrow \infty$ for all values of β but one. So,

$$\langle \sigma_i \sigma_j \rangle \equiv \langle \sigma_i \sigma_j \rangle - \langle \sigma_i \rangle \langle \sigma_j \rangle \sim e^{-|i-j|/\xi} \quad (5.94)$$

as $|i - j| \rightarrow \infty$ where ξ is a length scale which is called the **correlation length**. The correlation length is to some extent a measure of how interesting the system is. A short correlation length means that distant spins are very weakly correlated.

Now if there is an external magnetic field then the Hamiltonian becomes

$$\mathcal{H} = - \sum_{\langle i,j \rangle} \sigma_i \sigma_j - h \sum_i \sigma_i \quad (5.95)$$

where h is the magnetic field external parameter. The free energy is defined by

$$\exp[-\beta F] = \mathcal{Z} = \sum_{\sigma} \exp[-\beta \mathcal{H}] \quad (5.96)$$

where F is a function of the two parameters β and h and the choice of the finite volume (so that the Hamiltonian makes sense). The free energy per site f is then given by F divided by the number of sites. Then the magnetization is given by

$$m = -\partial_h f \quad (5.97)$$

When we defined the magnetization before there was no magnetic field term in \mathcal{H} . To obtain this magnetization we should set $h = 0$ after taking the derivative. However, we can talk about the magnetization of the system with an external

field present, $h \neq 0$. The magnetization is then a function of β, h . The derivative of the magnetization wrt h is called the magnetic susceptibility:

$$\chi = \partial_h m = -\partial_h^2 f. \quad (5.98)$$

By taking the derivative of βf wrt β one obtains the energy per site. The second derivative of f wrt β is known as the specific heat and denoted

$$C = \partial_\beta^2 f. \quad (5.99)$$

Phase transition - the role of the boundary conditions

We will start by asking whether the boundary conditions make a difference in the infinite volume limit. In a finite volume we might guess that + boundary conditions would make the spin at the origin a little more likely to be +1 than -1. What happens when $L \rightarrow \infty$? It turns out that the answer depends on β .

In this section, I will just list some theorems/lemmas without showing proofs. We will just focus on the results for now.

Theorem 5.7.1. If the number of dimensions is at least two, then there is a positive number β_c such that for $\beta < \beta_c$ the limit $\lim_{L \rightarrow \infty} \langle \sigma_0 \rangle_L^+$ exists and is zero while for $\beta > \beta_c$ the limit exists and is strictly greater than zero.

The notation $\langle \cdot \rangle^+$ denotes the positive boundary condition. This just means that all of the outside spins are +1 initially. The intuition is that when β is small, the spins should only be weakly correlated. The influence of the boundary condition will decay rapidly as we move in from the boundary.

Theorem 5.7.2. In any number of dimensions there is a positive number β_1 such that for $\beta < \beta_1$ the limits $\lim_{L \rightarrow \infty} \langle \sigma_0 \rangle_L^+$ and $\lim_{L \rightarrow \infty} \langle \sigma_0 \rangle_L^-$ are both zero.

Theorem 5.7.3. In two or more dimensions there is a positive number β_2 such that for $\beta > \beta_2$ the limit $\liminf_{L \rightarrow \infty} \langle \sigma_0 \rangle_L^+$ is strictly greater than zero for + boundary conditions.

Phase transitions and critical phenomena

Crudely speaking a phase transition is when an abrupt change occurs in the infinite volume system at some values of the parameters like temperature and magnetic field. The values of the parameters where this happens are called a *critical point*.

What typically happens is that when $T \rightarrow T_c$ some physical quantities diverge while others converge to zero. The way in which they do so, though, is interesting. Correlation length goes to zero as $T \rightarrow \infty$ and $T \rightarrow 0$. At T_c , the correlation length goes to ∞ . We can think of phase transition as the correlation length diverging.

As $T \rightarrow T_c$ the specific heat and magnetic susceptibility also diverge to ∞ . As $T \rightarrow T_c^-$ the magnetization m will converge to zero.

5.7.2 The Hubbard Model

In this section we will look at this [document](#) on the Hubbard Hamiltonian. In some sense the HH is quite similar to the Ising model, except that it is written in terms of fermionic creation and annihilation operators.

I'm not sure we have time to cover this partially (or even at all!), but here's the layout:

- Introduction
- Creation and Annihilation operators
- The Hubbard Hamiltonian
- Particle-hole symmetry
- The single-site limit
- The non-interacting Hubbard Hamiltonian
- Exact diagonalization: two-site HH
- Green functions: Mott gap and spectral function
- A peek at magnetism
- ...
- Conclusions

5.8 Locality in Quantum Systems

In this section we will look at this [paper](#) by Matthew Hastings. This is actually a compilation of lecture notes focusing on the application of ideas of locality, in particular Lieb-Robinson bounds, to quantum many-body systems. For our purposes, we will only look at a few topics included.

- We want to understand “locality” in the context of quantum simulation.
What, precisely, does it mean for a Hamiltonian to be “local”?
- What are the Lieb-Robinson bounds?
- What is the Lieb-Schulz-Mattis theorem?
- What are topologically ordered states?

5.8.1 Introduction & Notation

The basic problem studied in quantum many-body theory is to find the properties of the ground state of a given Hamiltonian. Expressed mathematically, the Hamiltonian \mathcal{H} is a Hermitian matrix, and the ground state, which we write ψ_0 , is an eigenvector of this matrix with the lowest eigenvalue. The properties we are interested in studying are expectation values of various observables: given a Hermitian matrix O , we would like to compute the expectation value $\langle \psi_0, O\psi_0 \rangle \equiv \langle O \rangle$.

While the method of “exact diagonalization” on a computer is an important technique in studying quantum systems, it is only a small part of how physical problems are studied. The feature that distinguishes the study of quantum many-body systems is *locality* of interactions. Consider a typical Hamiltonian, such as the one-dimensional transverse field Ising model (TFIM):

$$\mathcal{H} = -J \sum_{i=1}^{N-1} S_i^z S_{i+1}^z + B \sum_{i=1}^N S_i^x. \quad (5.100)$$

The very notation implicitly assumes local interactions. To be precise, throughout this section, we only consider quantum systems on a finite size lattice. We associate a D dimensional Hilbert space with each lattice site, and the Hilbert space of the whole system is tensor product of these spaces. We use N to represent the number of lattice sites, so that the Hilbert space on which a Hamiltonian such as the one above is defined in D^N dimensions. A term such as $S_i^z S_{i+1}^z$ is short hand for $I_1 \otimes I_2 \otimes \cdots \otimes S_i^z \otimes S_{i+1}^z \otimes I_{i+2} \otimes \cdots \otimes I_N$, where I_j is the identity operator on site j .

We let i, j, k, \dots denote the lattice site, and X, Y, Z to denote sets of lattice sites. We set Λ to denote the set of all lattice sites. We use $|X|$ to denote the cardinality of a set X .

An operator O is supported on a set A if we can write O as a tensor product of two operators $O = I_{\Lambda \setminus A} \otimes P$, where $I_{\Lambda \setminus A}$ is the identity operator on the sites not in A and P is some operator defined on the $D^{|A|}$ dimensional Hilbert space on the set A . For example, $S_i^z \otimes S_{i+1}^z$ is supported on the set $\{i, i+1\}$.

We let $\|O\|$ denote the “operator norm” for O . If O is Hermitian, the operator norm is the equal to the absolute value of the largest eigenvalue of O . For arbitrary operators O , the operator norm is given by

$$\|O\| = \max_{\psi, |\psi|=1} |O\psi|, \quad (5.101)$$

which one might have seen in real/functional analysis.

Using 2 simple assumptions: that the Hamiltonian has local interactions and that the Hamiltonian has a spectral gap, we will be able to prove a wide variety of results about the ground state of the Hamiltonian. We will consider two cases of a spectral gap: one where $E_1 - E_0 \geq \Delta E$, where ΔE is called the “spectral gap.” In the other case, we consider a Hamiltonian with several degenerate or approximately degenerate low energy states.

We will begin by considering properties of correlations in these systems, and prove an exponential decay of correlation functions. We then consider how the ground state of the system changes under a change in the Hamiltonian, and use this to prove a non-relativistic variant of Goldstone’s Theorem. Finally, we apply these techniques to more interesting systems with topological order, beginning with the Lieb-Schultz-Mattis theorem. However, before we can do any of this, we need to more precisely define the locality properties of the Hamiltonian and to prove a set of bounds on the propagation of information through the system called the Lieb-Robinson bounds.

5.8.2 Locality & The Lieb-Robinson bounds

To define locality, we need a notion of distance, so we introduce a metric on the lattice $d(i, j)$. One may consider the Manhattan or Euclidean metric. It’s natural to consider the metric

$$d(i, j) = \min_n |i - j + nN| \quad (5.102)$$

for a one-dimensional system with periodic boundary conditions and N sites. Next, we define the distance between two sets A, B to be

$$d(A, B) = \min_{a \in A, b \in B} d(a, b). \quad (5.103)$$

The diameter of a set A is given by

$$\text{diam}(A) = \max_{i, j \in A} d(i, j). \quad (5.104)$$

Next we consider Hamiltonians

$$\mathcal{H} = \sum_Z H_Z \quad (5.105)$$

where each H_Z is support on Z . We are interested in studying problems in which $|H_Z|$ decays rapidly with $\text{diam}(Z)$.

(to be continued)

5.9 Quantum Approximate Optimization Algorithm (QAOA)

Let's not worry about this now.

5.10 Quantum Variational Factorization (QVF)

Let's not worry about this now.

5.11 Variational Thermal Quantum Simulation via Thermofield Double States

Let's not worry about this now.

5.11.1 Thermofield Double States (TFD)

5.12 Generation of TFD and Critical Ground States with a Quantum Computer

5.13 Efficient variational simulation of non-trivial quantum states

In this section we look at this [paper](#), with title "Efficient variational simulation of non-trivial quantum states." The idea is the following. The paper provides an efficient and general route for preparing *non-trivial quantum states* (defined later) that are NOT adiabatically connected to unentangled product states. This approach will be a hybrid of quantum and classical variational protocols that incorporates a feedback loop between a quantum simulator and a classical computer. This turns out to be realizable on near-term quantum devices of synthetic quantum systems.

The paper provides explicit protocols which prepare with *perfect fidelities* the following non-trivial quantum states: GHZ, quantum critical state, and topologically ordered state, with $L = 2p$ variational parameters and physical runtimes $T \sim L$.

The paper also conjectures and numerically support (so, not analytically support) that the protocol can prepare with perfect fidelity and similar operational costs the ground state of every point in the 1D transverse field Ising model phase diagram (we will define this later).

5.13.1 Introduction

The paper demonstrates the efficient preparation of certain non-trivial states using a variational hybrid quantum-classical simulation, which utilizes resources of a quantum simulator and a classical computer in a feedback loop. Given Hamiltonians realizable on a quantum simulator, a quantum state $|\psi\rangle(\gamma, \beta)$ is produced with $(\gamma, \beta) \equiv (\gamma_1, \dots, \gamma_p, \beta_1, \dots, \beta_p)$ parameterizing a finite set of $2p$ variational angles (or times) that the Hamiltonians are run for.

A cost function, usually taken to be the energy of some target Hamiltonian, is then evaluated within the resulting state and optimized for in a classical computer, which yields a new set of $2p$ angles to be implemented to be fed back into the quantum simulator.

The entire process is iterated to convergence.

An example of an approach like this is the Quantum Approximate Optimization Algorithm (QAOA). There are a number of properties which make the variational quantum simulation appealing: it can be run on any quantum device (digital/tunable analog). The feedback loop of the protocol allows one to mitigate systematic errors that might be present in the experiment setups.

This variational quantum-classical simulations on local, uniform Hamilto-

nians can target GHZ state, the critical state of the 1D transverse field Ising model (TFIM), and the ground state of the 2d Toric code, all with *perfect fidelity*.

Later, we will see that the entire ground state phase diagram of the TFIM can be produced with perfect fidelities and similar operational costs. This is only conjectured and numerically supported.

Finally, we will look at the preparation of the ground states of the antiferromagnetic (AFM) Heisenberg chains. We will see that the protocol is able to achieve them efficiently and with *very good* fidelities (so, not perfect, but very good).

5.13.2 Non-trivial quantum states

In any case, let's look at some non-trivial quantum states. What are non-trivial quantum states? Consider a target state $|\psi_t\rangle$ and an untangled product state $|\psi_u\rangle$, both defined on a system with linear dimension L . $|\psi_t\rangle$ is said to be *non-trivial* if there does not exist a *local* unitary circuit \mathcal{U} of finite depth that connects the two: $|\psi_t\rangle = \mathcal{U}|\psi_u\rangle$. Instead, the depth of a local unitary circuit connecting the two must be at least $\mathcal{O}(L^\alpha)$.

Intuitively, non-trivial states have entanglement patterns fundamentally different from product states, i.e., these are states that are very “far” from being product states.

From the perspective of local Hamiltonians and gaps, such states are separated from product states by a gap-closing phase transition in the thermodynamic limit, and thus preparing them with the quantum adiabatic algorithm is hard.

Example 5.13.1 (GHZ states are non-trivial.). Let's first look at what GHZ states are. GHZ stands for Greenberger-Horne-Zeilinger. GHZ is a certain type of entangled quantum state that involved at least three subsystems. Extremely non-classical properties of the state have been observed.

Definition 5.13.1. The GHZ state is an entangled quantum state of $M > 2$ subsystems. If each system has dimension d , i.e., the local Hilbert space is isomorphic to \mathbb{C}^d , then the total Hilbert space of M partite system is $\mathcal{H}_{\text{tot}} = (\mathbb{C}^d)^{\otimes M}$. This GHZ state is also named as M -partite qubit GHZ state, which reads

$$|\text{GHZ}\rangle = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} \bigotimes_{i=0}^{d-1} |i\rangle. \quad (5.106)$$

In the case of each of the subsystems being two-dimensional, that is for qubits, it reads

$$|\text{GHZ}\rangle = \frac{|0\rangle^{\otimes M} + |1\rangle^{\otimes M}}{\sqrt{2}}. \quad (5.107)$$

The GHZ state is a quantum superposition of all qubits being in 0 with all being in 1. The GHZ state is a maximally entangled quantum state. The simplest being $|\text{GHZ}\rangle = (|000\rangle + |111\rangle)/2$.

