

# Quantifying Complexity

Han Yan Yoong

December 26th, 2022



MSc in Theoretical Physics  
The University of Edinburgh  
2022

## **Abstract**

This dissertation aims to provide the reader a basic notion of quantum complexity from three different perspectives. There exists discernible connections at a fundamental level among these perspectives which will be explored in the subsequent chapters. The content and level of this paper is aimed towards readers with some undergraduate knowledge in quantum mechanics and mathematics. We will state some important ideas and theorems from a few selected papers by some key physicists in the field of quantum information and quantum complexity, and then expand upon the ideas discussed. We also provide proof of an important theorem in Chapter 4 and discuss the significance of it. In Chapter 5, we shall discuss the ideas and concepts of quantum complexity geometry.

## Declaration

I declare that this dissertation was composed entirely by myself.

Chapter 2 provides an introduction to the basics of quantum computation and quantum mechanics, which will be the main tools in understanding quantum complexity. These ideas and definitions are not original as they are referenced from standard mathematical and quantum information literature online.

Chapter 3 provides an introduction to the notion of complexity growth of  $K$ -qubit system in the Hilbert space from the first part of L.Susskind's paper[1]. In Section 3.3.3, I provide an independent analysis of the volume of  $SU(N)$  which is my own original work in the form Table 3.1, and compared it to the very rough estimation provided by L.Susskind in this paper[1] on page 10, which is quite ambiguous.

In Chapter 4, I state the Solovay Kitaev theorem and provide the proof for the theorem which is largely my own work. I used ideas and strategies from some papers in order to help strengthen my intuition in proving the theorem using sources which are all referenced in this paper.

Chapter 5 presents the works mainly by Michael Nielsen et al in these papers[2], [3] and [4]. The discussions of these papers are not of my own work, but I have consciously extracted the concepts and ideas I deemed important for the discussion of the topic of this dissertation.

Finally, the conclusion of the dissertation is composed entirely of my own thoughts.

All the diagrams in this dissertation are referenced from the relevant papers apart from Figure 2.1, Diagram 2.3.1, Diagram 2.3.2, Diagram 2.5.1 and Figure 4.1 are drawn entirely by myself using  $\LaTeX$  package.

## Personal Statement

The first two weeks of the project consisted of me mostly reading up and trying to understand the concepts and ideas from the papers recommended by my supervisor, Dr. Joan Simon. From his recommendations, the paper by L.Susskind[1] about complexity and blackholes piqued my interest and encouraged me to learn about his perspective on quantum complexity, especially in analysing the growth of complexity and qubits in Hilbert space. Whilst reading his paper, I found some ambiguity in his approximation of the volume of  $SU(N)$  and thereupon carried out my own analysis. My supervisor and I then discussed about my findings gave me some feedback on my analysis.

In the third week, my supervisor recommended that I read up on the Solovay-Kitaev theorem which is an important idea central to discrete gate complexity involving efficient unitary approximation. I found that this theorem is very relevant to the topic of my dissertation and decided to delve deeper into it by providing the proof for the theorem. However before embarking to prove the theorem, I had to read up on topics in Lie Group theory and some basics in manifolds. I relied heavily of the quantum computation and quantum information textbook by Nielsen and Chuang[5] to solidify my knowledge on universal quantum gates, matrix norm and approximating quantum circuits which served as a basis in proving the theorem.

In the fourth week, I became a father as my son was born. This was one of the happiest and most amazing events of my life but I was aware about the time constraint and pressure that this circumstance would impose upon me. I got in touch immediately with the university and Dr. Joan to inform them of this event and they were very supportive and understanding of my situation. I applied for a one week extension to submit my dissertation and was granted owing to this circumstance. After a few days of the birth of my son, I resumed work on the Solovay-Kitaev theorem and in the subsequent meeting with my supervisor we discussed the ideas behind the theorem and my intention to provide the proof for it.

In the sixth week meeting with my supervisor, he mentioned to me about looking at quantum complexity in terms of geometry. I found the idea interesting and was encouraged to read up on some papers recommended by my supervisor. Whilst doing so, I learned about the idea of metricizing complexity with Riemannian geometry and using geodesic to define complexity between unitary operators in the  $SU(N)$  manifold. Since I did not have any prior experience on manifolds and differential geometry, it took me a week to grasp the idea of complexity as geometry as outlined in the papers[2] and [3] by reading up on those subjects.

Shortly thereafter my wife, my newborn son and I unfortunately contracted COVID-19 on the last week of November and it rendered me unable to work on the thesis for more than a week. Consequently I had to cancel the weekly meeting with my supervisor and instead continue reading and making notes at home. In the meantime I started writing my thesis, typing out the preliminary ideas and background knowledge for each chapter. I had by then decided to write about three different notions of complexity in my dissertation, namely the ideas of L.Susskind, the Solovay-Kitaev theorem and complexity as geometry as conjectured by Nielsen et al.

As the deadline drew closer, I started to write Chapter 5, which is the last chapter of the

dissertation which discusses the idea of geometry and complexity. Chapter 5 for me was the most challenging chapters of all due to the fact that it requires a good understanding of topological space, manifold and Lie Group theory, which I had not previously studied. In the first section of Chapter 5, I present ideas of the geodesic equation of the Riemannian manifold in  $SU(2)$  using the Lax equation, a new equation which I learned from reading the papers. In the next part of the chapter, through the geometric picture of complexity I reviewed the ideas by Nielsen[3] that we can approximate a unitary  $U$  up to an order of polynomial number of one and two-qubit gates. In the last section, I wrote about the lower and upper bound of quantum complexity in the geometric picture and I particularly struggled with writing this section especially with Hessian and the introduction of Finsler metric. I had to spend a week reading the papers[4],[6] and did a lot of research online in order to write up this section.

I spent the final two weeks to write up the whole project and proofread it to the best of my knowledge. It has been a rather challenging and difficult journey in writing this dissertation. However despite all the challenges I have enjoyed learning the different perspectives and concepts of quantum complexity through writing this dissertation.

## Acknowledgements

First and foremost I would like to extend my gratitude and thanks to my supervisor Dr. Joan Simon Soler for being very patient and kind towards me, and constantly replying to my emails in a timely manner. Thank you very much for your invaluable advice on the project and for sharing your incredible knowledge and expertise with me in writing this dissertation.

I would like to thank all the wonderful and amazing people whom I met during my time at the university. Special shout out to Ng Yi Sheng and Euan MacKay for being so supportive and kind towards me. I am very glad to have met you guys and I truly appreciate the patience and help you guys have given me. Without this, I believe I would not have been able to progress thus far in my MSc. Also not forgetting Ben Karsberg for being very generous and kind in sharing his knowledge and making time for me. I would like to thank all of you from the bottom of my heart, who made my time in Edinburgh so much more memorable and wonderful.

Finally, none of this would ever be possible without the special support and love from my wife Vickie Yoong for being the rock in my life.

# Contents

List of Tables	vii
List of Figures	1
<b>1 Introduction</b>	<b>2</b>
<b>2 The Basics of Quantum Computation</b>	<b>4</b>
2.1 Postulates of Quantum Mechanics. . . . .	4
2.2 Qubits and the Space State. . . . .	5
2.2.1 The Qubits . . . . .	5
2.2.2 The Bloch Sphere. . . . .	7
2.3 Quantum Circuit . . . . .	8
2.4 Single Qubit Operations. . . . .	9
2.5 Universal Quantum Gates . . . . .	11
<b>3 Exploring the Qubits in the Hilbert Space.</b>	<b>14</b>
3.1 The Lie Group: Unitary $U(n)$ and Special Unitary $SU(n)$ Group. . . . .	14
3.1.1 Matrix Lie Groups . . . . .	15
3.1.2 Matrix Exponentiation. . . . .	16
3.1.3 Lie Algebra . . . . .	17
3.2 Complex Projective Hilbert Space $\mathbb{C}P^N$ . . . . .	18
3.3 Counting States and Unitary Operators. . . . .	20
3.3.1 The Size of the Hilbert Space . . . . .	20
3.3.2 The Volume of $\mathbb{C}P^N$ . . . . .	21
3.3.3 Counting Unitary Operators in $SU(2^K)$ . . . . .	21
3.4 Relative Complexity of Unitaries . . . . .	24
3.5 Complexity: A Graph Theory Perspective. . . . .	25
3.6 The Second Law of Quantum Complexity. . . . .	27

3.6.1	Hamiltonian Evolution. . . . .	29
<b>4</b>	<b>Efficient Unitary Approximation: The Solovay-Kitaev Theorem</b>	<b>31</b>
4.1	The Basic Idea . . . . .	31
4.1.1	Matrix Norm . . . . .	32
4.2	Distance measure in $SU(2)$ . . . . .	33
4.2.1	The $\epsilon$ -net . . . . .	33
4.2.2	The Lie Algebra of $SU(2)$ . . . . .	33
4.2.3	Distance relations of the $SU(2)$ . . . . .	35
4.3	The "Shrinking" Lemma. . . . .	37
4.4	The Solovay-Kitaev Theorem . . . . .	39
4.5	Generalization to $SU(d)$ . . . . .	41
<b>5</b>	<b>The Geometry of Quantum Complexity</b>	<b>43</b>
5.1	Preliminaries . . . . .	44
5.2	The Riemannian Metric of the Lie Group $SU(2^N)$ . . . . .	48
5.2.1	The Geodesic Equation . . . . .	49
5.3	The Geometry of Quantum Circuits . . . . .	50
5.4	Complexity of the upper and lower bounds . . . . .	55
<b>6</b>	<b>Discussion and Conclusion</b>	<b>59</b>
6.1	Overview and Summary of Dissertation. . . . .	59
6.2	Future Areas for Research. . . . .	60
	<b>References and Bibliography</b>	<b>61</b>



# List of Tables

3.1 : Superfactorial vs Modified Stirling's approximation. . . . .	23
--	----

# List of Figures

2.1	The Bloch Sphere . . . . .	8
3.1	: The relative complexity can be thought of in terms of a discrete curve from $I$ to $U$ . . . . .	24
3.2	: Diagrammatic notion of a circuit graph visualized by a decision tree in $SU(2^N)$ . This is a regular tree with degree 8 and depth 2. The number of branches or edges $d$ at a vertex is called the degree of that vertex. The black dots (endpoints) represents the unitary operators(gates) and is called leaves. . . . .	26
3.3	: This is a regular tree with degree 3. The graph ceases being a tree and doubles back. Since a group space is homogeneous, the structure must look the same from every vertex. The shortest loops (girth of the graph) is of order $4^K$ . This the red loop is too short and the loops must look like the blue loop. . . . .	28
3.4	: Evolution of complexity with time. The ragged red curve is the evolution for a specific instance of an ensemble. The smooth curve is the ensemble average. . . . .	29
4.1	: Diagrammatic notion of an $\epsilon$ -net . . . . .	33
5.1	: Diagrammatic illustration of tangent space $T_x M$ at point $x$ on $M$ where the $v$ is the tangent vector of $\gamma(t)$ the parametrised curve. [20] . . . . .	45
5.2	: Constructing a quantum circuit to approximate $U$ . . . . .	54

# Chapter 1

## Introduction

Computer science has evolved over the past several decades garnering significant interest and investment as a result. Currently, we are living in the formative years of quantum computing since its ideas were introduced in the 1980s by physicists. In the following years thereafter, physicists David Deutsch and Richard Josza demonstrated that quantum algorithm is exponentially faster than any possible deterministic classical algorithm[7]. In 1994, Peter Shor developed a quantum algorithm for finding the prime factors of an integer with the potential to decrypt RSA-encrypted communications[8]. However, such quantum algorithm will require quantum computers to run and harness its full potential. Quantum computers are hard to build due to controlling or removing quantum decoherence. Therefore progress in building practical quantum computers is slow and still a work in progress at this time.

Quantum computers operate using qubits, the most basic unit of information in quantum computing and quantum information, which is analogous to bits in classical computing. For these qubits to perform any meaningful task in quantum computation, they require some quantum gates which are used to perform some unitary transformations on these qubits in order to solve some computational problems. Hence, any quantum computation can be thought of as a unitary transformation on qubits (sort of the same way a classical computation can be thought of as an arbitrary function from inputs to outputs).

Quantum computation can be visualised using the quantum circuit model. In the most basic form, such a circuit consists of some input qubits, quantum gates and an output measurement. The number of qubits of a circuit is called the *width* and the number of gates is called the *size*. This very basic notion is used as a fundamental model for quantum computers. Since quantum gates are crucial in finding the solution of a problem through the input qubits, they will be the main subject of interest in this paper.

The quantum complexity of a unitary transformation or quantum state is defined as the size of the shortest quantum computation that executes the unitary or prepares the state. Quantum complexity is also defined as the inherent hardness or difficulty of a quantum computational problem. It can also be seen as costs or a measure of resources i.e energy, time and entanglement required to solve problems. Therefore, quantum complexity is imperative in the design and build of quantum computers since it allows us to quantify the viability, scale and resources needed in implementing quantum computations.

In recent years, the notion of quantum complexity has been extended to the theory of gravity and blackhole by Leonard Susskind, et al. He conjectured that the growth of the inside of the blackhole is proportional to the growth of the complexity of a quantum circuit of  $n$ -qubit[1]. Besides that, complexity theory has also found its application in the field of condensed matter and many-body physics, to name a few. Quantum complexity theory offers a viable tool for us to study and understand the growth and behaviour of blackholes beyond the event horizon, which has always been a very challenging subject.

The main objective of this paper is to discuss the basic ideas of quantum complexity in three different perspectives which are outlined in Chapter 3, 4 and 5. We shall review the different notions of quantum complexity with the aim of giving the reader a basic idea of the application of quantum complexity in different areas of physics. This paper is targeted to readers with some background knowledge in quantum mechanics, quantum information and some basic undergraduate knowledge in geometry and group theory. For the sake of completeness, we start off each chapter with the basic definition of key concepts and ideas which are required to fully grasp and appreciate each respective topic as we proceed with each chapter.

In Chapter 2, we begin by giving some key definitions on the important tools and terminology which are used to work with quantum information and complexity. Moving on to Chapter 3, we start off by discussing the idea of complexity growth of  $n$ -qubit quantum states is proportional to the number of unitary operators in the Lie group  $SU(2)$  from the paper by L.Susskind[1]. We then move on to discuss a few relevant topics from Susskind's paper namely, relative complexity of unitaries, graph theory perspective on complexity and the second law of complexity. Susskind argues that the complexity is equal to the entropy of an *auxiliary system*.

In Chapter 4 we discuss the notion of discrete gate complexity with the emphasis on the Solovay-Kitaev theorem. The Solovay Kitaev theorem states that for any unitary gate  $SU$  on a single qubit can be approximated to a precision of a given constant, which we will discuss in quite some detail and give a non-exhaustive proof of the theorem. The theorem essentially says that we can approximate any unitary operator from a finite set of universal gates, which directly translates to a finite *size* of a quantum circuit, which gives an idea of quantifying complexity.