GHZ states are in contrast to the so-called “cat states,” which are quantum superpositions of two macroscopically distinct states. Note that a cat state, unlike GHZ states, are not necessarily an entangled state because a cat state could be of one or more modes or particles.

Let us look at why the GHZ states are non-trivial. Consider the GHZ state:

$$|\text{GHZ}\rangle = \frac{1}{\sqrt{2}} \left(\bigotimes |Z=1\rangle + \bigotimes |Z=-1\rangle \right). \quad (5.108)$$

Suppose there exists a local, finite-depth unitary \mathcal{U} that takes the completely polarized product state $|+\rangle \equiv \bigotimes |X=1\rangle$ to $|\text{GHZ}\rangle$. Then due to locality, there exists a Lieb-Robinson bound which limits the spread of information and entanglement under this evolution, implying that \mathcal{U} can only generate a finite correlation length ξ for the final state. Measuring a long-range spin-spin correlator gives

$$\langle \text{GHZ} | Z_i Z_j | \text{GHZ} \rangle = 1, \quad (5.109)$$

while on the other hand the same quantity can be expressed as

$$1 = \langle + | \mathcal{U}^\dagger Z_i \mathcal{U} \mathcal{U}^\dagger Z_j \mathcal{U} | + \rangle, \quad (5.110)$$

which in the limit $|i - j| \gg \xi$ decomposes as

$$\langle + | \mathcal{U}^\dagger Z_i \mathcal{U} | + \rangle \langle + | \mathcal{U}^\dagger Z_j \mathcal{U} | + \rangle = \langle \text{GHZ} | Z_i | \text{GHZ} \rangle \langle \text{GHZ} | Z_j | \text{GHZ} \rangle = 0 \neq 1 \quad (5.111)$$

which is a contradiction.

Similar arguments apply to critical states which have polar-law correlations, and topologically ordered states which have long-range correlations in loop operators and non-zero topological entanglement entropy.

5.13.3 Variational Quantum-Classical Simulation (VQCS)

Now, we define the variational quantum-classical simulation (VQCS). This is motivated by the QAOA and the VQE.

The protocol is envisioned to be run on a quantum simulator that can realize certain interactions between qubits and single qubit rotations, which we denote schematically by H_1, H_2 . The aim is to produce a good approximation to the ground state of a target many-body Hamiltonian H_T , which we will assume to be a linear combination of H_1 and H_2 .

The VQCS begins with the ground state of H_1 , denoted $|\psi_1\rangle$, and evolves it with the following sequence alternating between H_1, H_2 :

$$|\psi(\boldsymbol{\gamma}, \boldsymbol{\beta})\rangle_p = e^{-i\beta_p H_1} e^{-i\gamma_p H_2} \dots e^{-i\beta_1 H_1} e^{-i\gamma_1 H_2} |\psi_1\rangle. \quad (5.112)$$

This is motivated by the quantum adiabatic algorithm, which works because of the adiabatic theorem. Basically the idea is that we are starting out in a ground state of some Hamiltonian H_1 . The goal is to evolve the Hamiltonian H_1 slowly so that it turns into the target Hamiltonian H_2 . The ground state of H_1 as a result will turn into the ground state of H_2 as well. Now, the way we slowly vary from one to another is by letting the ground state of H_1 evolve slowly under sequential linear combinations of H_1 and H_2 . This means the initial ground state is going to be evolved by

$$e^{i(H_1 + \epsilon H_2)}. \quad (5.113)$$

Now, we don't want to do this. Rather, we want to use Trotterization to factor out and turn the above operator into

$$\sim e^{iH_1} e^{iH_2} \quad (5.114)$$

Of course, we want to do this sequentially, so the result is that we have alternating evolution operator as above.

Of course, there are many ways to get to the $|\psi(\boldsymbol{\gamma}, \boldsymbol{\beta})\rangle$ state. But this just happens to be a pretty common way to do that.

For a fixed integer p , there are $2p$ variational angles

$$(\boldsymbol{\gamma}, \boldsymbol{\beta}) \equiv (\gamma_1, \dots, \gamma_p, \beta_1, \dots, \beta_p). \quad (5.115)$$

A cost function, such as the energy expectation value of the target Hamiltonian

$$F_p(\boldsymbol{\gamma}, \boldsymbol{\beta}) =_p \langle \psi(\boldsymbol{\gamma}, \boldsymbol{\beta}) | H_T | \psi(\boldsymbol{\gamma}, \boldsymbol{\beta}) \rangle_p, \quad (5.116)$$

is then evaluated. A classical computer then performs an optimization to produce a new set of angles (or times) $(\boldsymbol{\gamma}, \boldsymbol{\beta})$ which are then fed back into the quantum simulator and the process is repeated until the cost function is desirably minimized. The state corresponding to these optimal angles is therefore the optimal state that can be prepared by the protocol given this cost function.

As the VQCS is envisioned to be run on near-term quantum simulators which are inherently noisy (NISQ), the physical runtimes $t = \sum_i^{p=L/2} (\gamma_i + \beta_i)$ of the VQCS constitute an important measure of the feasible of the protocol – in general, shorter runtimes lead to less noise encountered and a better implementation.

5.13.4 Using VQCS to prepare nontrivial quantum states

In this section we look at how the VQCS generates some non-trivial quantum states.

GHZ state

Recall the GHZ state looks like

$$|\text{GHZ}\rangle = \frac{1}{\sqrt{2}} \left(\bigotimes |Z=1\rangle + \bigotimes |Z=-1\rangle \right). \quad (5.117)$$

So it is a superposition of two possible ways that the system is completely aligned. The GHZ state can be taken to be the ground state of the following Hamiltonian

$$\mathcal{H}_T = \sum_{i=1}^L Z_i Z_{i+1}, \quad (5.118)$$

in the symmetry sector $S = \prod_i^L X_i = 1$. Just to clarify the notation, Z_i gives $+1$ when the state is “up”, i.e., if the state is $|\uparrow\rangle$. This means the Hamiltonian “favors” (i.e., achieve lowest energy) when everybody is aligned. Further, for simplicity we assume periodic boundary conditions.

We choose in this case $H_1 = -\sum_i X_i$ with the product ground state $|\psi_1\rangle = |+\rangle = \bigotimes_i^L |+\rangle_i$. Recall that on the Bloch sphere, the $|+\rangle$ and $|-\rangle$ states lie in the horizontal plane, and are eigenstates of the X operator, i.e., $X |\pm\rangle = \pm 1 |\pm\rangle$. We also know that H_1 has a straightforward implementation in both the analog and digital settings.

Next, choose $H_2 = \mathcal{H}_T$. On a digital quantum simulator, this can be achieved using elementary two-site gates $Z_i Z_{i+1}$:

$$e^{-i\gamma H_2} = \prod_{i=1}^{L/2} e^{-i\gamma Z_{2i} Z_{2i+1}} \prod_{i=1}^{L/2} e^{-i\gamma Z_{2i-1} Z_{2i}} \quad (5.119)$$

such that each unitary $e^{-i\gamma H_2}$ in the VQCS protocol can be considered as a depth-2 quantum circuit, so that overall the VQCS_p can be realized as a quantum circuit with depth at most $3p$. If we start in $|+\rangle$ then since the VQCS respects this symmetry, optimization of $F_p(\gamma, \beta)$ as $p \rightarrow \infty$ yields the GHZ

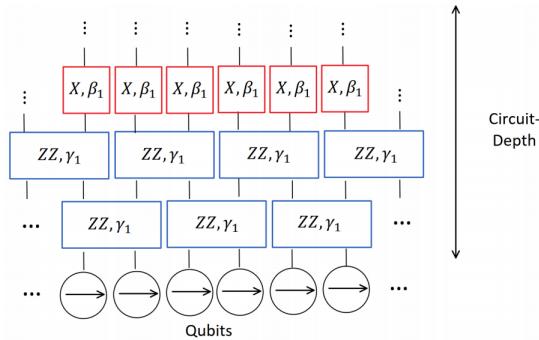


Figure 2: Quantum circuit realization of the VQCS protocol on a digital quantum simulator. Shown is the first layer of the protocol (5) using angles (γ_1, β_1) corresponding to time evolution by two-site and one-site gates $Z_i Z_{i+1}$ and X_i respectively. Subsequent layers utilize different angles (γ_n, β_n) . For a total of p layers corresponding to VQCS_p , the quantum circuit has a depth of at most $3p$.

state, with $\lim_{p \rightarrow \infty} F_p(\boldsymbol{\gamma}, \boldsymbol{\beta})/L = -1$.

Let's see how this performs numerically. To do this, we implement the VQCS, finding numerically the optimal angles $(\boldsymbol{\gamma}_*, \boldsymbol{\beta}_*)$ that minimizes $F(\boldsymbol{\gamma}, \boldsymbol{\beta})$ (which, remember, is the energy of the system) via a search by gradient descent of the parameter space $\gamma_i, \beta_i \in [0, \pi/2]$, for system sizes $L \leq L_1 = 18$ and for $p \leq L_1/2$. Note that for fixed L , a brute force search of this parameter space takes an exponentially long time $t \sim \mathcal{O}(M^{2p})$ in p . So, we have to ensure that the total number of runs performed is large enough to ensure convergence of the search algorithm to the global minimum. Here's the result given in the paper:

The figure is quite self-explanatory so I won't go into much detail there. We will just note that the preparation achieves perfect fidelity to machine precision.

Let's also note that various quantum circuits are known to also exactly prepare the GHZ state (using Hadamard and CNOT, and so on). Furthermore, experimentally, GHZ states of various sizes have been prepared with high fidelity using the Molmer-Sorensen technique in trapped ions. The VQCS is complementary in that it provides a uniform circuit that achieves the same result.

Critical State

Consider the preparation of the critical state, namely the ground state of the critical 1d transverse field Ising model (TFIM) on a ring:

$$\mathcal{H}_T = - \sum_{i=1}^L Z_i Z_{i+1} - \sum_{i=1}^L X_i. \quad (5.120)$$

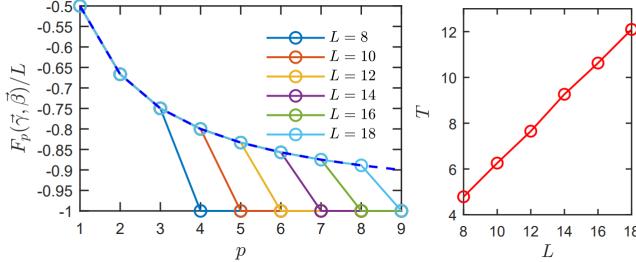


Figure 3: Preparation of GHZ state. (Left) Optimal cost function (6). One sees that $F_p(\gamma, \beta)/L = -1$ for $p \geq L/2$; in other words, the GHZ state is created with perfect fidelity using $\text{VQCS}_{p \geq L/2}$. We have also plotted a conjectured analytic expression $-p/(p+1)$ (from [26]) for the optimal cost function as the dashed blue line. (Right) Total minimum time $T = \min_{(\gamma, \beta)} \left[\sum_i^{p=L/2} (\gamma_i + \beta_i) \right]$ required for the VQCS to produce the GHZ state with perfect fidelity using $\text{VQCS}_{p=L/2}$. The minimization is performed over all the numerical solutions found. One sees a linear trend $T \sim L$.

As before, we assume the same operators $H_1 = -\sum_{i=1}^L X_i$ and $H_2 = -\sum_{i=1}^L Z_i Z_{i+1}$ as before, though now we want to minimize the new cost function $F(\gamma, \beta)$ using this H_T . In this case, we benchmark the simulation by computing the overlap $|\langle \psi_p | \psi_t \rangle|^2$ where the ground state of the system $|\psi_t\rangle$ is obtained (theoretically) by exact diagonalization.

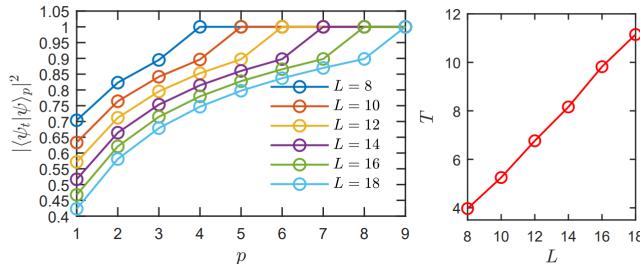


Figure 4: Preparation of critical state. (Left) Many-body overlap $|\langle \psi_t | \psi \rangle_p|^2$ of the prepared state with the target ground state of (9) found by exact diagonalization. One sees perfect fidelity for $p \geq L/2$. (Right) Total minimum time $T = \min_{(\gamma, \beta)} \left[\sum_i^{p=L/2} (\gamma_i + \beta_i) \right]$ required for the VQCS to produce the critical state with perfect fidelity using $\text{VQCS}_{p=L/2}$. One sees a linear trend $T \sim L$.

Once again we observe that perfect fidelity can be obtained.

Ground states of the TFIM at generic points in the phase diagram

The perfect fidelities achieved for both the GHZ and critical cases using the VQCS for $p = L/2$ suggest that other points g in the phase diagram of the TFIM

$$H_{\text{TFIM}} = - \sum_i^L Z_i Z_{i+1} - g \sum_i^L X_i \quad (5.121)$$

might be similarly targeted. The paper conjectures that for a 1d system of even L spin-1/2s with period boundary conditions, any state produced by VQCS for arbitrary p using the H_2, H_1 above can also be achieved perfectly by the same procedure. This would imply that it is possible to achieve the ground state of H_{TFIM} at any point g in the phase diagram, which in particular would cover the GHZ and critical cases.

Let's note that while perfect fidelities achieved (to numerical precision) suggest that an analytic understanding may be possible. However, while the model and unitary gates can be mapped to free fermions, the minimization of the VQCS cost function maps to a nonlinear optimization problem involving an extensive number of variables, which is highly nontrivial. s

Ground state of the Toric code

This is complicated, so we'll come back to this later. The upshot is that the VQCS procedure also works well here (perfect fidelity and so on) with runtime increasing linearly with L , which something we also saw before.

One thing to note here is just that Toric code are those with Hamiltonians with four interacting neighboring qubits. For example, the Hamiltonian looks like

$$\mathcal{H} = - \sum_i^L \sum_j^L \sigma_{i,j+1}^x \sigma_{i+1,j+1}^y \sigma_{i=1,j}^x \sigma_{i,j}^y. \quad (5.122)$$

The Toric code is an example of topologically-ordered quantum state. Toric codes can describe really exotic behaviors that are too advanced for our guide here. For more detail, consult Wikipedia.