In Chapter 5, we will delve briefly into the geometric notion of quantum complexity as geodesics in the Lie group  $SU(2^n)$  based on the ideas of Nielsen, et al. We explore the topological and geometrical aspects of quantum complexity in the manifold of  $SU(2^n)$  and discuss the notion of geodesics and the mathematical tools used in defining them. We also discuss briefly the meaning of the lower and upper bound of quantum gate complexity in its relation to optimal control cost and finding minimal distances on certain Riemannian and Finslerian manifolds[4]. The benefit of the analysis on this topic provides a unified and generalized framework for deriving connections between quantum gate complexity and optimal control.

Lastly, we present a summary of the chapters and make some comments on each chapter. We also mention the possibility of future research from the ramification of ideas discussed in this paper.

# Chapter 2

## The Basics of Quantum Computation

We begin with the review of and recap of some of the most important concepts of quantum physics before we proceed into the details of the chapters of this paper. Therein, we state three fundamental postulates of quantum mechanics[9] which are essential in understanding quantum computation. We also define briefly the qubit, quantum circuits, single quantum gates and universal quantum gates which will be key to understanding complexity.

Equipped with these notions, the reader should be able to appreciate and grasp the ideas outlined in this paper.

### 2.1 Postulates of Quantum Mechanics.

**Postulate 1: State space and States.** The state space of any closed quantum system is a Hilbert space  $\mathcal{H}$ , i.e. a complex vector space with an inner product. Quantum states are vectors  $|\psi\rangle$  in this Hilbert space, i.e.  $|\psi\rangle \in \mathcal{H}$ , with norm one, i.e.  $\langle\psi|\psi\rangle = 1$ . More precisely, quantum states correspond to equivalence classes of vectors differing by a multiplicative overall phase, i.e. we say  $|\psi\rangle$  is equivalent to  $|\psi'\rangle$  if  $|\psi\rangle \sim e^{i\alpha} |\psi'\rangle$ ,  $\forall \alpha \in \mathbb{R}$ . These classes are sometimes referred to as rays.

**Postulate 2: Time evolution.** The *time evolution* of a closed quantum system is described by a unitary operator  $U$ . Superficially given a quantum system in state  $|\psi(t)\rangle \in \mathcal{H}$  at time  $t$ , the state of the system at time  $t'$  is

$$|\psi(t')\rangle = U(t', t) |\psi(t)\rangle \tag{2.1.1}$$

Thus while Postulate 1 tells us what the state space of our system is, Postulate 2 tells us how an initial quantum state for the same quantum system evolves in time in the same state space.

**Postulate 3: Observables and Measurements.** Observables are *self-adjoint* (Hermitian)<sup>1</sup> operators. Standard measurements in quantum mechanics are *projective* measurements. These are characterised by a collection of projection operators  $\{P_n\}$ , i.e.  $P_n^2 = P_n \forall n \in \mathbb{Z}^+$ , satisfying  $\sum_n P_n = \mathbb{I}$ , where the positive integer  $n$  labels the different outcomes of measurement. Furthermore, if the quantum state of the system prior to the measurement is the vector  $|\psi\rangle$ , the probability that outcome  $n$  occurs is given by

$$p(n) = \langle \psi | P_n | \psi \rangle \quad (2.1.2)$$

Whenever the outcome  $n$  occurs, the state of the quantum state immediately *after* the measurement is

$$\frac{P_n |\psi\rangle}{\sqrt{\langle \psi | P_n | \psi \rangle}} \quad (2.1.3)$$

Unless stated otherwise, we will assume the Hilbert space  $\mathcal{H}$  is *finite* dimensional in this paper.

## 2.2 Qubits and the Space State.

The qubit is a simple representation of the classical bit analogue to quantum mechanics. This is one of the most important building blocks in quantum information and computation. The applications and properties of qubits will be the main subject of investigation and examination throughout this dissertation. In order to gain more meaningful understanding of qubits, we shall explore the geometry and the space where the qubits live in.

### 2.2.1 The Qubits

The qubit can be thought of as the "quantum" counterpart to the classical bit.

A bit is expressed in binary numbers of 0 and 1. In the most simple case of a classical bit, a bit can only be in the state of 0 or 1. However as for the quantum analogue, a qubit can be expressed in a superposition of both the binary state of 0 and 1.

A quantum state(s) is usually represented using the Dirac bracket. *Ad demonstrationem*, let  $\psi = (a_1, \dots, a_n) \in \mathbb{C}^n$  be a point in  $\mathbb{C}^n$ .

The ket  $|\psi\rangle$  represents the entries as a column vector,

---

<sup>1</sup>Let  $H$  be a Hermitian operator such that  $H = H^\dagger$ .

$$|\psi\rangle = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \quad (2.2.1)$$

The bra  $\langle\psi|$  is the adjoint of the ket and it represents the row vector,

$$\langle\psi| = (a_1^* \ a_2^* \ \dots \ a_n^*) \quad (2.2.2)$$

This state of a qubit  $|\psi\rangle$  in the Hilbert space  $\mathcal{H}$  can be written as a superposition of eigenstates such as:

$$|\psi\rangle = \alpha |0\rangle + \beta |1\rangle \quad (2.2.3)$$

Where  $\alpha$  and  $\beta$  are normalised complex constants such that,

$$|\alpha|^2 + |\beta|^2 = 1 \quad (2.2.4)$$

for  $\alpha, \beta \in \mathbb{C}$  be any normalized complex number.

**Definition 2.2.1:** A complex Hilbert space is a complex vector space  $\mathcal{H}$ , equipped with an inner product such that the norm turns  $\mathcal{H}$  into a complete metric space.

The quantum state of  $|0\rangle$  and  $|1\rangle$  are known as the computational basis state in the  $\hat{\mathbf{z}}$ -axis on a Bloch Sphere.

The state of a quantum system can be represented as the spin, energy, angular momentum and magnetic moment of an elementary particle. Without the loss of generality, the state  $|0\rangle$  and  $|1\rangle$  will denote spin up and spin down of the electron in this paper. .

**Definition 2.2.2:** a *pure state* of a  $n$ -quantum level system is a vector in  $|\psi\rangle$  in  $\mathbb{C}^n$  such that  $\langle\psi|\psi\rangle = 1$ . Furthermore a pure state is a state which cannot be written as a mixture of others states.

In quantum mechanics, it is not possible to distinguish physical states that are the same up to a global phase factor, such that the states  $|\psi\rangle$  and  $e^{i\theta} |\psi\rangle$  are indistinguishable. For this reason, the following equivalence relation holds,

$$\psi \sim \psi' \iff \psi = e^{i\theta} \psi' \quad (2.2.5)$$

for  $\theta \in \mathbb{R}$ .

This equivalence relation leads to the equivalence classes of complex lines through the origin in  $\mathbb{C}^n$ . These lines form the *complex projective space*,

$$\mathbb{C}P^{n-1} := \frac{\mathbb{C}^n - \{0\}}{z \sim \lambda z}, \quad \lambda \in \mathbb{C} \quad (2.2.6)$$

There is a bijective (one to one) correspondence between  $\mathbb{C}P^{n-1}$  and  $n$ -quantum level systems.

Therefore, we can formally define the qubit as,

**Definition 2.2.3:** A qubit is a two state or two level quantum system equivalent to a complex projective line  $\mathbb{C}P^{n-1}$ . A pure qubit state is a point in  $\mathbb{C}P^1$ .

### 2.2.2 The Bloch Sphere.

A qubit can be geometrically visualised in a Bloch Sphere.

The *projectivization*<sup>2</sup> of two-dimensional complex Hilbert space is the complex projective line  $\mathbb{C}P^1$  and is defined to be the Bloch sphere. The state space which describes one single qubit is simply the two-dimensional complex Hilbert space.