Ground state of AFM Heisenberg chain

Here we consider targeting the ground state of the AFM spin-1/2 Heisenberg chains with open boundary conditions,

$$\mathcal{H}_T = \sum_{i=1}^{L-1} \mathbf{S}_i \cdot \mathbf{S}_{i+1}. \quad (5.123)$$

Here we use

$$H_1 = \sum_{i=1}^{L/2-1} \mathbf{S}_{2i} \cdot \mathbf{S}_{2i+1}, \quad H_2 = \sum_{i=1}^{L/2} \mathbf{S}_{2i-1} \cdot \mathbf{S}_{2i}, \quad (5.124)$$

whilst evaluating the energy as the cost function. Note that the initial state is now the product state of Bell pairs $\bigotimes_i (1/\sqrt{2}) (|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)_{2i-1,2i}$, and that angles can be restricted to $\gamma, \beta \in [0, 2\pi)$.

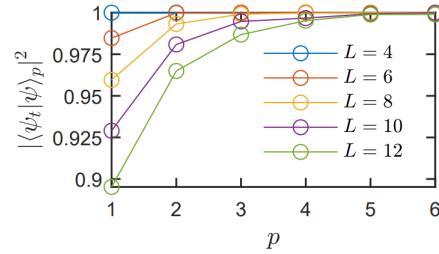


Figure 5: Fidelities in the preparation of ground states of AFM spin-1/2 Heisenberg chains of various sizes using VQCS.

In this case, while the preparation of the ground states is generically not perfect for any finite p , the many-body fidelities are already very good for very low p 's, at least for small system sizes (small L). This illustrates the utility and generality of VQCS in preparing non-trivial quantum states.

5.14 Cluster States & Measurement-based QC

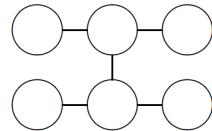
In this section we look at concepts such as *cluster states* and *measurement-based quantum computation* as introduced in this [paper](#) by M. Nielsen. I will assume familiarity with the basics of the quantum circuit model and go straight to the cluster state model.

We will explain the cluster state model of quantum computation, and explain how cluster states can be used to efficiently simulate quantum circuits.

5.14.1 How it works

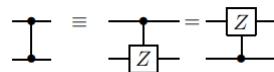
A cluster-state computation begins with the preparation of a special entangled many-qubit quantum state, known as the *cluster state*, followed by an adaptive sequence of single-qubit measurements, which process the cluster, and finally read-out of the computation's result from the remaining qubits.

The term *cluster state* refers not to a single quantum state, but rather a class of quantum states. To any graph G on n vertices we can define an associated n -qubit cluster state, by first associating to each vertex a corresponding qubit, and then applying a graph-dependent preparation procedure to the quits. For example, this graph represents a six-qubit cluster state: The cluster state



associated to the graph may be defined as the result of applying the following preparation procedure:

- Prepare each of the n qubits in the state $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$.
- Applying the controlled-PHASE gates between qubits whose corresponding graph vertices are connected. Recall that the controlled-PHASE looks like and remember that in the first case the Z gate is applied to the second



qubit if the first qubit is 1. In the second case, the Z gate is applied to the first qubit if the second qubit is 1. Note that if the “dot” in the circuit is a circle then Z is applied when the controlled qubit is 0.

Because the controlled-PHASE (or controlled- Z) gates commute with each other, we don't worry about the order in which these gates are applied. Also, although we have described the preparation of the cluster in terms of applying

quantum gates, later in the paper we briefly describe how to prepare clusters using measurement gates, and so the cluster-state model may be regarded as a truly measurement-only model of quantum computation.

In the original paper by Briegel, the term *graph state* is used instead of *cluster states*. The term *cluster state* was introduced to refer to the case where the graph G is a two-dimensional square lattice. This is the class of states which could be used as a substrate for quantum computations. The term *graph state* refers to states associated with more general graphs G . In any case, we will stay with the term *cluster state* for now, for consistency.

Once the cluster state is prepared, (note that the cluster state has two components: the vertices of qubits and edges of operators) the next step in the computation is to perform a sequence of *processing measurements* on the state. These measurement satisfy:

- They are single-qubit measurements
- The choice of measurement basis may depend on the outcomes or earlier measurements, i.e., feedforward of classical measurement results is allowed
- Measurement results may be processed by a classical computer to assist in the feedforward, so that choice of the basis may be a complicated function of earlier measurement results.

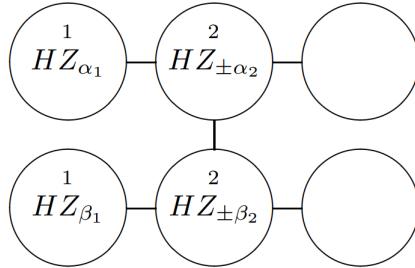
Note that for the cluster state computation to be efficient we must constrain the classical computation to be of polynomial size.

There are two ways to define the output of the cluster-state computation:

- We can define the output in terms of the *quantum state* resulting from the computation. This is the quantum state of the qubits which remain when the sequence of processing measurements has terminated.
- We can also define in terms of the set of *read-out measurements*, a sequence of single-qubit measurements applied to the qubits which remain when the processing measurements are complete. In this case the output of the computation is a classical bit string.

Example 5.14.1. We will consider an example of a cluster-state computation to see how it might work explicitly. In this example, labels indicate qubits on which processing measurements occur, while unlabeled qubits are those which remain as the output of the computation when the processing measurements are complete. The labeled qubits are labeled with (1) a positive integer n and (2) a single-qubit unitary, which we refer to as a generic unitary U . Here, $U = HZ_{\pm\alpha_j}, HZ_{\pm\beta_j}$.

The positive number n indicates the time-ordering of the processing measurements. Qubits of the same n can be measured in either order or simultaneously.

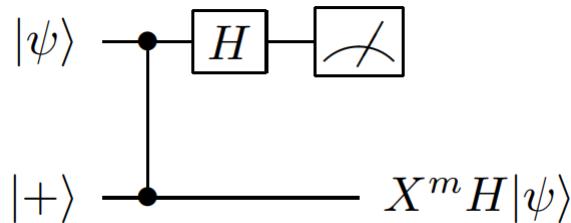


The time-ordering is important, because it determines which measurement results can be fed forward to control later measurement bases.

The unitary label U indicates the basis in which the qubits are measured, denoting a rotation by the unitary U , following by a computational basis measurement. In other words, a single-qubit measurement in the basis $\{U^\dagger |0\rangle, U^\dagger |1\rangle\}$ is performed. The notation \pm denotes the fact that the choice of sign depends on the outcomes of earlier measurements, in some specified manner. We will look at an explicit example later.

5.14.2 Simulating quantum circuits in the cluster-state model

In this part we look at how cluster-state computation can simulate quantum circuits. The key idea underlying this example is the one-bit teleportation:



Here m is the outcome of the computational basis measurement on the first qubit. Let's verify this. Suppose we start out with $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ and $|+\rangle = (1/\sqrt{2})(|0\rangle + |1\rangle)$. Then the composite quantum state is given by

$$|\psi+\rangle = \alpha|00\rangle + \beta|11\rangle + \alpha|01\rangle + \beta|10\rangle. \quad (5.125)$$

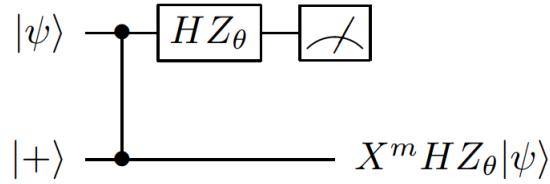
A controlled- Z , followed by a Hadamard on the first qubit gives

$$\begin{aligned} (\mathcal{H} \otimes \mathcal{I}) C_Z |\psi+\rangle &= \alpha|+0\rangle - \beta|-1\rangle + \alpha|+1\rangle + \beta|-0\rangle \\ &= \alpha|++\rangle + \beta|--\rangle. \end{aligned} \quad (5.126)$$

This state can be re-written as

$$\begin{aligned} (\mathcal{H} \otimes \mathcal{I}) C_Z |\psi+\rangle &= \frac{1}{\sqrt{2}}(|0\rangle \otimes \mathcal{H}|\psi\rangle + |1\rangle \otimes X\mathcal{H}|\psi\rangle) \\ &= \frac{1}{\sqrt{2}}(|0\rangle \otimes X^0\mathcal{H}|\psi\rangle + |1\rangle \otimes X^1\mathcal{H}|\psi\rangle). \end{aligned} \quad (5.127)$$

So the result follows. However, we can generalize this a bit further to get the following teleportation circuit:



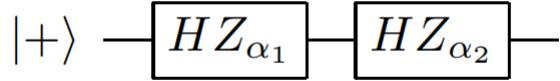
The proof is similar. We note that Z_θ commutes with the controlled- Z , so the output of the circuit is the same as would have been output from the previous circuit had $Z_\theta|\psi\rangle$ been input instead of $|\psi\rangle$.

What do learn from this? We see that even though the first qubit is measured, no quantum information is lost, for no matter what measurement outcome, the posterior state of the second qubit is related by a known unitary transformation to the original input $|\psi\rangle$: $|\psi\rangle$ vs. $X^m HZ_\theta|\psi\rangle$.

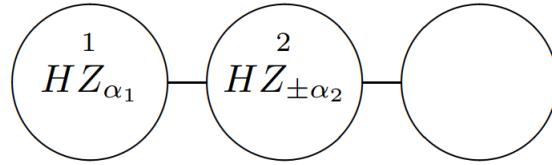
This might seem unsurprising. However, notice that by varying the basis in which the first qubit is measured, i.e., by varying θ , we can vary the unitary transformation acting on the second qubit without destroying any quantum information.

Single-qubit case

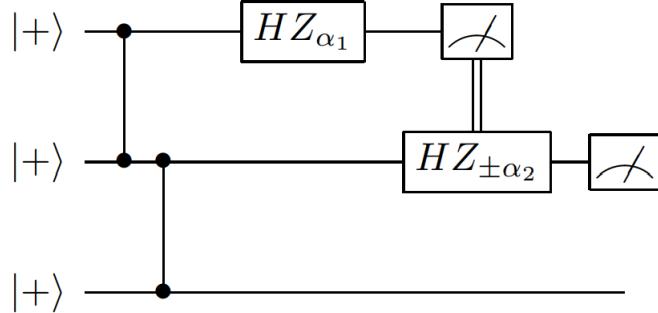
Now, we will use this example to explain how the cluster-state computation can simulate quantum circuits. Let us explain how to simulate a single-qubit circuit of the form



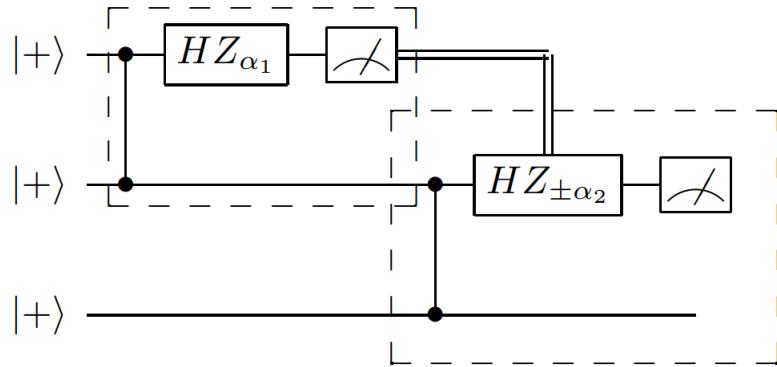
This looks trivial, but is very important. Here we are assuming that the qubit starts at $|+\rangle$ and single-qubit gates are of the form $\mathcal{H}Z_\alpha$ for convenience. The associated cluster-state computation used to simulate the given circuit is



Note that in order to simulate this one-qubit quantum circuit, we require 03 qubits in the cluster-state model. This will be a recurring theme in the cluster-state computation: the complexity in the quantum circuit is prepackaged into the initial quantum state, so that the complexity of the measurements is reduced. By construction, this cluster-state model is equivalent to the following quantum circuit:



where we have assumed all qubits are starting out in the $|+\rangle$ state. Further, we can also equivalently delay the controlled- Z between the second and third qubits to after $\mathcal{H}Z_{\alpha_1}$ to split the circuit into two parts like this:



It can be readily checked that these circuits are equivalent. This can be done

by thinking of the “arrow of time” in the cluster-state model as pointing out of the page.

By splitting the circuit like this, we can see that the two highlighted boxes are of the form of the one-bit teleportation circuit that we just discussed. This means that two teleportations occur in this circuit. The first teleportation send $|\psi\rangle = |+\rangle$ to the second qubit, which has the form:

$$\text{After 1st teleportation: } |\psi_1\rangle = X^{m_1} \mathcal{H} Z_{\alpha_1} |+\rangle \quad (5.128)$$

as shown earlier.

The second teleportation sends this $|\psi_1\rangle$ to the third qubit, which has the form

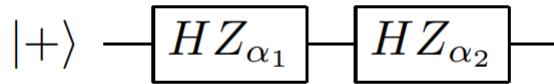
$$\text{After 2nd teleportation: } |\psi_2\rangle = X^{m_2} \mathcal{H} Z_{\pm\alpha_2} X^{m_1} \mathcal{H} Z_{\alpha_1} |+\rangle \quad (5.129)$$

as shown earlier. Of course, m_1, m_2 are the outputs of the measurements on the first and second qubits, respectively.

Observe further that feedforward can be used to choose the sign of $\pm\alpha_2$ so that $Z_{\pm\alpha_2} X^{m_1} = X^{m_1} Z_{\alpha_2}$. Also, it is well-known that $\mathcal{H} X^{m_1} = Z^{m_1} \mathcal{H}$, so we can write the output as

$$\begin{aligned} X^{m_2} \mathcal{H} Z_{\pm\alpha_2} X^{m_1} \mathcal{H} Z_{\alpha_1} |+\rangle &\rightarrow X^{m_2} Z^{m_1} \mathcal{H} Z_{\alpha_2} \mathcal{H} Z_{\alpha_1} |+\rangle \\ &= X^{m_2} Z^{m_1} (\mathcal{H} Z_{\alpha_2} \mathcal{H} Z_{\alpha_1} |+\rangle). \end{aligned} \quad (5.130)$$

We see that, up to known Pauli matrices X^{m_2}, Z^{m_1} , the output is identical to the output of the circuit



which we introduced earlier.