The basis state of a qubit can be visualised in a form of a Bloch Sphere in Figure 2.1

---

<sup>2</sup>Projectivization is a procedure which associates with a non-zero vector space  $V$  a projective space  $\mathbb{P}(V)$ , whose elements are one-dimensional subspaces of  $V$



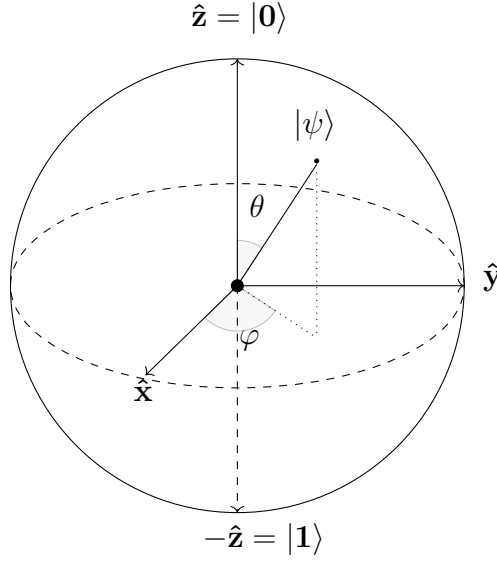


Figure 2.1 The Bloch Sphere

These qubit states in the z-axis can be expressed in column matrices as:

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ for spin up, and } |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ for spin down} \quad (2.2.7)$$

The state of a qubit is expressed in a two dimensional complex Hilbert space comprising of the eigenstate  $|0\rangle$  and  $|1\rangle$ .

In quantum mechanics, the inner product of a quantum state  $|\Psi\rangle$  must preserve the probability,

$$\langle\Psi|\Psi\rangle = 1 \quad (2.2.8)$$

Under this condition and representing  $|\Psi\rangle$  as a point on the surface of the Bloch sphere,

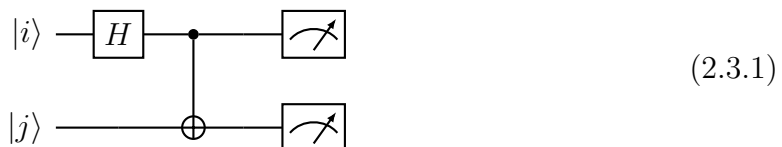
$$|\Psi\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\varphi}\sin\left(\frac{\theta}{2}\right)|1\rangle \quad (2.2.9)$$

where  $0 \leq \theta \leq \pi$  and  $0 \leq \varphi \leq 2\pi$ .

## 2.3 Quantum Circuit

A sequence of unitary gates, initialization of qubits and measurements constitute a simple model for quantum computation which defines a quantum circuit.

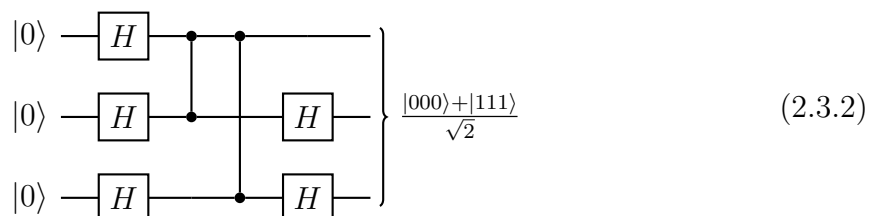
Diagram 2.3.1 depicts a two qubit quantum circuit initialised for state  $|i\rangle$  and  $|j\rangle$  which can be used to create entangled states. The first qubit  $|i\rangle$  also known here as the controlled qubit, is acted upon by a Hadamard gate and a CNOT gate (controlled-NOT) and a measurement. The second qubit  $|j\rangle$ , also known here as the target qubit is connected by a CNOT gate with the first qubit and a measurement.



A sequence of gates make a circuit and a quantum circuit consists of:

- Total number of qubits  $K$  is called the width of the circuit.
- The size of a quantum circuit is dictated by the number of gates (unitary operators).
- The number of steps in a circuit is called its depth. The path length is an integer number representing number of gates to execute in that path.

Diagram 2.3.2 shows an example of quantum circuit with width = 3, size = 7 and depth = 5.



This is a three qubit (width) circuit with seven gates (size) namely, five Hadamard gates and two Controlled- $X$  gates and a depth of five, such that the path of the qubits are through four unitary gates and one measurement operator going from left to right of the circuit.

## 2.4 Single Qubit Operations.

Having described the notions of the qubit and introduced some basic tools to work with it, we shall look at ways on how to operate on the qubit. Operations on a qubit must preserve the inner product with itself in order to maintain the probabilities of the states. The quantum operators acting on qubit are simply  $2 \times 2$  unitary matrices.

Recall that a pure state qubit is presented by a point  $z = (z_0, z_1) \in \mathbb{CP}^1$ , so a unitary transformation  $U \in U(2)$  will act as a matrix vector product  $U \cdot x$ . We can then provide a definition analogous to the classical logical gate in the quantum case.

**Definition 2.3.1:** A quantum gate for a pure state qubit is a unitary transformation  $U : \mathbb{CP}^1 \rightarrow \mathbb{CP}^1$ .

Some of the most important and common quantum gates are the Pauli matrices:

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \quad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (2.4.1)$$

Three other quantum gates which play an important role in quantum computation are the Hadamard gate  $H$ , phase gate  $S$  and the  $\pi/8$  gate  $T$ :

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}; \quad S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}; \quad T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix} \quad (2.4.2)$$

As we will see in the next section,  $H$  and  $T$  are a universal set of quantum gates or equivalently they generate a dense subset in  $U(2)$ .

Recall that a single qubit in the state  $a|0\rangle + b|1\rangle$  can be visualized as a point  $(\theta, \varphi)$  on a unit sphere where  $a = \cos(\theta/2)$ ,  $b = e^{i\varphi} \sin(\theta/2)$  and  $a$  can be taken to be real because the overall phase of the state is unobservable. This is called the Bloch sphere representation and the vector  $(\cos \varphi \sin \theta, \sin \varphi \sin \theta, \cos \theta)$  is called the Bloch vector.[5]

Since a qubit is a point on the Bloch sphere, a operation transforms these points to another, hence this transformation can be thought of a rotation over some axis. The Pauli matrices give rise to rotation operators in the  $\hat{x}, \hat{y}, \hat{z}$  axes and is defined by the following equations:

$$R_x(\theta) \equiv e^{-\theta X/2} = \cos \frac{\theta}{2} I - i \sin \frac{\theta}{2} X = \begin{bmatrix} \cos \frac{\theta}{2} & -i \sin \frac{\theta}{2} \\ -i \sin \frac{\theta}{2} & \cos \frac{\theta}{2} \end{bmatrix} \quad (2.4.3)$$

$$R_y(\theta) \equiv e^{-\theta Y/2} = \cos \frac{\theta}{2} I - i \sin \frac{\theta}{2} Y = \begin{bmatrix} \cos \frac{\theta}{2} & -\sin \frac{\theta}{2} \\ \sin \frac{\theta}{2} & \cos \frac{\theta}{2} \end{bmatrix} \quad (2.4.4)$$

$$R_z(\theta) \equiv e^{-\theta Z/2} = \cos \frac{\theta}{2} I - i \sin \frac{\theta}{2} Z = \begin{bmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{bmatrix} \quad (2.4.5)$$

An arbitrary unitary operator on a single qubit can be written in many ways as a combination of rotations, together with global phase shifts on the qubits. The following theorem provides a means of expressing an arbitrary single qubit rotation[5].

**Theorem 2.3.2: (Z-Y decomposition for a single qubit)** Suppose  $U$  is a unitary operation on a single qubit. Then there exist real number  $\alpha, \beta, \gamma$  and  $\delta$  such that,

$$U = e^{i\alpha} R_z(\beta) R_y(\gamma) R_z(\delta) \quad (2.4.6)$$

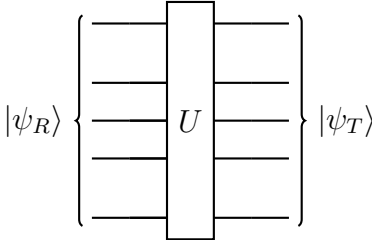
## 2.5 Universal Quantum Gates

In classical computing, algorithm(s) is used to solve a computational problem, which consist of an input, a sequence of instructions(logic gates) and an output(solution) to the problem.

In quantum computation, a similar model is implemented with the use of quantum gates. These quantum gates are unitary operators which perform transformation on the input qubits in order to perform quantum computation. These transformations can be thought of as a concatenation of a set of quantum gates  $g_n$  which prepares the unitary operator  $U$  which solves the computational problem.

To get an intuition, refer to diagram 2.5.1 as a basic example of universal in which a finite set of discrete universal gates can be used to approximate a target unitary operator  $U_T$ .

The reference state  $|\psi_R\rangle$  can be thought of as an input of a computational problem where  $|\psi_T\rangle$  the target state is the output or solution of the problem.



$$\left. |\psi_R\rangle \right\} \left[ \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{array} \right] U \left[ \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{array} \right] \left. |\psi_T\rangle \right\} \quad (2.5.1)$$

$$|\psi_T\rangle \approx U |\psi_R\rangle \quad (2.5.2)$$

$$\approx g_n g_{n-1} \dots g_1 |\psi_R\rangle \quad (2.5.3)$$

The target unitary  $U_T$  in this instance is an *algorithm* consisting of a concatenation of universal gates which is used to find the solution of the problem. In another words, the sequence of gates determines a unitary evolution  $U$  performed by a quantum computer.

In an ideal world, the set of gates  $g_n$  would be reduced to just a single exact quantum operation  $U_{exact}$ . However the preparation of  $U_{exact}$  is not possible in real world because we will require an infinite sequence of gates. Hence the closest to  $U_{exact}$  can only be achieved through *approximating* a desired quantum gate  $U$  from a finite set of *universal* gates to within an accuracy of  $\epsilon > 0$ , where epsilon is an arbitrarily small value.

Through the Solovay-Kitaev theorem, a meaningful and efficient approximation of these gate can be achieved, which will be discussed in more details in Chapter 4.

A set of quantum gates is said to be universal if any unitary transformation of the quantum data can be efficiently approximated arbitrarily well as a sequence of gates in the set. We assume that if  $g$  is in allowed set, so is  $g^\dagger$ . In another words, we say that a finite subset  $\mathcal{G} \subseteq SU(2)$  is a universal gate set if  $\langle \mathcal{G} \rangle$  is dense in  $SU(2)$  and  $\mathcal{G}$  is closed under inverses.

As outlined in the previous chapter, the notion of complexity can be thought of as a discrete set of unitary gates in a quantum circuit which prepares or constitute a unitary operator  $U$ .

From our definition of quantum gate, it can be generalised to systems to  $n$ -qubits as unitary matrices of dimension  $2^n \times 2^n$ . Any unitary operation on  $n$ -qubits can be decomposed as operations on a single qubit and an extra gate called the CNOT (controlled NOT) gate in  $U(4)$ . The CNOT gate acts on two qubits

The set of quantum operations will be finite when used in computation with a quantum computer. These unitary operations translate to physical operations in some system, so restrictions such as noise, etc are applied. Moreover these physical operations are difficult to control in a laboratory settings, so in principle we only have access to a small number of operations. This is why it is useful to find a set of *finite universal quantum gates*. We shall introduce some basic definitions for this concept.

**Definition 2.5.1:** Let  $\ell \leq 1$  be integers and  $\mathcal{G}$  a finite set of elements of  $SU(2)$ . A word of length  $\ell$  of elements in  $\mathcal{G}$  is given by  $w_\ell = g_1 g_2 \dots g_\ell$  with  $g_i \in \mathcal{G}$ . Let  $\mathcal{G}_\ell$  be the set of all words in  $\mathcal{G}$  of length at most  $\ell$ : and by  $\langle \mathcal{G} \rangle$  the set of all words in  $\mathcal{G}$  of finite length:

$$\mathcal{G}_\ell = \{w_k = g_1 g_2 \dots g_k : g_i \in \mathcal{G}, k \leq \ell\} \quad (2.5.4)$$

$$\langle \mathcal{G} \rangle = \bigcup_{\ell < \infty} \mathcal{G}_\ell \quad (2.5.5)$$

**Remark 2.5.2:** Note that in general, such a word  $w_\ell = g_1 g_2 \dots g_\ell \in \mathcal{G}_\ell$  is not necessarily an element of  $\mathcal{G}$

**Remark 2.5.3:** Since matrices in  $U(2)$  and  $SU(2)$  differ from a constant factor, we will give most of the definitions in  $SU(2)$ .

A set of quantum gates  $\mathcal{G}$  is defined as a set of universal quantum gates if it can be expressed as some word length  $w_\ell$  from  $\mathcal{G}_\ell$  for a unitary operation  $U$ . Note that  $SU(2)$  has infinite elements, and the number of finite sequences from a finite set is countable. To solve this problem we only require that any quantum operation can be approximated by a sequence of gates from this finite set. Furthermore, we can approximate unitaries of a constant qubit rather efficiently using the Solovay-Kitaev Theorem, which we will discuss in some detail in Chapter 4.

**Definition 2.5.3:** A set of quantum gates  $\mathcal{G}$  is called a set of universal quantum gates if  $\langle \mathcal{G} \rangle$  is dense in  $SU(2)$ , i.e if and only if for every element  $U \in SU(2)$  and for all  $\epsilon > 0$ . There exist  $g \in \langle \mathcal{G} \rangle$  such that  $d(g, U)^3 < \epsilon$ .

---

<sup>3</sup>The distance measure refers to the distance of matrices as defined in [4.1.1](#)

From this definition,  $d(g, U)$  can be understood as the error of implements  $g$  instead of  $U$ . More precisely,

**Definition 2.5.4:** Let  $U, V \in U(2)$ . The error when  $V$  is implemented instead of  $U$  is defined by:

$$E(U, V) := \max_{\|\psi\|=1} \|(U - V) |\psi\rangle\| \quad (2.5.6)$$

From section 4.1.1, this definition of error is equivalent of the distance between the operators  $U$  and  $V$  induced by the operator norm. Moreover the error can be defined using another distance like the trace norm. Thus, the concept of universal set of quantum gates can be viewed as a set that can implement an arbitrary unitary operation with an arbitrary non-zero error.

**Proposition 2.5.5:** Given two sequences  $V_1, V_2, \dots, V_m$  and  $U_1, U_2, \dots, U_m$  of quantum gates, the error of implementing the first sequence instead of the second satisfies:

$$E(U_m \dots U_2 U_1, V_m \dots V_2 V_1) \leq \sum_{i=1}^m E(U_i, V_i). \quad (2.5.7)$$

The proof of Proposition 2.5.5 can be found in [5] on page 195.

**Theorem 2.5.6:** The Hadamard gate  $H$  and the  $\pi/8$ -gate  $T$ , are a set of universal quantum gates for  $SU(2)$ . The proof for this theorem can be found in [5] on page 196.

This theorem states that the Hadamard and  $\pi/8$ -gates can be used to approximate any single qubit unitary operation to arbitrary accuracy. There are other universal gates such as the phase gate, controlled-NOT and Toffoli gate that are discussed in [5].