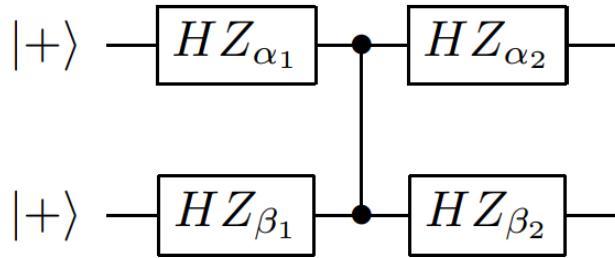
Remark: We can (easily) generalize this example to larger single-qubit circuits containing gates of the form $\mathcal{H} Z_{\alpha_j}$. Here’s the general proof:

- Rewrite the cluster-state computation in terms of an equivalent quantum circuit.
- Reinterpret the quantum circuit as a sequence of circuits of the *teleportation form*, i.e., move the controlled-Z’s around to make the highlighted boxes.

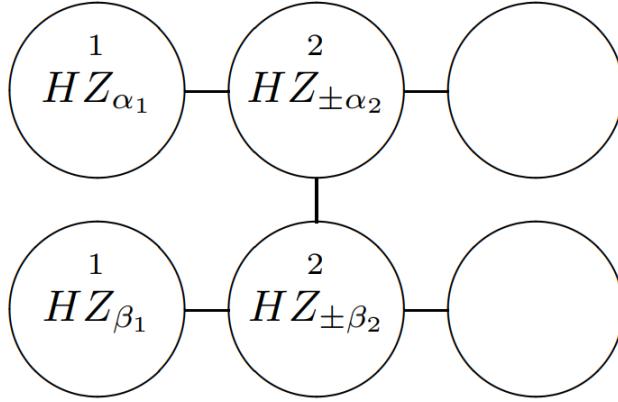
- In the resulting expression for the output state, commute operators of the form X^m all the way to the left, using feedforward to choose signs on the terms of the form $Z_{\pm\alpha}$ to ensure that after commutation they are of the form Z_α .
- The result is a state which, up to known factors of Pauli matrices, is equivalent to the output of the single-qubit quantum circuit.

Multi-qubit case

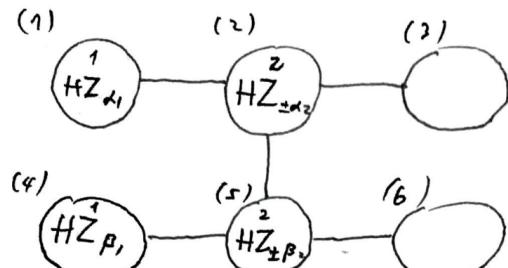
These ideas generalize to multi-qubit quantum circuits. For example, consider the quantum circuit:



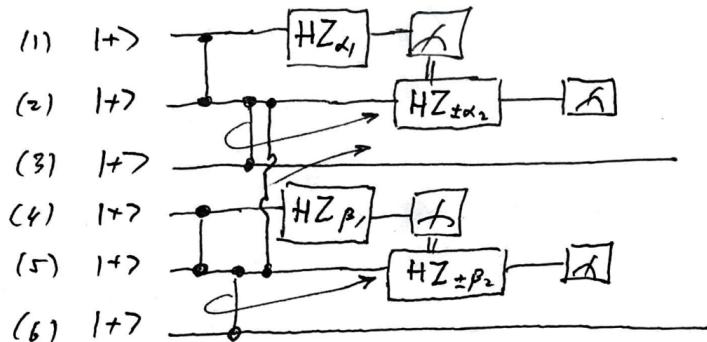
I claim that this quantum circuit can be simulated using the cluster-state computation



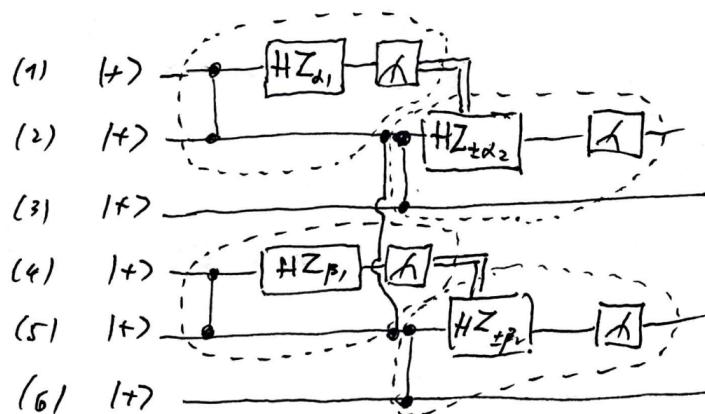
We notice a (weak) pattern here. The number of qubits necessary for the cluster to simulate a 2-qubit quantum circuit is now 6. (last time it was 3 qubits to simulate a 1-qubit circuit). In any case, the proof of this equivalence follows exactly the same lines as in the single-qubit case, and is notationally more complicated. It suffices to prove this diagrammatically. To do this, we first label the qubits like so:



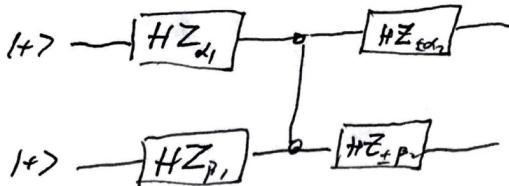
By definition, the cluster state computation has the following associated quantum circuit:



Since the controlled-PHASE/controlled-Z can be moved around (as we have argued before), this circuit is equivalent to



Based on the highlighted boxes, we can see that this quantum circuit is equivalent to



With this we're done.

So, summing p, we have shown how the cluster-state model of computation can be used to efficiently simulate any quantum circuit whose inputs are all $|+\rangle$ states, and whose gates are either controlled-PHASE gates, or gates of the form $\mathcal{H}Z_\alpha$. This set of resources is *universal* for quantum computation, and thus the cluster-state model is capable of efficiently simulating any quantum circuit.

Conversely, it is straightforward to see that any cluster-state model can be efficiently modeled by a quantum circuit model, and thus the two models are computationally equivalent.

5.14.3 Properties

Now we discuss some properties of the cluster-state model of computation. We address two simple but fundamental questions about cluster states, and present some progress on answering these questions:

- Can cluster states arise as ground states of some naturally occurring Hamiltonian?
- What general properties of a quantum state enable it to serve as a substrate for quantum computation, in a manner similar to the cluster state?

Cluster states and ground states of many-body quantum systems

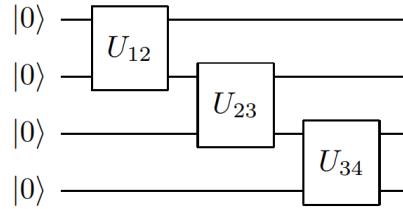
We won't focus too much on the details in this section. I will just focus on the fact that cluster states are in general not ground states of naturally/realistic Hamiltonians. However, this does not mean we can't prepare them in a lab or on a digital quantum computer. In particular, clusters arising in simulations of non-trivial multi-qubit quantum circuits cannot be non-degenerate ground states.

The question of obtaining a general classification of which clusters satisfy these conditions is still open. It would be of great interest to develop a general understanding of which clusters can and cannot arise as ground states of physically reasonable Hamiltonians.

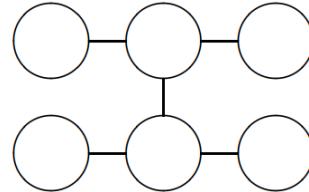
Linearly assembled quantum states can be efficiently simulated on a quantum computer

Now we look at what makes cluster states useful substrates for quantum computation. We will show that spatial dimension plays a role, proving that quantum states which can be linearly prepared in one dimension can always be simulated efficiently on a classical computer, and thus are **not** useful for quantum computation.

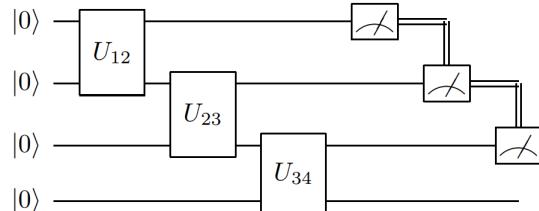
Definition 5.14.1 (Linearly preparable). A linearly preparable quantum state $|\psi\rangle$ can be prepared using a circuit of the form of cascading sequence of two-qubit quantum gates applied to some product starting state like so:



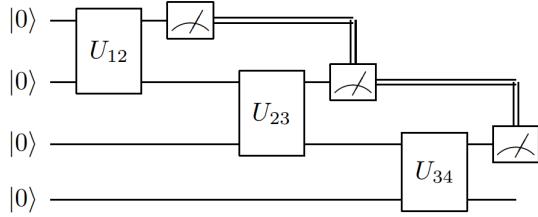
We note that the two-dimensional cluster states we have been using are not linearly preparable:



Now, suppose we perform a sequence of quantum measurements on a linearly assembled quantum state, e.g.,



where earlier measurement results may be used to control later measurement bases. Note that the measurement gates above may be in any single-qubit basis, not just the computational basis. This is equivalent to the following circuit

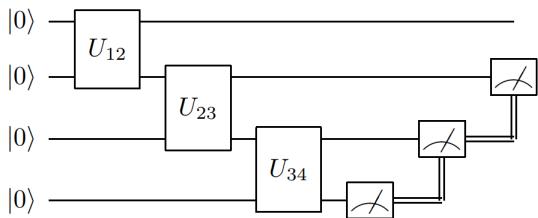


We claim that we can use a classical computer to simulate such a circuit.

- First, we consider the result of applying U_{12} to $|00\rangle$ and compute the probabilities for the measurement on qubit 1.
- We then sample from this distribution to produce a posterior state for qubit 2, which is used as an input for U_{23} .
- Then, repeat this cycle of computing probabilities, sampling, and computing posterior states.
- To make the overall result accurate, it suffices to carry out computations to $\mathcal{O}(\log(1/n))$ bits of precision, and thus the entire computation can be carried out using $\mathcal{O}(n \log^c(1/n))$ operations, where c is some constant arising from the need to multiply floating point numbers.

The point here is that efficient classical simulation of such a computation can be performed.

An elaboration of this argument takes care of the case where we measure the qubits in some other order, for example:



I won't go into the details here.

This efficient classical simulation procedure can easily be generalized to a sequence of measurements performed in any order on the qubits including adaptive orderings. It is also straightforward to extend these results to some simple variants on the linear topology, including rings, and a small number of interlinking rings.

It remains an interesting open problem to determine what class of states can be used as a substrate for quantum computation. The examples in this subsection indicate that the geometry underlying the preparation procedure plays a key role.

5.15 Mesurement-based quantum computing: Jozsa

This [paper](#) by Richard Jozsa introduces two schemes of measurement-based quantum computation: the one-way quantum computer (1WQC) (or equivalently the cluster state model) and the teleportation quantum computation (TQC). The former was formulated by Briegel et al while the latter by Gottesman, Chuang, Nielsen, and Leung. While I want to discuss both models, we will want to read this paper with a focus on *cluster state computation*. It turns out that there's a close relationship between these models, so it's probably useful to know both. But we'll see how much we want to understand teleportation-based QC.

Remember that in the formalism of measurement based quantum computation we start with a given fixed entangled state of many qubits and perform single-qubit measurements to designed qubits in designated bases. The choice of basis for later measurements may depend on earlier measurement outcomes and the final result of the computation is determined from the classical data of all the measurement outcomes. This is in contrast to the more familiar gate array model in which computational steps are unitary operations.

In contrast to unitary evolution, measurements are irreversibly destructive, involving much loss of potential information about a quantum states identity. Thus it is interesting, and at first sight surprising, that we can perform universal quantum computation using only measurements as computational steps. These “measurement based” models are especially interesting for fundamental issues: they have no evident classical analogues and they offer a new perspective on the role of entanglement in quantum computation.

5.15.1 Review

It's never a bad time to review the ubiquitous gates used in this QC. First, the Pauli matrices:

$$I = \mathcal{I}, \quad X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad iY = ZX = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}. \quad (5.131)$$

\oplus denotes addition mod 2. The controlled-NOT gate is given by

$$CX |i\rangle |j\rangle = |i\rangle |i \oplus j\rangle. \quad (5.132)$$

The controlled-PHASE is given by

$$CZ |i\rangle |j\rangle = (-1)^{ij} |i\rangle |j\rangle. \quad (5.133)$$

Note that in contrast to the CX gate, CZ is symmetrical in two qubits. The Hadamard is given by

$$\mathcal{H} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad (5.134)$$

with the basis states $|\pm\rangle$. The Bell basis states are

$$\begin{aligned} |B_{00}\rangle &= \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = I \otimes I |B_{00}\rangle \\ |B_{01}\rangle &= \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle) = X \otimes I |B_{00}\rangle \\ |B_{10}\rangle &= \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle) = Z \otimes I |B_{00}\rangle \\ |B_{11}\rangle &= \frac{1}{\sqrt{2}}(|10\rangle - |01\rangle) = iY \otimes I |B_{00}\rangle \end{aligned} \tag{5.135}$$

Succinctly, we have

$$|B_{cd}\rangle = Z^c X^d \otimes I |B_{00}\rangle. \tag{5.136}$$

The maximally entangled states is given by

$$|\mathcal{H}\rangle = \frac{1}{\sqrt{2}}(|0+\rangle + |1-\rangle) = \frac{1}{\sqrt{2}}(|+0\rangle + |-1\rangle) = CZ |++\rangle. \tag{5.137}$$

5.15.2 Adaptive Measurements

In this section we introduce some identities that are useful for us when we perform adaptive measurements. Our goal is to elaborate on how one makes *adaptive measurements*, i.e., make measurements on some basis such that the state acquired by the subsequent qubit is “nice.” Here, “nice” means that we can use known rules of Pauli propagation (e.g. $HXH = Z$) or commute things around, and so on.