The theorem and propositions tell us that finding a set of universal gates is not a difficult. In the paper [10], it has been proven that almost any set of two qubit quantum gates is universal.

# Chapter 3

## Exploring the Qubits in the Hilbert Space.

In this chapter, we review the first part of the lecture by L.Susskind's paper[1], which consisted of a three part lecture on complexity and blackholes. The author has made a few interesting conjecture about the complexity growth of  $K$ -qubit system in the Hilbert space that scale proportionally with the number of unitary operators in the Lie group manifold  $SU(2^K)$ . We shall explore and discuss some of his idea on the growth of unitaries of a quantum system of  $K$ -qubits in  $SU(2^K)$ , the notion of relative complexity between two unitaries in the Lie manifold, basic introduction of complexity theory using graph theory and the second law of complexity which relates complexity to the entropy of an auxiliary system. For this paper, we shall not discuss the complexity theory of blackholes.

We begin this chapter with some basic definitions of general group theory and complex projective Hilbert space which will be essential in this chapter, before delving into the ideas from L.Susskind's paper.

### 3.1 The Lie Group: Unitary $U(n)$ and Special Unitary $SU(n)$ Group.

Lie Group is a group of differential manifold (5.1) which provides a natural model for the concept of continuous symmetry. This notion of continuous symmetry is extremely important in physics as nature exhibits aspects of symmetry in its most fundamental form. Symmetry is defined as invariance under a set of transformation and therefore, one defines a group as a as collection of transformations.

Of all the Lie Groups, the  $U(n)$  and  $SU(n)$  groups are the ones of our interest in quantum computing. This is because they represent the possible quantum logic gate operations in a quantum circuit with  $n$  qubit, thus the  $2^n$  basis states.

### 3.1.1 Matrix Lie Groups

Since we will be working with groups and in particular the Lie group  $SU(n)$  throughout the paper, we shall give a brief definition of a group with properties of symmetry transformation:

**Definition 3.1.1:** A group  $(G, \circ)$  is a set  $G$  together with a binary operation  $\circ$  defined on  $G$  that satisfies the following axioms,

- Closure: For all  $g_1, g_2 \in G, g_1 \circ g_2 \in G$
- Identity element: There exists an identity element  $e \in G$  such that for all  $g \in G, g \circ e = g = e \circ g$ .
- Inverse element: For each  $g \in G$ , there exists a inverse element  $g^{-1} \in G$  such that  $g \circ g^{-1} = e = g^{-1} \circ g$ .
- Associativity: For all  $g_1, g_2, g_3 \in G, g_1 \circ (g_2 \circ g_3) = (g_1 \circ g_2) \circ g_3$ .

The set of all transformations that leave a given object invariant is called a symmetry group.

**Definition 3.1.2:** A Lie group is a smooth manifold  $G$  which is also a group and such that the group product:

$$\mu : G \times G \rightarrow G \quad (3.1.1)$$

and the inverse map are smooth. These conditions can be obtained into the single requirement that the map  $G \times G \rightarrow G$  given by  $(x, y) \rightarrow x^{-1}y$  is smooth.

**Definition 3.1.3:** Let  $G$  and  $H$  be Lie groups. A Lie group homomorphism<sup>1</sup> from  $G$  to  $H$  is a group homomorphism  $\Phi : G \rightarrow H$  which is also a smooth map. Additionally, if  $\Phi$  is one to one and onto and the inverse map  $\Phi$  is smooth, then  $\Phi$  is called a Lie group isomorphism.

**Definition 3.1.4:** The unitary group  $U(n)$  is the group of  $n \times n$  complex matrix  $A$  such that:

$$\sum_{l=1}^k (A^*)_{jl} A_{lk} = \delta_{jk} \quad (3.1.2)$$

where  $\delta_{jk}$  denotes the Kronecker delta. From this definition, we have  $A^*A = AA^* = I$ .

Note that if  $A$  belongs to  $U(n)$  then  $\det(A) = e^{i\phi}$  with  $\phi \in \mathbb{R}$ .

**Definition 3.1.5:** The special unitary group  $SU(n)$  is a subgroup of  $U(n)$  consisting of unitary matrices with determinant one. It is a  $n \times n$  matrix with determinant 1. Its dimension as a real manifold is  $n^2 - 1$ .

---

<sup>1</sup>A homomorphism is a map between two algebraic structures of the same type (that is of the same name), that preserves the operations of the structures



### 3.1.2 Matrix Exponentiation.

The exponential of a matrix is important in the theory of Lie groups it connects the Lie groups with their corresponding Lie algebras. We shall state some basic properties of the exponential of a matrix.

**Definition 3.1.6:** The exponential of an  $n \times n$  matrix  $A$  is defined as,

$$e^A = \sum_{m=0}^{\infty} \frac{A^m}{m!} \quad (3.1.3)$$

As usual  $A^0 = I$  and  $A^m$  is the product of  $A$ ,  $m$  times with itself.

**Definition 3.1.7:** For  $A$  is a diagonal matrix, exponentiation can be performed by exponentiating the diagonal elements.

$$A = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_k \end{pmatrix} \quad (3.1.4)$$

$$e^A = \begin{pmatrix} e^{a_1} & 0 & \cdots & 0 \\ 0 & e^{a_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{a_k} \end{pmatrix} \quad (3.1.5)$$

**Definition 3.1.8:** The following are some helpful properties and identities for two arbitrary matrices  $A, B$  of the same dimension.

- $e^0 = I$
- $(e^A)^* = e^{A^*}$
- $e^A$  is invertible and  $(e^A)^{-1} = e^{-A}$
- If  $[A, B] = 0$ , then  $e^{A+B} = e^A e^B = e^B e^A$
- If  $C$  is invertible, then  $e^{CAC^{-1}} = C e^A C^{-1}$
- $\frac{d}{dt} e^{tA} = A e^{tA} = e^{tA} A$
- $\det(e^A) = e^{\text{tr} A}$
- $e^{iAx} = \cos(x)I + i \sin(x)A$

### 3.1.3 Lie Algebra

The Lie Algebra consists of all elements the *generator* of  $X$  that result in an element of a corresponding group  $G$ , such that  $e^X \in G$  with the Lie algebra's binary operation known as the Lie bracket  $[\cdot]$  between the Lie algebra elements.

The Lie algebra is a vector space  $\mathfrak{g}$  with the binary operation  $[\cdot]$ :  $\mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$  satisfying,

1. Bilinearity:  $[aX + bY, Z] = a[X, Z] + b[Y, Z]$  and  $[Z, aX + bY] = a[Z, X] + b[Z, Y]$ , for  $a, b$  arbitrary constant and  $\forall X, Y, Z \in \mathfrak{g}$
2. Anticommutativity:  $[X, Y] = -[Y, X] \forall X, Y \in \mathfrak{g}$
3. The Jacobi Identity:  $[X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] = 0 \forall X, Y, Z \in \mathfrak{g}$

**Example 3.1.9:** For an element of the Lie group  $g \in G$  close to the identity is given,

$$g(\epsilon) = I + \epsilon X \quad (3.1.6)$$

where  $\epsilon$  is a very small number and  $X$  is the generator.

Repeating such an infinitesimal transformation will result in a finite transformation, such as a rotation  $h(\theta)$ . We can write the idea of repeating small transformation many times,

$$h(\theta) = (I + \epsilon X)(I + \epsilon X)(I + \epsilon X) \dots = (I + \epsilon X)^k \quad (3.1.7)$$

where  $k$  is the number of times the small transformation is repeated. For  $\theta$  to represent a finite transformation,  $N$  is some big number so that the element  $g$  can be expressed as close as possible to the identity,

$$g(\theta) = I + \frac{\theta}{N} X \quad (3.1.8)$$

For the transformation to be the smallest possible we require  $N \rightarrow \infty$ . Hence the the rotation is simply repeating the infinitesimal transformation infinitely,

$$h(\theta) = \lim_{N \rightarrow \infty} (I + \frac{\theta}{N} X)^N = e^{\theta X} \quad (3.1.9)$$

where the limit is just exponential function.

We can calculate the generator  $X$  of a given transformation(group element)  $h(\theta)$  by,

$$\left. \frac{d}{d\theta} h(\theta) \right|_{\theta=0} = \left. \frac{d}{d\theta} e^{\theta X} \right|_{\theta=0} = X e^{\theta X} \Big|_{\theta=0} = X \quad (3.1.10)$$

The Pauli matrices form the basis the generator of the Lie Algebra and they are given,

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (3.1.11)$$

## 3.2 Complex Projective Hilbert Space $\mathbb{C}P^N$ .

The states of a physical system correspond to vectors in a Hilbert space  $\mathcal{H}$  and the Born's rule gives the probability for a system in state  $|\psi\rangle$  to be in the state  $|\phi\rangle$  by,

$$P(\psi, \phi) = \frac{|\langle\psi|\phi\rangle|^2}{|\langle\psi|\psi\rangle\langle\phi|\phi\rangle|} \quad (3.2.1)$$

For any  $c \in \mathbb{C} - \{0\}$ ,  $P(c\psi, \phi) = P(\psi, c\phi) = P(\psi, \phi)$ , therefore  $P(\psi, \psi) = P(\psi, c\psi) = 1$  holds, hence  $c|\psi\rangle$  is the same state as  $|\psi\rangle$ .

**Definition 3.2.1:** A ray is the set of all vectors describing the same state by this condition. It is the one-dimensional subspace spanned by any of all the states.

For  $|\psi\rangle$ , the associated ray  $R_\psi$  is the set,

$$R_\psi := \{|\phi\rangle \in \mathcal{H} | \exists c \in \mathbb{C} : |\phi\rangle = c|\psi\rangle\} \quad (3.2.2)$$

The quantum-mechanical states are associated with rays in the vector space and the projective space is the space of rays.

"*Belonging to a ray*" is an equivalence relation on the Hilbert space, and hence can be divided out in the sense that we simply say two vectors are the same object in the space of rays if they lie in the same ray - the rays are the equivalence classes. Formally, we set up the relation,

$$\psi \sim \phi \iff \psi \in R_\phi \quad (3.2.3)$$

and the space of rays of projective Hilbert space to be,

$$P(\mathcal{H}) := (\mathcal{H} - \{0\}) / \sim. \quad (3.2.4)$$

This notion can be extended and applied to the complex projective Hilbert space  $\mathbb{C}P^N$

The complex projective space is the projective space with respect to the field of complex numbers. We shall endeavour to define the complex projective Hilbert space  $\mathbb{C}P^N$  in the perspective of quantum mechanics. In doing so we shall not explore the more rigorous and detailed definition in the topological picture.

Mathematically, quantum mechanics is described in the finite or infinite dimension projective Hilbert space. The Bloch Sphere represents the two dimensional Hilbert space of a qubit and the space of all qubit states is called the complex projective line  $\mathbb{C}P^1$ .

**Definition 3.2.2a:** For  $N \in \mathbb{N}$  the complex  $N$ -dimensional complex projective space is the complex manifold (topological space) defined as the quotient,

$$\mathbb{C}P^N := (\mathbb{C}^{N+1} - \{0\}) / \sim \quad (3.2.5)$$

of the Cartesian product of  $(N+1)$ -copies of the complex plane with the origin removed, by the equivalence relation,

$$(z \sim w) \iff (z = \kappa \cdot w) \quad (3.2.6)$$

for some  $\kappa \in \mathbb{C} - \{0\}$ .

**Definition 3.2.2b** For a finite dimension Hilbert space, a complex projective space  $\mathbb{C}P^N$  is defined as the set of all lines that pass through the origin of the  $\mathbb{C}^{N+1}$  space, with an equivalence relation that every single line is regarded as a point in this multi-set. An equivalence relation  $z \sim \kappa z$  in the complex space  $\mathbb{C}^{N+1}$ ,

$$(z_1, z_2, \dots, z_{N+1}) \sim (\kappa z_1, \kappa z_2, \dots, \kappa z_{N+1}), \quad \forall \kappa \in \mathbb{C} - \{0\} \quad (3.2.7)$$

For  $z = (z_1, z_2, \dots, z_{N+1}) \in \mathbb{C}^{N+1}$  and  $(z_1, z_2, \dots, z_{N+1}) \neq (0, 0, \dots, 0)$

The equivalence relation can be expressed as,

$$[z] = [z_1, z_2, \dots, z_N] \quad (3.2.8)$$

where every point in  $[z]$  represents a point in  $\mathbb{C}P^N$ . The points in  $\mathbb{C}P^N$  are equivalent to point on the surface of a  $S^{2N+1}$ -sphere of  $\mathbb{C}^{N+1}$ -space, multiplied by a phase factor of  $Z^k \sim e^{i\alpha} Z^k$ , where  $\alpha, k \in \mathbb{R}$

Therefore,

$$\mathbb{C}P^N \cong S^{2N+1}/U(1) \cong S^{2N+1}/S^1 \cong \mathbb{C}^N \cup \mathbb{C}P^{N-1} \quad (3.2.9)$$

For the special case here with the dimension of the complex projective space equals to 1,  $\mathbb{C}P^1 \simeq S^2$ , s.t the Bloch Sphere is isomorphic the Riemann Sphere  $S^2$ .

In addition to that, the transitive group of  $\mathbb{C}P^N$  is  $SU(N+1)$ , that is to say for every  $g \in SU(N+1)$ ,  $p, q \in \mathbb{C}P^N$ , there is a map,

$$\xi_g : SU(N+1) \times \mathbb{C}P^N \rightarrow \mathbb{C}P^N \quad (3.2.10)$$

$$p \mapsto \xi_g(p) = q \quad (3.2.11)$$

The isotropic group of  $\mathbb{C}P^N$  is the  $U(N)$  group, hence,

---

<sup>2</sup>The 1-sphere is a manifold(topological space) which is commonly called a circle. It has a nontrivial fundamental Abelian Lie group structure  $U(1)$ ; the circle group. It is homeomorphic to the real projective line.

$$\mathbb{C}P^N \cong SU(N+1)/U(N) \quad (3.2.12)$$

### 3.3 Counting States and Unitary Operators.

On this section, we shall explore the notion of complexity as outlined by L.Susskind on the first chapter of his paper[1]. In this paper he discussed the relationship between the growth of complexity and the number of qubits of unitary operators.

#### 3.3.1 The Size of the Hilbert Space

Since the qubit is described in the Hilbert space, it is natural to consider the space of states of  $K$  qubits. The state  $|\psi\rangle$  of a quantum system in Hilbert space  $\mathcal{H}$  is represented as vectors in complex vector space,

$$|\psi\rangle = \sum_{i=1}^{2^K} \alpha_i |i\rangle, \quad \alpha_i \in \mathbb{C} \quad (3.3.1)$$

To gain some perspective on the size of the Hilbert space spanned by all vectors, we could attempt to quantify the total number of vectors in  $\mathcal{H}$ . However this would be infinite since there are an infinite number of vectors<sup>3</sup> spanning the Hilbert space.

We can regulate the infinities by restricting each  $\alpha_i$  to be any finite constant  $m$  values.

Not taking into the account of normalizing  $|\psi\rangle$  or the overall complex phases of the state, the total number of states is,

$$\#\text{states} = m^{2^K} = \exp 2^K \log m \quad (3.3.2)$$

We can show the relationship between  $K$  and the number of states by a simple example. For  $K = 4$  and  $m = 4$ , the number of states is  $\approx 4.3 \times 10^9$ .