Here’s the idea. Suppose we wish to perform arbitrary sequences of gates from a universal set by measurements. For simplicity, consider 1-qubit gates: to perform $\dots U_3 U_2 U_1 |\psi\rangle$ we successively teleport the three gates but instead of the described result we get something like $\dots P_3 U_3 P_2 U_2 P_1 U_1 |\psi\rangle$ where P_i are the Pauli operators depending on the measurement outcomes To deal with this awkwardness we introduce *adaptive choices* of measurements. This was introduced in the previous section as “feeding forward the result of the last measurement into the next measurement.”

Let us assume that all 1-qubit operations are x or z rotations. The following Pauli propagation relations are easily verified:

$$R_x(\theta)X = XR_x(\theta), \quad R_x(\theta)Z = ZR_x(-\theta), \quad R_z(\theta)X = XR_z(-\theta), \quad R_z(\theta)Z = ZR_z(\theta). \tag{5.138}$$

These make sense: rotating in the same axis commutes. Rotating in the different axes don’t commute. To swap the order of rotation, we also have to flip the

direction of the rotation, so that it undoes the parity change.

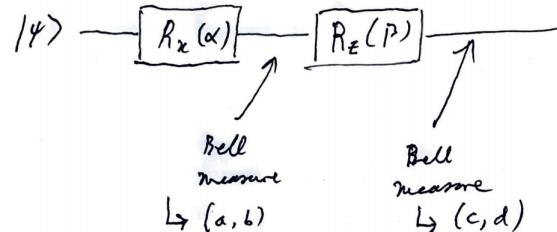
We also have

$$W(\theta)X = ZW(-\theta), \quad W(\theta)Z = XW(\theta) \quad (5.139)$$

where

$$W(\theta) = HP(\theta) = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & e^{i\theta} \\ 1 & e^{-i\theta} \end{bmatrix}. \quad (5.140)$$

Now, suppose we want to do $R_z(\beta)R_x(\alpha)|\psi\rangle$. Now, there's a lot of shorthand notation here, but the important thing we have to keep in mind is that $|\psi\rangle$ is actually initially adjoined by $|B_{00}\rangle = (1/\sqrt{2})(|00\rangle + |11\rangle)$. With this, it makes sense to make a Bell measurement on the qubit pair and get teleportation (in the Jozsa sense). Consider the object we want to evaluate:



After the first unitary, the state of interest becomes $R_x(\alpha)|\psi\rangle$. A Bell measurement gives outcomes a, b and returns the teleported state $Z^a X^b R_x(\alpha)|\psi\rangle$ where $a, b \in \{0, 1\}$. The proof of this are everywhere (and quite easy), so let's not worry about the details.

Now we notice that commuting $Z^a X^b$ to the left through $R_z(\beta)$ causes β to change to $(-1)^b \beta$. So, if instead of $R_z(\beta)$ we next applied the Bell measurement for $R_z((-1)^b \beta)$ then

$$R_z((-1)^b \beta) Z^a X^b R_x(\alpha) |\psi\rangle = Z^a R_z((-1)^b \beta) X^b R_x(\alpha) |\psi\rangle = Z^a X^b R_z(\beta) R_x(\alpha) |\psi\rangle \quad (5.141)$$

as wanted. Suppose this second measurement with basis adapted to measurement outcome b has an outcome c, d say, and the state is now $Z^{a+c} X^{b+d} R_z(\beta) R_x(\alpha) |\psi\rangle$. Continuing in this way, adapting measurement bases to earlier measurement results we get

$$\dots X_2^{m_2} Z_2^{n_2} X_1^{m_1} Z_1^{n_1} (\text{the correct wanted U}) |\psi\rangle. \quad (5.142)$$

Here $|\psi\rangle$ is a state of many qubits and X_i, Z_i are Pauli operators on the i th qubit. The indices m_i, n_i are accumulations (sum mod 2) of measurement outcomes.

The same idea applies to the 2-qubit gate CZ too where the situation is even better. We have the propagation relations:

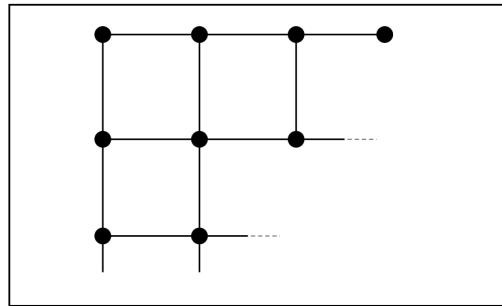
$$CZ(Z \otimes I) = (Z \otimes I)CZ, \quad CZ(X \otimes I) = (X \otimes Z)CZ. \quad (5.143)$$

Thus we can propagate Pauli operators through CZ while keeping CZ the same, so no basis adaption is required.

If U is a total unitary effect of a gate array on a multi-qubit state $|\psi\rangle$ then using the above methods we are able to generate a state of the form $\dots X_2^{m_2} Z_2^{n_2} X_1^{m_1} Z_1^{n_1} U |\psi\rangle$. In a quantum computation we finally measure at least some qubits of $U |\psi\rangle$ in the Z_i bases $\{|0\rangle, |1\rangle\}$. Now, notice that $Z_i^{n_i}$ has no effect on the measurement outcomes (only phases are shifted) and $X_i^{m_i}$ requires only some simple re-interpretation of the output results (since all X does is flipping bits). This means we simple need to reinterpret measurement outcome $k_i \in \{0, 1\}$ as $k_i \oplus m_i$ where $m_i \in \{0, 1\}$ is the exponent of the corresponding X Pauli matrix.

5.15.3 Cluster-state computation/1WQC

We just looked at this model in the previous section. However, in this section we will go into a little more detail. Remember that this model was introduced by Raussendorf and Briegel. The model starts with a 2D grid of $|+\rangle$ states, followed by applications of CZ to each nearest neighbor pair in vertical and horizontal directions. The CZ 's commute so they can be applied in parallel, as a process of constant quantum depth. The resulting state is an entangled states of many qubits, called a *cluster state*.



As we have seen, any quantum gate array can be implemented as a pattern of 1-qubit measurements on a sufficiently large 2D cluster state. The only measurements used are in the bases $M_z = \{|0\rangle, |1\rangle\}$ and $M_\theta = \{|0\rangle \pm e^{i\theta} |1\rangle\}$ for some θ . $\theta = 0$ corresponds to the X basis measurements $\{|+\rangle, |-\rangle\}$. Measurement outcomes are always labeled 0 or 1 and they are always uniformly random.

Also, as we have seen before, quantum gates are implemented only up to known Pauli correction matrices $X^a Z^b$ where a, b depend on the measurement

outcomes. Recall that this is due to teleportation. In the 1WQC literature, these are called the *bi-product Pauli operators*. Of course, the quantum computation is “one way” in the sense that the initial resource of the pure cluster state is irreversibly degraded as the computation proceeds in its layers of measurement.

Example 5.15.1. This first example is take from this [paper](#) by Briegel et al. For this example, we look at measurement patterns for 1-qubit gate (where a 1-dimensional cluster state suffices, as we have seen).

We know that any 1-qubit unitary U can be expressed as a three alternating rotations:

$$U = R_x(\zeta)R_z(\eta)R_x(\xi), \quad (5.144)$$

for some ζ, η, ξ where of course

$$R_x(\theta) = e^{-i\theta X}, \quad R_y(\theta) = e^{-i\theta Y}, \quad R_z(\theta) = e^{-i\theta Z} \quad (5.145)$$

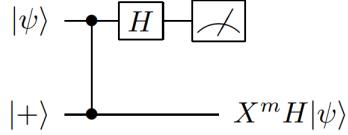
are rotation matrices (in the Bloch sphere). This is quite intuitive, so I won’t explain why this is the case. Now, to apply U to $|\psi\rangle$ by the 1WQC we start with $|\psi\rangle$ in a line with 04 $|+\rangle$ qubits like this:

qubit number	1	2	3	4	5
states	$ \psi\rangle$	$ +\rangle$	$ +\rangle$	$ +\rangle$	$ +\rangle$
entangle with CZ	*	●	●	●	●
measurements	X	$M(-\xi(-1)^{s_1})$	$M(-\eta(-1)^{s_2})$	$M(-\zeta(-1)^{s_3+s_4})$	
outcomes	s_1	s_2	s_3	s_4	

Remember that because there are 3 total rotations, we will need four $|+\rangle$ “ancillary” qubits. Next, entangle all neighboring pairs with CZ gates and subsequently measure qubits 1, 2, 3, 4 adaptively in the bases shown in the figure. It may be shown that qubit 5 acquires the state $X^{s_2+s_4}Z^{s_1+s_3}U|\psi\rangle$, where s_i is the outcome of the measurement on qubit i .

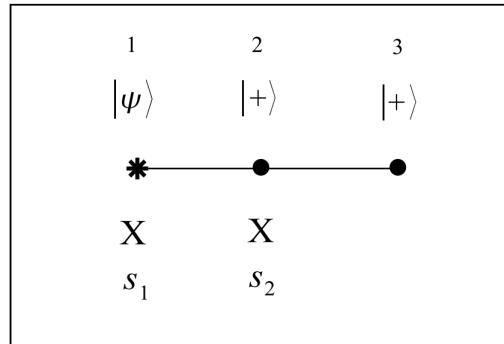
We won’t worry about the details of this example now, for we will return to this paper by Briegel et al later. But the idea is this. Recall that in the teleportation circuit we have something like

where a rotation by Z_α gives a state of the form $X^m|\phi\rangle$. By a similar argument, because $HXH = Z$, we can argue that a rotation by X_α gives a state of of the form Z^m . This means that after the first measurement in the X



basis, the resulting state is going to be Z^{s_1} . Next, we have a rotation by R_z , so the next factor, roughly speaking, is going to be X^{s_2} , and so on. Of course, the reality is that commutation rules also dictate which power goes where in the bi-product operator. So, a careful check is necessary. In any case, we'll come back to this in the Briegel's paper, which is a more mathematically rigorous approach to all of this.

Example 5.15.2. In this example we consider a simpler pattern:



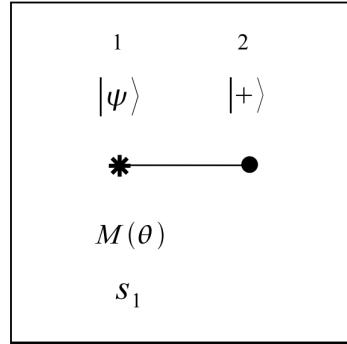
This is more familiar to us at this point. Again, by exchanging the role of Z and X between our previous examples and this one, we can see that qubit 3 acquires the state $Z^{s_1} X^{s_2} |\psi\rangle$ (remember, this is just teleportation). The two X measurements may be done in parallel but we can also consider this process as a sequence of two identical steps as we have seen: given a state $|\psi\rangle$, adjoint this to $|+\rangle$, entangle them with CZ and then X -measure the first qubit, giving an outcome $s \in \{0, 1\}$. Qubit 2 is then left in the state $X^s \mathcal{H} |\psi\rangle$. This means the two step chain ends up in the state

$$(X^{s_2} \mathcal{H})(X^{s_1} \mathcal{H}) |\psi\rangle = X^{s_2} Z_{s_1} |\psi\rangle. \quad (5.146)$$

where we have used the fact that $\mathcal{H}X\mathcal{H} = Z$.

Example 5.15.3. Here we look at a single step operation for the general measurement basis $M(\theta)$.

The state $|\psi\rangle$ and $|+\rangle$ are entangled by the CZ . Measuring the first qubit in the basis $M(\theta) = \{|0\rangle \pm e^{i\theta} |1\rangle\}$ leaves qubit 2 in the state $X^{s_1} W(-\theta) |\psi\rangle$.



Here,

$$W(\theta) = HP(\theta) = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & e^{i\theta} \\ 1 & -e^{i\theta} \end{bmatrix}. \quad (5.147)$$

Let's check this to make sure we're still on track. According to the circuit, we first have a CZ between $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ and $|+\rangle$. This gives a new state

$$\begin{aligned} CZ|\psi+\rangle &= CZ \frac{1}{\sqrt{2}} [\alpha|00\rangle + \alpha|01\rangle + \beta|10\rangle + \beta|11\rangle] \\ &= \frac{1}{\sqrt{2}} [\alpha|00\rangle + \alpha|01\rangle + \beta|10\rangle - \beta|11\rangle]. \end{aligned} \quad (5.148)$$

To check that after measuring in the basis $M(\theta) = \{|0\rangle \pm e^{i\theta}|1\rangle\}$ we get $X^{s_1}W(-\theta)|\psi\rangle$ it suffices to check that

$$CZ|\psi+\rangle = (|0\rangle + e^{i\theta}|1\rangle)X^0W(-\theta)|\psi\rangle + (|0\rangle - e^{i\theta}|1\rangle)X^1W(-\theta)|\psi\rangle. \quad (5.149)$$

This is not hard to check, so I'll leave this to the reader. The point is to show that qubit 2 acquires the state $X^{s_1}W(-\theta)|\psi\rangle$.

The point of this example is to show that if we are not concerned with issues of parallelizability then any 1D measurement pattern may be viewed as a sequential application of the process this example, measurements in the $M(\theta)$ basis, applied repeatedly.

This shows that the unitary operation $W(\theta)$ plays a fundamental role in 1WQC/cluster state computation. For instance, in the previous example, we can reorder all operations as follows: (entangle 12, measure 1), (entangle 23, measure 2), and so on. This is because the 12-entangling operation and first measurement both commute with all subsequent entangling operations and measurements. Of course, the basis $M(\theta)$ is identically X if $\theta = 0$.

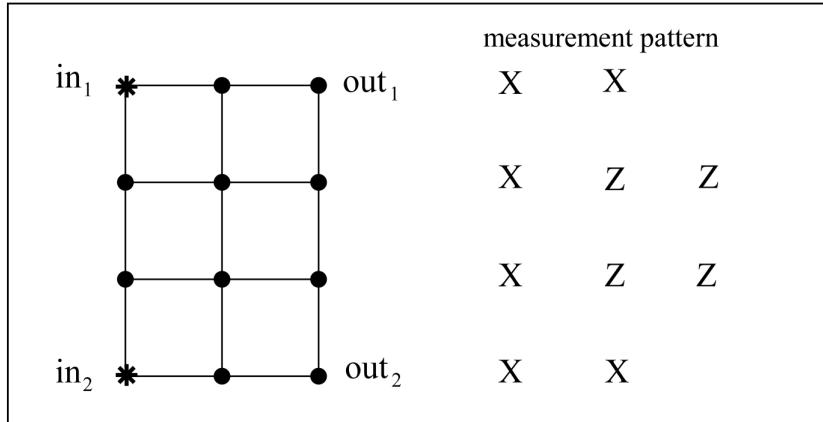
From there examples, we can concatenate gates. Suppose we wish to apply U_1 then U_2 to some input $|\psi\rangle$. Now, measurement pattern 1 for U_1 gives an

qubit number	1	2	3	4	5
states	$ \psi\rangle$	$ +\rangle$	$ +\rangle$	$ +\rangle$	$ +\rangle$
entangle with CZ	*	●	●	●	●
measurements	X	$M(-\xi(-1)^s)$	$M(-\eta(-1)^{s_2})$	$M(-\zeta(-1)^{s_1+s_3})$	
outcomes	s_1	s_2	s_3	s_4	

output that is the input for U_2 . But all measurements in pattern 1 commute with all entangling operations of pattern 2 (recall we can move things the entangling and measurement 1 around in the previous section). So, we can apply all entangling operations for both patterns first, to get a longer single cluster state and then apply the measurements.

Finally, we can eliminate the input state $|\psi\rangle$ altogether, to get a formalism based entirely on a starting state that is a fully standard cluster state (or slightly longer length). This is because we can implement any 1-qubit unitary U , we can take out input starting state to be $|+\rangle$ and prefix the desired process with an initial measurement pattern for a unitary operation that takes $|+\rangle$ to $|\psi\rangle$.

This 1D formalism may be generalized using 2D cluster grids to incorporate 2-qubit gates. For universal computation it suffices to be able to implement just the CZ gate in addition to 1-qubit gates. For example:

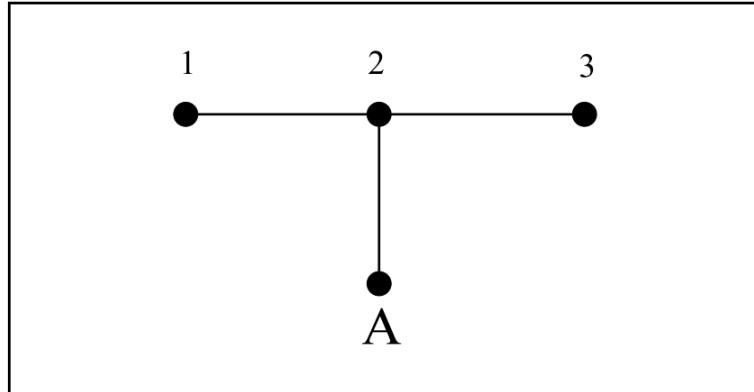


The dots denote the $|+\rangle$ states and connecting lines denote CZ applications between the dots. If the measurement pattern shown on the right is applied at the sites, the only unmeasured sites are of interest and contain $(P_1 \otimes P_2)CZ |\psi_{in}\rangle$ where $P_1 \otimes P_2$ is a Pauli operation that depends on the measurement outcomes.

The Role of Z measurements

Measurement patterns for general gate arrays, which are implemented on a sufficiently large 2D cluster, will generally have extraneous sites not used in the measurement patterns. Z -measurements are used to delete such extraneous sites. To see how this works, we consider two examples.

Example 5.15.4. Consider the following cluster:



This has quite an irregular shape. Suppose we only want to use sites 1, 2, 3, i.e., we want to delete site A . Consider Z_A , a Z measurement at site A , with outcome k_A . To see its effect we need to write down what the state resulting from figure 9 is. Now, measurement Z_A commutes with CZ_{12} and CZ_{23} . However,

$$CZ_{A2} |++\rangle = |0+\rangle + |1-\rangle. \quad (5.150)$$

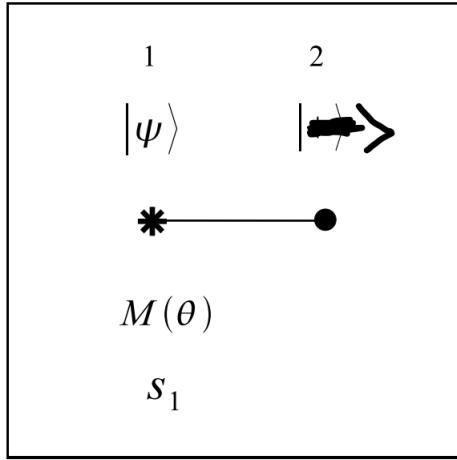
So, after the Z_A measurement (with outcome k_A) the sites 1, 2, 3 contain $|+\rangle |(-1)^{k_A}\rangle |+\rangle$, where $|(-1)^{k_A}\rangle$ is one of the two Hadamard basis states:

$$\sim [1 \mp \sqrt{2} \quad 1]^\top \quad (5.151)$$

This means that the effect of the Z_A measurement is to have used the Hadamard basis state $|\pm\rangle = |(-1)^{k_A}\rangle$ at site 2 instead of the standard $|+\rangle$ and then entangle as usual.

So, to summarize, by applying the measurement Z_A , we end up with the total state $|+(-1)^{k_A}+\rangle$ at 1, 2, 3, and get no more entanglement with A .

Example 5.15.5. Consider the following figure again:



Suppose we start with $|\psi-\rangle$ instead of with $|\psi+\rangle$ as before. Applying CZ , then measure $M(\theta)$ at qubit 1 for output at qubit 2. The reader can check that this is the same as (i) start with $|\psi+\rangle$, (ii) apply CZ , (iii) measure $M(\theta)$ on qubit 1, (iv) applying the unitary $I \otimes Z_2$. Now, since Z_2 commutes with CZ and $M(\theta)$, we can apply $I \otimes Z_2$ before applying CZ and note that

$$I \otimes Z_2 |\psi+\rangle = |\psi-\rangle. \quad (5.152)$$

In summary, we see that the measurement Z_A at site A with outcome k_A adds in an extra Z^{k_A} Pauli correction at all neighboring sites of A , in addition to the usual Pauli operations arising from measurement patterns in a standard cluster state that had site A absent from the start.

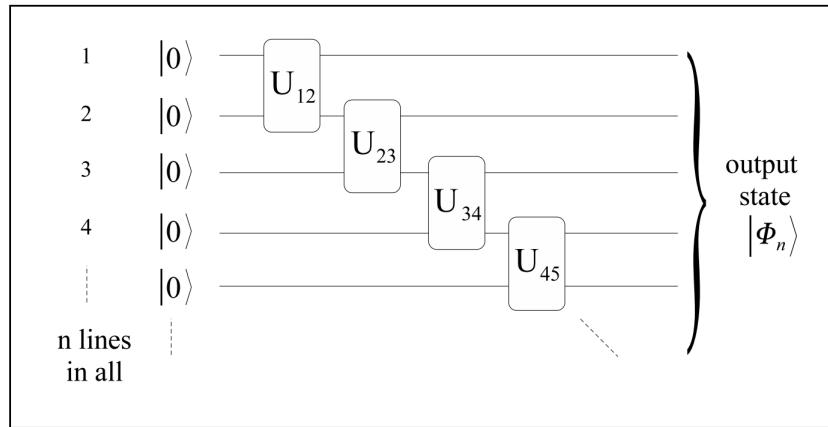
Summary

To summarize the cluster state/1WQC model, we have seen that

- Any quantum gate array can be translated into a pattern of 1-qubit measurements on a sufficiently large 2D cluster state.
- Choices of measurement bases are generally adaptive: possibly depending on previous measurement outcomes.

5.15.4 Selected extra features: Cluster state model

Theorem 5.15.1. 1WQC/Cluster state model with 2D cluster states is universal for quantum computation.



However, as we have discussed, any 1WQC/cluster state process based on only 1D cluster states can be simulated classically efficiently (in polynomial time in the number of qubits). Thus, the universality of any such model would imply that quantum computation is no more powerful than classical computation.

5.16 QAOA with Cluster States

In this section, we demonstrate that it is possible to explicitly translate the QAOA/Trotterization algorithm into the cluster state language. To this end, we first will show that it is possible to construct a universal set of gates for quantum computation using cluster states. Next, we apply this universal set of gates to run the QAOA to prepare critical ground state of TFIM.

5.16.1 Universal gates with CC

to be updated soon

Rotation in X

to be updated soon

Rotation in Z

to be updated soon

CNOT

to be updated soon

5.16.2 Cluster States for TFIM state preparation

to be updated soon

5.17 Variational Imaginary Time Ansatz (VITA)

In this section we look this [paper](#) by Tim Hsieh et al. The reason we want to do this is because this method motivates our development of the cluster state model to make a more efficient (faster than linear) critical state preparation protocol. Recall that the real time Trotterization QAOA protocol scales linearly in the number of qubits. This VITA allows us to go as $\log L$ where L is the system size.

The paper introduces a variational wave function for many-body ground states that involves imaginary time evolution with two different Hamiltonians in an alternating fashion with variable time intervals. The VITA's on one- and two-dimension TFIM are studied and the scaling for the 1-dimension model at criticality is studied. The total imaginary time required scales logarithmically with system size, in contrast to the linear scaling in conventional quantum Monte Carlo. This is perhaps due to unique dynamics permitted by alternating imaginary time evolution, including exponential growth of bipartite entanglement.

5.17.1 Introduction

Recent developments of quantum computers and simulators has motivated variational approaches which involve applying a sequence of unitary operators, parametrized by several variables onto an easy-to-prepare initial state. The parameters are chosen to minimize some cost function. Examples of these include the QAOA/VQCS method where we alternate between evolving with the Hamiltonian and a transverse field.

Motivated by this, the paper proposes VITA which involves sequentially evolving with different Hamiltonians in imaginary time. As proof of concept, the paper demonstrate the efficiency of the ansatz in representing the ground state of the TFIM at criticality.

5.17.2 The Ansatz

Many Hamiltonians are naturally a linear combination of two components

$$H = H_a + gH_B \quad (5.153)$$

where H_A, H_B are individually tractable to analyze. The TFIM Hamiltonian is one example of such Hamiltonians. Motivated by the QAOA procedure, we consider the following variational imaginary time ansatz (VITA) for the ground state of H :

$$|\psi_p(\boldsymbol{\alpha}, \boldsymbol{\beta})\rangle = \mathcal{N} e^{-\beta_P H_B} e^{-\alpha_P H_A} \dots e^{-\beta_1 H_B} e^{-\alpha_1 H_A} |\psi_0\rangle \quad (5.154)$$

where P is the number of pairs of variational parameters $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_P), \boldsymbol{\beta} = (\beta_1, \dots, \beta_P)$, $|\psi_0\rangle$ is an initial state, and \mathcal{N} is a normalization factor. The total

imaginary time is defined as

$$\tau = \frac{1}{2} \sum_{p=1}^P (\alpha_p + \beta_p). \quad (5.155)$$

The standard projector method for attaining the ground state of H involves evaluating $e^{-\tau H}$ for $\tau \sim 1/\Delta$ where Δ is the many-body spectral gap. This can be decomposed via Trotterization into a sequence of the VITA form with parameters $\alpha_p = \beta_p = \tau/2P$ for large P , which guarantees that VITA can exactly represent the ground state in the $P \rightarrow \infty$ limit. However, this method is expensive for critical systems where $\Delta \sim 1/L$ and hence τ scales polynomially with L .

Thus, one can consider the VITA form as a nonuniform Trotterization with large and variable time steps. It turns out the remarkably high fidelities can be attained even with τ that is exponentially smaller compared to the aforementioned estimate from the standard projector method.

5.17.3 VITA on TFIM

Consider the TFIM Hamiltonian:

$$H = H_{ZZ} + hH_X = - \sum_{i=1}^N Z_i Z_{i+1} - h \sum_{i=1}^N X_i \quad (5.156)$$

with periodic boundary conditions on a system with N spins. We start in the ground state of H_X , $\bigotimes |+\rangle$, and alternate between H_{ZZ} and H_X . For a fixed P , optimization involves finding the minima of the energy cost function

$$E_p(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \langle \psi_p(\boldsymbol{\alpha}, \boldsymbol{\beta}) | H | \psi_p(\boldsymbol{\alpha}, \boldsymbol{\beta}) \rangle. \quad (5.157)$$

We can also compute the error in fidelity:

$$1 - f = 1 - |\langle \psi_{\text{exact}} | \phi_p(\boldsymbol{\alpha}, \boldsymbol{\beta}) \rangle|^2. \quad (5.158)$$

It turns out the efficiency is remarkable. For $L = 64$, $P = 2$ is already sufficient to approximate the critical state to within around 10^{-4} in relative energy and 10^{-2} in fidelity. Recall that the error in fidelity provides an upper bound for the error in any observable.

5.18 The Jordan-Wigner Transformation

In this section we study the Jordan-Wigner transformation, which is an important tool for analyzing the TFIM with nearest-neighbor interactions. This is interesting because by being able to map a system of qubits to a system of Fermions, anything we understand about one type of system can be immediately applied to learn something about the other type of system.

First, we will consider the motivation for the transformation and learn about the basics of the Jordan-Wigner transformation from this [notes](#) by Michael Nielsen.

Then, we will look at how the Jordan-Wigner transformation is often written in the literature, based on materials from various sources.

After that, we might explore what the Jordan-Wigner transformation is applied in context of QAOA and perhaps measurement-based quantum computation.

5.18.1 Fermionic canonical commutation relations

Introduction

Suppose we have a set of operators a_1, \dots, a_n acting on some Hilbert space V . We say that these operations satisfy the *canonical commutation relations (CCRs) for Fermions* if

$$\{a_j, a_k^\dagger\} = \delta_{jk} I; \quad \{a_j, a_k\} = 0 \quad (5.159)$$

where

$$\{A, B\} = AB + BA \quad (5.160)$$

is the anti-commutator. Note that when we take the conjugate of the second of these relations we find

$$\{a_j^\dagger, a_k^\dagger\} = 0 \quad (5.161)$$

which is sometimes also referred to as one of the CCRs. It is also frequently useful to set $j = k$, giving $a_j^2 = (a_j^\dagger)^2 = 0$. This is also the statement of the Pauli exclusion principle.

One way of thinking about the CCRs is in an axiomatic mathematical sense. We may think of these conditions as mathematical conditions that can be imposed on a set of matrices: for a given set of matrices, we can simple check and verify whether those matrices satisfy the CCRs.

For instance, when the state space V is that of a single qubit, we can easily verify that the operator $a = |0\rangle\langle 1|$ satisfies the Fermionic CCRs. Now, the

question is what consequences about the structure of V and the operators a_j can be deduced from the fact that the CCRs hold.

Of course, there's also a representation-theoretic way of understanding things, but let's not go there.

Finally, the physical way to understanding the CCRs is the following. When we want to describe a system of Fermions, one way to begin is to start by writing down a set of operators satisfying the CCRs, and then try to guess what sort of Hamiltonian involving those operators could describe the interactions observed in the system.

We shall approach the Fermionic CCRs from a purely mathematical point of view, asking the question “What can we deduce from the fact that a set of operators satisfying the CCRs exists?” It turns out that we can deduce quite a lot about the structure of V and a_j simple from the fact that the a_j ’s satisfy the CCRs.

Consequences of the Fermionic CCRs

Assume that V is finite dimensional, and that there are n operators a_1, \dots, a_n on V satisfying the Fermionic CCRs. We will do the following things:

- Start by showing that the operators $a_j^\dagger a_j$ are positive Hermitian operators with eigenvalues 0, 1.

Proof. Observe that the $a_j^\dagger a_j$ are positive matrices. Now, since $(a_j^\dagger a_j)^2 = a_j^\dagger a_j$:

$$(a_j^\dagger a_j)^2 = a_j^\dagger a_j a_j^\dagger a_j = a_j^\dagger (1 - a_j^\dagger a_j) a_j = -(a_j^\dagger)^2 a_j^2 + a_j^\dagger a_j = a_j^\dagger a_j \quad (5.162)$$

where we have used the CCRs $\{a_j^\dagger, a_j\} = 1$ and $\{a_j, a_j\} = 0$. Thus, the eigenvalues of $a_j^\dagger a_j$ are 0 and 1. \square

- a_j acts as a *lowering operator* for $a_j^\dagger a_j$, in the sense that if $|\psi\rangle$ is a normalized eigenstate of $a_j^\dagger a_j$ with eigenvalue 1 then $a_j |\psi\rangle$ is a normalized eigenstate of $a_j^\dagger a_j$ with eigenvalue 0. If $|\psi\rangle$ is normalized eigenstate of $a_j^\dagger a_j$ with eigenvalue 0, then $a_j |\psi\rangle$ vanishes.

Proof. Let a normalized $|\psi\rangle$ be given, and $a_j^\dagger a_j |\psi\rangle = |\psi\rangle$, that is $|\psi\rangle$ is the 1-eigenstate of $a_j^\dagger a_j$. Then

$$\langle \psi | a_j^\dagger a_j | \psi \rangle = \langle \psi | \psi \rangle = 1 \quad (5.163)$$

so $a_j |\psi\rangle$ is normalized. Now, note that because $\{a_j, a_j\} = 0$, we have

$$(a_j^\dagger a_j) a_j |\psi\rangle = 0. \quad (5.164)$$

Finally, suppose $a_j^\dagger a_j |\psi\rangle = 0$, then

$$a_j |\psi\rangle = a_j \{a_j^\dagger, a_j\} |\psi\rangle = 0. \quad (5.165)$$

□

- Similarly, a_j^\dagger acts as a *raising operator* for $a_j^\dagger a_j$, in the sense that if $|\psi\rangle$ is a normalized eigenstate of $a_j^\dagger a_j$ with eigenvalue 0 then $a_j^\dagger |\psi\rangle$ is a normalized eigenstate of $a_j^\dagger a_j$ with eigenvalue 1. If $|\psi\rangle$ is normalized eigenstate of $a_j^\dagger a_j$ with eigenvalue 1, then $a_j^\dagger |\psi\rangle$ vanishes.

Proof. The proofs are not too hard (when in doubt, use the CCRs), so I will leave them to the reader. □

- Prove that the operators $a_j^\dagger a_j$ form a mutually commuting set of Hermitian matrices, and thus there exists a state $|\psi\rangle$ which is a simultaneous eigenstate of $a_j^\dagger a_j$ for all values of j .

Proof. We first want to show that the $a_j^\dagger a_j$ form a mutually commuting set of observables. To see this, let $j \neq k$, and apply the CCRs repeated to obtain

$$a_j^\dagger a_j a_k^\dagger a_k = a_k^\dagger a_k a_j^\dagger a_j \quad (5.166)$$

which is the desired commutativity. From linear algebra, we know that a family of commuting matrices is simultaneously diagonalizable, so we're done. Basically, we simply use the fact that there exists a normalized state $|\psi\rangle$ which is a simultaneous eigenstate of all the $a_j^\dagger a_j$ operators. In particular for all j we have

$$a_j^\dagger a_j |\psi\rangle = \alpha_j |\psi\rangle \quad (5.167)$$

where for each j , $\alpha_j \in \{0, 1\}$. It will be convenient to assume that $\alpha_j = 0$ for all j . This assumption can be made without loss of generality, by applying the lowering operators to the $|\psi\rangle$ for each j such that α_1 , resulting in a normalized state $|\text{vac}\rangle$ such that

$$a_j^\dagger a_j |\text{vac}\rangle = 0 \quad (5.168)$$

for all j . □

- By raising and lowering the state $|\psi\rangle$ in all possible combinations, we'll construct a set of 2^n orthonormal states which are simultaneous eigenstates of the $a_j^\dagger a_j$. The corresponding vector of eigenvalues uniquely labels each state in this orthonormal basis.

Proof. Here we want to define an orthonormal basis. For any vector $\alpha = (\alpha_1, \dots, \alpha_n)$ where each $\alpha_j = 0$ or 1, define a corresponding state

$$|\alpha\rangle = (a_1^\dagger)^{\alpha_1} \dots (a_n^\dagger)^{\alpha_n} |\text{vac}\rangle \quad (5.169)$$

It is clear that there are 2^n such states $|\alpha\rangle$, and that they form an orthonormal set spanning a subspace of V that we call W . \square

- Suppose the vector space spanned by these 2^n simultaneous eigenstates is W . At this point, we know that a_j, a_j^\dagger map $W \rightarrow W$, and indeed, we know everything about the action a_j, a_j^\dagger have on W .

Proof. Here we want to describe the action of a_j, a_j^\dagger on W . Applying the CCRs and the definitions of the states $|\alpha\rangle$ it is easy to check that the action of a_j is as follows:

- If $\alpha_j = 0$, then $a_j |\alpha\rangle = 0 - a_j$ makes the existing lower state vanish.
- If $\alpha_j = 1$, then let α' be that vector which results when the j th entry of α is changed to 0. Then

$$a_j |\alpha\rangle = -(-1)^{s_\alpha^j} |\alpha'\rangle \quad (5.170)$$

where

$$s_\alpha^j = \sum_{k=1}^{j-1} \alpha_k \quad (5.171)$$

For example, if $|\alpha\rangle = |111\rangle$, i.e.,

$$|\alpha\rangle = (a_1^\dagger)^1 (a_2^\dagger)^1 (a_3^\dagger)^1 |\text{vac}\rangle \quad (5.172)$$

then

$$a_2 |111\rangle = -(-1)^1 (a_1^\dagger)^1 (a_2^\dagger)^0 (a_3^\dagger)^1 |\text{vac}\rangle \quad (5.173)$$

which one can verify via the applying the CCR on $a_2 a_1^\dagger$.

In any case, the action of a_j^\dagger on W similar, but quite opposite:

- If $\alpha_j = 0$, then let α' be that vector which results when the j th entry of α is changed to 1. Then

$$a_j^\dagger |\alpha\rangle = -(-1)^{s_\alpha^j} |\alpha'\rangle \quad (5.174)$$

where again

$$s_\alpha^j = \sum_{k=1}^{j-1} \alpha_k. \quad (5.175)$$

- Suppose $\alpha_j = 1$. Then $a_j^\dagger |\alpha\rangle = 0$.

□

- Suppose we define W_\perp to be the orthocomplement of W in V (that is, $V = W \oplus W_\perp$). Then we'll show that the a_j, a_j^\dagger map $W_\perp \rightarrow W_\perp$, and their restrictions to W_\perp satisfy the Fermionic CCRs. Repeating the procedure above, and we'll be able to identify a 2^n -dimensional subspace of W_\perp on which we know the action of the a_j, a_j^\dagger exactly.

Proof. Here we want to look at the action of a_j, a_j^\dagger on W_\perp . We first want to show that they actually map W_\perp to itself. To see this, let $|\psi\rangle \in W_\perp$ and consider $a_j |\psi\rangle$. To show that $a_j |\psi\rangle \in W_\perp$, we want to show that $\forall |\phi\rangle \in W, \langle \phi | a_j |\psi\rangle = 0$. Well, observe that

$$\langle \psi | a_j^\dagger |\phi\rangle = \overline{\langle \phi | a_j |\psi\rangle} = 0 \quad (5.176)$$

because $|\psi\rangle \in W_\perp$ and $a_j^\dagger |\phi\rangle \in W$ since $|\phi\rangle \in W$. The proof for a_j^\dagger is similar. So, a_j, a_j^\dagger map W_\perp to itself. □

- Iterate this procedure until W_\perp is the trivial vector space, at which point it is no longer possible to continue. At this point we have established an orthonormal basis for the whole of V wrt which we can explicitly write down the action of both a_j, a_j^\dagger .

Proof. Now, we see that once we restrict these operators to acting only on W_\perp , so long as W_\perp is nontrivial, we can repeat the argument and show that W_\perp can be decomposed further into $U \oplus U^\perp$ and so on. Since V is finite dimensional, the process will terminate. At the point of termination we will have broken up V as a direct sum of some finite number d of orthonormal 2^n -dimensional vector spaces, W_1, \dots, W_d and on each each vector space we will have an orthonormal basis with respect to which the action of a_j, a_j^\dagger is known precisely.

In other words, we can introduce an ONB $|\alpha, k\rangle$ for V where α runs over all n -bit vectors and $k = 1, \dots, d$ and such that the action of the a_j and a_j^\dagger is to leave k invariant, and to act on $|\alpha\rangle$ as described above. In this representation it is clear that V can be regarded as a tensor product $C^{2^n} \otimes C^d$ with the action of a_j, a_j^\dagger trivial on the C^d component. We will call this the **occupation number representation** for the Fermi algebra a_j . □

Diagonalizing a Fermi quadratic Hamiltonian

Suppose a_1, \dots, a_n satisfy the Fermionic CCRs, and we have a system with Hamiltonian

$$H_{\text{free}} = \sum_j \alpha_j a_j^\dagger a_j, \quad (5.177)$$

where $\alpha_j \geq 0$ for each value of j . In physical terms this is the Hamiltonian used to describe a system of free, i.e., non-interacting Fermions.

The positivity of $a_j^\dagger a_j$ implies that

$$\langle \psi | H_{\text{free}} | \psi \rangle \geq 0, \forall |\psi\rangle \quad (5.178)$$

But we also know that there is at least one $|\text{vac}\rangle$ for which $a_j^\dagger a_j |\text{vec}\rangle = 0$ for all j , so the ground state energy of this Hamiltonian is exactly 0, with $|\text{vac}\rangle$ being the corresponding ground state.

When α_j have arbitrary signs, then we just take the ground state energy of this Hamiltonian to be $\sum_j \min\{\alpha_j, 0\}$, and the ground state $|\psi\rangle$ is obtained from $|\text{vac}\rangle$ by applying a_j^\dagger for all j with $\alpha_j < 0$. More generally, the allowed energies of the excited states of this system correspond to sums over subsets of the α_j .

Now, let us generalize this result to the spectrum of a more general interacting Hamiltonian:

$$H = \sum_{j,k} \alpha_{jk} a_j^\dagger a_k \quad (5.179)$$

For H to be Hermitian, $\alpha_{jk} = \alpha_{jk}^*$, which means they are necessarily real. So, the matrix α whose entries are the α_{jk} is itself Hermitian.

Suppose we introduce new operators b_1, \dots, b_n defined by

$$b_j = \sum_{k=1}^n \beta_{jk} a_k \quad (5.180)$$

where β_{jk} are complex numbers. We are going to try to choose the β_{jk} so that

- The operators b_j satisfy the Fermionic CCRs.

To this end, we look for conditions on the complex number β_{jk} such that the b_j operators satisfy the Fermionic CCRs. Observe that

$$\{b_j, b_k^\dagger\} = \sum_{lm} \beta_{jl} \beta_{km}^* \{a_l, a_m^\dagger\} = \sum_{lm} \beta_{jl} \beta_{km}^* \delta_{lm} I = (\beta \beta^\dagger)_{jk} I \quad (5.181)$$

where $\beta \beta^\dagger$ denotes the matrix product of the matrix β with entries β_{jl} and its adjoint β^\dagger .

To $\{b_j, b_k\}$ we use the linearity of the anti-commutator bracket in each term to express $\{b_j, b_k\}$ as a sum over terms of the form $\{a_l, a_m\}$, each of

which is 0, by the CCRs. As a result, we have

$$\{b_j, b_k\} = 0 \quad (5.182)$$

Thus, provided that $\beta\beta^\dagger = I$, i.e., provided that β is unitary, the operators b_j will satisfy the CCRs.

- When expressed in terms of the b_j , this Hamiltonian will take the form as H_{free} and thus can be diagonalized.

To this end, let us assume that β is indeed unitary, and change our notation, writing $u_{jk}\beta_{jk}$ in order to emphasize the unitarity of this matrix. We now have

$$b_j = \sum_k u_{jk} a_k \quad (5.183)$$

By the unitarity of u we can invert this equation to get

$$a_j = \sum_k u_{jk}^\dagger b_k. \quad (5.184)$$

Substituting this expression and its adjoint into H and doing some simplifications gives us

$$H = \sum_{lm} (u\alpha u^\dagger)_{lm} b_l^\dagger b_m \quad (5.185)$$

Now notice that that α is Hermitian, we can choose u so that $u\alpha u^\dagger$ is diagonal, with entries λ_j being the eigenvalues of α , giving us

$$H = \sum_j \lambda_j b_j^\dagger b_j \quad (5.186)$$

This is the same form as H_{free} , an thus the ground state energy and excitation energies may be computed in the same way as we described earlier.

What about the ground state of H ? Assuming that all the λ_j are non-negative, it turns that that a state $|\psi\rangle$ satisfies $a_j^\dagger a_j |\psi\rangle = 0$ for all j **if and only if** $b_j^\dagger b_j |\psi\rangle = 0$ for all j . So the ground state for the two sets of Fermi operators is the same. This comes from the fact that $a_j^\dagger a_j |\psi\rangle = 0 \iff a_j |\psi\rangle = 0$: If $a_j^\dagger a_j |\psi\rangle = 0$ for all j , then $a_j |\psi\rangle = 0$ for all j , thus $b_j |\psi\rangle = 0$ since b_j is just a linear combination of a_j 's, from which it follows that $b_j^\dagger b_j |\psi\rangle = 0$.

So, the ground state for the two sets of Fermi operators a_j and b_j is in fact the same.

Now it turns out that we can generalize even further. The Hamiltonian $H = \sum_{jk} \alpha_{jk} a_j^\dagger a_k$ can be generalized to any Hamiltonian which is quadratic in Fermi operators, by which we mean it may contain terms of the form

$$a_j^\dagger a_k, \quad a_j a_K^\dagger, \quad a_j a_k, \quad a_j^\dagger a_k^\dagger. \quad (5.187)$$

Note that we will not allow linear terms of the form $a_j + a_j^\dagger$. Additive constants such as γI are easily incorporated, so we won't worry about those.

There are many ways to write such Hamiltonians, but the following form turns out to be especially convenient:

$$H = \sum_{jk} (\alpha_{jk} a_j^\dagger a_k - \alpha_{jk}^* a_j a_k^\dagger + \beta_{jk} a_j a_k - \beta_{jk}^* a_j^\dagger a_k^\dagger) \quad (5.188)$$

where this is written up to an additive constant γI , and with α Hermitian, and β antisymmetric.

The following is the procedure with which we will diagonalize this matrix:

- First, we define operators b_1, \dots, b_n :

$$b_j = \sum_k (\gamma_{jk} a_k + \mu_{jk} a_k^\dagger) \quad (5.189)$$

Once again, we want the b_j 's to satisfy the Fermionic CCRs, and when expressed in terms of the b_j , H has the same form as H_{free} so that it can be diagonalized.

- A calculation shows that

$$\{b_j, b_k^\dagger\} = \delta_{jk} I \iff \gamma\gamma^\dagger + \mu\mu^\dagger = I \quad (5.190)$$

while the condition

$$\{b_j, b_k\} = 0 \iff \gamma\mu^\top + \mu\gamma^\top = 0 \quad (5.191)$$

Upon writing out explicitly what the μ, γ matrices do, these two conditions are equivalent to the statement that the matrix

$$T = \begin{bmatrix} \gamma & \mu \\ \mu^* & \gamma^* \end{bmatrix} \quad (5.192)$$

is unitary. With this, we can check (via writing everything out explicitly) that

$$\begin{bmatrix} a_1 \\ \vdots \\ a_n \\ a_1^\dagger \\ \vdots \\ a_n^\dagger \end{bmatrix} = T^\dagger \begin{bmatrix} b_1 \\ \vdots \\ b_n \\ b_1^\dagger \\ \vdots \\ b_n^\dagger \end{bmatrix} \quad (5.193)$$

- The next step is to rewrite the Hamiltonian in terms of the b_j operators. To do this, observe that the expression for H can be written out explicitly as the following

$$H = [a_1^\dagger \dots a_n^\dagger \quad a_1 \dots a_n] \begin{bmatrix} \alpha & -\beta^* \\ \beta & -\alpha^* \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_n \\ a_1^\dagger \\ \vdots \\ a_n^\dagger \end{bmatrix} \quad (5.194)$$

From Eq. 5.193 we see that

$$[a_1^\dagger \dots a_n^\dagger \quad a_1 \dots a_n] = [b_1^\dagger \dots b_n^\dagger \quad b_1 \dots b_n] T. \quad (5.195)$$

This allows us to write

$$H = [b_1^\dagger \dots b_n^\dagger \quad b_1 \dots b_n] T M T^\dagger \begin{bmatrix} b_1 \\ \vdots \\ b_n \\ b_1^\dagger \\ \vdots \\ b_n^\dagger \end{bmatrix} \quad (5.196)$$

where

$$M = \begin{bmatrix} \alpha & -\beta^* \\ \beta & -\alpha^* \end{bmatrix}, \quad (5.197)$$

Of course, to save space, we can let

$$[b^\dagger \quad b] = [b_1^\dagger \dots b_n^\dagger \quad b_1 \dots b_n] \quad (5.198)$$

from which H can be expressed more succinctly as

$$H = [b^\dagger \quad b] T M T^\dagger \begin{bmatrix} b \\ b^\dagger \end{bmatrix} \quad (5.199)$$

- Suppose we can choose T such that TMT^\dagger is diagonal, we see that the Hamiltonian can be expressed in the form of H_{free} , and the energy spectrum found, following our earlier methods.
- Since α is Hermitian and β is anti-symmetric it follows that M is also Hermitian, and so can be diagonalized for some choice of a unitary T . However, the fact the the b_j 's must satisfy the CCRs constrains the class of T 's available to us. It remains to show that there always **exists** such a T that can be used to do the diagonalization. It turns out that there is a proof out there for this involving the CS-decomposition for unitary matrices. However, we won't go into that.

5.18.2 Jordan-Wigner transformation

In this section we describe the Jordan-Wigner transform, explaining how it can be used to map a system of qubits (spin-1/2 systems) to a system of Fermions, and vice versa. To illustrate the utility of the transformation, we will look at how we can solve a one-dimensional quantum spin system using the J-W transformation.

Let an n -qubit system be given with the usual state space C^{2^n} and Pauli operators X_j, Y_j, Z_j acting on qubit j . We are going to use these operators to define a set of a_j operators acting on C^{2^n} which satisfy the Fermionic CCRs.

For the sake of argument, suppose we have found such a set of operators. Then from our earlier discussion the action of the a_j operators in the occupation number representation

$$|\alpha\rangle = |\alpha_1, \dots, \alpha_n\rangle \quad (5.200)$$

must be as follows:

- Suppose $\alpha_j = 0$ then $a_j |\alpha\rangle = 0$.
- Suppose $\alpha_j = 1$. Let α' be that vector which results when the j th entry of α is changed to 0. Then

$$a_j |\alpha\rangle = -(-1)^{s_\alpha^j} |\alpha'\rangle \quad (5.201)$$

where once again

$$s_\alpha^j = \sum_{k=1}^{j-1} \alpha_k. \quad (5.202)$$

If we identify the occupation number state $|\alpha\rangle$ with the corresponding computational basis state $|\alpha\rangle$, then this suggests taking as our *definition*:

$$a_j \equiv - \left(\bigotimes_{k=1}^{j-1} Z_k \right) \otimes \sigma_j$$

(5.203)

where σ_j denotes the matrix $\sigma \equiv |0\rangle\langle 1|$ acting on the j th qubit. In some sense, this is the “lowering operator” on the single-qubit computational basis state.

It is easy to verify that these operators satisfy the Fermionic CCRs. This definition of the a_j is known as the **Jordan-Wigner transform**. It allows us to define a set of operators a_j satisfying the Fermionic CCR in terms of the usual operators we use to describe qubits, or spin-1/2 systems.

The J-W transformation can be inverted, allowing us to express the Pauli operators in terms of the Fermionic operators a_1, \dots, a_n . In particular, we have

$$Z_j = a_j a_j^\dagger - a_j^\dagger a_j \quad (5.204)$$

Indeed, it is easy to check that the expression above keeps $|0\rangle$ constant and turns $|1\rangle$ to $-|1\rangle$.

Next, observe that

$$X_j = \sigma_j + \sigma_j^\dagger = |0\rangle\langle 1| + |1\rangle\langle 0|. \quad (5.205)$$

So, it follows that

$$X_j = -(Z_1 \dots Z_{j-1})(a_j + a_j^\dagger). \quad (5.206)$$

By substituting in the expressions for Z_1, \dots, Z_{j-1} in terms of the Fermionic operators. Similarly, we have

$$Y_j = i(Z_1 \dots Z_{j-1})(a_j^\dagger - a_j), \quad (5.207)$$

which together with the expression for the Z_j operators, enables us to write Y_j solely in terms of the Fermionic operators.

Now, obviously the expressions for X_j, Y_j are rather inconvenient as they involve products of large numbers of Fermi operators. Remarkably, however, for certain simple *products* of Pauli operators it is possible to obtain simple expressions in terms of the Fermi operators. In fact, one can check with algebra that

$$Z_j = a_j a_j^\dagger - a_j^\dagger a_j \quad (5.208)$$

$$X_j X_{j+1} = (a_j^\dagger - a_j)(a_{j+1} + a_{j+1}^\dagger) \quad (5.209)$$

$$X_j Y_{j+1} = i(a_j^\dagger - a_j)(a_{j+1}^\dagger - a_{j+1}) \quad (5.210)$$

$$Y_j Y_{j+1} = -(a_j^\dagger + a_j)(a_{j+1}^\dagger - a_{j+1}) \quad (5.211)$$

$$Y_j X_{j+1} = i(a_j^\dagger + a_j)(a_{j+1}^\dagger + a_{j+1}) \quad (5.212)$$

Suppose now that we have an n -qubit Hamiltonian H that can be expressed as a sum over operators from the set $Z_j, X_jX_{j+1}, Y_jY_{j+1}, X_jY_{j+1}, Y_jX_{j+1}$. An example of such a Hamiltonian is the transverse field Ising model,

$$H = \alpha \sum_j Z_j + \beta \sum_j X_i X_{i+1} \quad (5.213)$$

which describes a system of magnetic spins with nearest neighbor couplings of strength β along their \hat{x} axes, and in an external magnetic field of strength α along the \hat{z} axis.

For any such Hamiltonian, we see that it is possible to re-express the Hamiltonian as a Fermi quadratic Hamiltonian. As we saw in an earlier section, determining the energy levels is then a simple matter of finding the eigenvalues of a $2n \times 2n$ matrix, which can be done very quickly. In particular, finding the ground state energy is simply a matter of finding the smallest eigenvalue of that matrix, which is often particularly easy. In the case of models like the transverse Ising model, it is even possible to do this diagonalization analytically, giving rise to exact expressions for the energy spectrum. Details can be found in the paper by Lieb, Schulz and Mattis mentioned earlier, or books such as the well-known book by Sachdev on quantum phase transitions.

5.18.3 Summary: The Jordan-Wigner Transformation

The following presentation of the Jordan-Wigner transformation is perhaps more standard than Nielsen's introduction, so I think it's a good idea to have it here.

Remember that we have the operators $\sigma_j^+, \sigma_j^-, \sigma_j^z$ acting on site j . We find that $\{\sigma_j^+, \sigma_j^-\} = 1$ as would be expected from fermionic creation and annihilation operators. So, we might be tempted to set

$$\sigma_j^+ = \frac{1}{2}(\sigma_j^x + i\sigma_j^y) = f_j^\dagger \quad (5.214)$$

$$\sigma_j^- = \frac{1}{2}(\sigma_j^x - i\sigma_j^y) = f_j \quad (5.215)$$

$$\sigma_j^z = 2f_j^\dagger f_j - 1. \quad (5.216)$$

Now, remember that the CCRs are the following:

$$\boxed{\{a_j, a_k^\dagger\} = \delta_{jk}I; \quad \{a_j, a_k\} = 0; \{a_j^\dagger, a_k^\dagger\} = 0} \quad (5.217)$$

Notice that we now have $\{f_j^\dagger, f_j\} = I$ as the desired same-site relation, but $[f_j^\dagger, f_k] = 0$ where $j \neq k$. And so spins on different sites commute unlike fermions which anti-commute. We must fix this problem to get the correct analogy.

A transformation which recovers the true fermion commutation relations from spin-operators was done in 1928 by Jordan and Wigner.

Take a chain of fermions, and define a new set of operators:

$$\boxed{a_j^\dagger = e^{+i\pi \sum_{k=1}^{j-1} f_k^\dagger f_k} \cdot f_k^\dagger} \quad (5.218)$$

$$\boxed{a_j = e^{-i\pi \sum_{k=1}^{j-1} f_k^\dagger f_k} \cdot f_k} \quad (5.219)$$

$$\boxed{a_j^\dagger a_j = f_j^\dagger f_j} \quad (5.220)$$

Recall that $f_j^\dagger f_j$ is the *number* operator and $f_j^\dagger f_j \in \{0, 1\}$, so the quantity

$$\sum_{k=1}^{j-1} f_k^\dagger f_k \quad (5.221)$$

is the occupation number in modes $k = 1, \dots, j-1$ of the field. The phase is equal to **ONE** if the occupied modes is **EVEN**, and **MINUS ONE** if the occupied modes is **ODD**.

This phase is often expressed, using the relation $\{f_k^\dagger, f_k\} = I$ and that $\sigma_k^z - 2f_k^\dagger f_k - 1$, as

$$\exp\left(\pm i\pi \sum_{k=1}^{j-1} f_k^\dagger f_k\right) = \prod_{k=1}^{j-1} e^{\pm i\pi f_k^\dagger f_k} = \prod_{k=1}^{j-1} (1 - 2f_k^\dagger f_k) = \prod_{k=1}^{j-1} (-\sigma_k^z) \quad (5.222)$$

where the second equality makes use of the fact that $f_k^\dagger f_k \in \{0, 1\}$.

With these, the transformed spin operators now have the appropriate fermionic commutation relations:

$$\boxed{\{a_j^\dagger, a_j\} = \delta_{i,j}, \quad \{a_j^\dagger, a_j\} = 0, \quad \{a_i, a_j\} = 0} \quad (5.223)$$

The inverse transformation is given by

$$\boxed{\sigma_j^+ = e^{-i\pi \sum_{k=1}^{j-1} a_k^\dagger a_k} \cdot a_j^\dagger} \quad (5.224)$$

$$\boxed{\sigma_j^- = e^{+i\pi \sum_{k=1}^{j-1} a_k^\dagger a_k} \cdot a_j} \quad (5.225)$$

$$\boxed{\sigma_j^z = 2a_j^\dagger a_j - 1} \quad (5.226)$$

Note that the definition of the fermionic operators is nonlocal with respect to the bosonic operators because we have to deal with an entire chain of operators to the left of the site the fermionic operators are defined with respect to. This is also true the other way around.

5.18.4 Jordan-Wigner transformation in QAOA

5.19 Quantum Critical State from GHZ?