We take the logarithm of the number of states to show small number,

$$\log(\#\text{states}) = 2^K \log m \approx 22 \quad (3.3.3)$$

---

<sup>3</sup>A state vector can be thought of as a point in the Bloch Sphere and there are infinitely many points in a unit sphere.

It can be observed that the growth of the states is exponential (strong) on  $K$  and it is logarithmic(weak) on the regulator parameter  $m$ .

### 3.3.2 The Volume of $\mathbb{CP}^N$

We would like to count the states in  $\mathbb{CP}(2^K - 1)$  or unitary operators in  $SU(2^K)$ . First we calculate the volume of  $\mathbb{CP}(N)$  as follows[11],

$$V[\mathbb{CP}(N)] = \frac{\pi^N}{N!} \approx \left(\frac{e\pi}{N}\right)^N \quad (3.3.4)$$

To count states, we need to regulate the geometry. One way to do this is to replace points by small balls of radius  $\epsilon$ .

The volume of  $\epsilon$ -radius ball in  $\mathbb{CP}(N)$  is given as,

$$V[B(2N)] = \frac{\pi^N}{N!} \epsilon^{2N} \quad (3.3.5)$$

Now we want to find the number of  $\epsilon$ -balls in  $\mathbb{CP}(N)$ . This can be easily achieved by dividing the volume of  $\mathbb{CP}(N)$  by the volume of  $\epsilon$ -ball which gives  $\epsilon^{-2N}$ .

We identify the number of states with the number of  $\epsilon$ -balls in  $\mathbb{CP}(2^K - 1)$  by,

$$\#states = \frac{1}{\epsilon}^{2(2^K - 1)} \quad (3.3.6)$$

or

$$\log(\#states) = 2(2^K - 1) \log\left(\frac{1}{\epsilon}\right) \quad (3.3.7)$$

Observe that the strong growth in the number of states with respect to  $K$  and is only logarithmically sensitive to the cutoff parameter  $\epsilon$ .

### 3.3.3 Counting Unitary Operators in $SU(2^K)$

Next we show the number of operators in  $SU(N)$  by dividing its volume by the volume of an  $\epsilon$ -ball.

The formula for the volume of  $SU(N)$  is given as[11],

$$\text{Vol}(SU(N)) = \sqrt{\frac{N}{2(N-1)} \frac{N-1}{2(N-2)} \cdots \frac{3}{2(2)}} \prod_{K=1}^{N-1} \frac{2\pi^{K+1}}{K!} \quad (3.3.8)$$

$$= \sqrt{\frac{N}{2^{(N-1)}}} \prod_{K=1}^{N-1} \frac{2\pi^{K+1}}{K!} \quad (3.3.9)$$

$$= \sqrt{N \cdot 2^{(N-1)}} \pi^{(N-1)(N+2)/2} \prod_{K=1}^{N-1} \frac{1}{K!} \quad (3.3.10)$$

The volume of  $\epsilon$ -radius ball of dimension  $N^2 - 1$  in  $SU(N)$  can be calculated using the volume for an  $n$ -sphere or hypersphere,

$$\text{vol}(\epsilon(n)) = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2} + 1)} \epsilon^n \quad (3.3.11)$$

$$= \frac{\pi^{(N^2-1)/2}}{\Gamma(\frac{N^2-1}{2} + 1)} \epsilon^{(N^2-1)} \quad (3.3.12)$$

$$= \frac{\pi^{(N^2-1)/2}}{((N^2-1)/2)!} \epsilon^{(N^2-1)} \quad (3.3.13)$$

We have substituted  $n = N^2 - 1$  in (3.3.12)

The Stirling formula is given as,

$$N! \sim \sqrt{2\pi N} \left(\frac{N}{e}\right)^N \quad (3.3.14)$$

Now dividing the volume of  $SU(N)$  by the volume of  $\epsilon$  ball,

$$\# \text{unitaries} = \sqrt{N \cdot 2^{(N-1)}} \pi^{(N-1)(N+2)/2} \prod_{K=1}^{N-1} \frac{1}{K!} \times \frac{((N^2-1)/2)!}{\pi^{(N^2-1)/2} \epsilon^{2(N^2-1)}} \quad (3.3.15)$$

$$= \frac{\sqrt{N \cdot 2^{(N-1)}} \pi^{(N-1)(N+2)/2}}{\pi^{((N^2-2)/2)} \sqrt{2\pi N}} \left(\frac{e}{N}\right)^{N^2/2} \left(\frac{N^2}{2e}\right)^{N^2/2} \frac{\sqrt{\pi} N}{\epsilon^{2(N^2-1)}} \quad (3.3.16)$$

$$\propto \left(\frac{N}{\epsilon^2}\right)^{N^2/2} \quad (3.3.17)$$

Some remarks on (3.3.13) and (3.3.14),

1. We have assumed  $N^2 - 1 \approx N^2$  for large  $N$
2. The prefactor is calculated to be  $2^{\frac{2N-3}{4}} \cdot N^{5/4} \pi^N$ , which is will be dominated by the exponent term, hence it has been omitted in (3.3.14).
3. The size of  $N$  is  $2^K$ , for  $K$ -qubit. Non-trivially,  $K \geq 1$

Taking the log of # unitaries,

$$\log(\#\text{unitaries}) = \frac{N^2}{2}(\log N + 2\log(\frac{1}{\epsilon})) \quad (3.3.18)$$

$$\approx \frac{K \cdot 4^K}{2} \log 2 + 4^K \log(\frac{1}{\epsilon}) \quad (3.3.19)$$

It is clear that the number of unitaries has a strong dependence on  $K$  and weak logarithmic dependence of the radius  $\epsilon$ .

We have noticed that the superfactorial term  $\prod_{K=1}^{N-1} \frac{1}{K!}$  equation 3.3.15 on this section as outlined by L.Susskind[1] page 10 gives a rather crude approximation. Therefore we will show the difference between the superfactorial which he used and our version of the modified Stirling's approximation to compare the results quantitatively between these two.

Ignoring the prefactor and constants, we analyse the superfactorial,

$$\prod_{K=1}^N K! \sim \prod_{K=1}^N \sqrt{2\pi K} \left(\frac{K}{e}\right)^K \approx \left(\sqrt{2\pi K} \left(\frac{K}{e}\right)^{K^2}\right)^{1/2} \quad (3.3.20)$$

For large  $K$ .

Comparing (3.3.15) between Superfactorial and Modified Stirling Approximation		
$K$	$\prod_{K=1}^N K!$	$\left(\sqrt{2\pi K} \left(\frac{K}{e}\right)^{K^2}\right)^{1/2}$
1	1	0.96027...
2	2	1.01923...
3	12	3.24747...
4	288	49.22480...
5	34560	4816.49926...
6	$2.48 \times 10^7$	$3.83 \times 10^6$
7	$1.25 \times 10^{11}$	$2.99 \times 10^{10}$
8	$5.06 \times 10^{15}$	$4.18 \times 10^{15}$
9	$1.83 \times 10^{21}$	$3.13 \times 10^{21}$
10	$6.66 \times 10^{27}$	$5.43 \times 10^{28}$
11	$2.66 \times 10^{35}$	$1.55 \times 10^{37}$
12	$1.27 \times 10^{44}$	$7.97 \times 10^{46}$
13	$7.93 \times 10^{53}$	$8.10 \times 10^{57}$

Table 3.1: Superfactorial vs Modified Stirling's approximation.

Note that as  $K$  grows larger, our modified version of Stirling Approximation will be greater than the superfactorial. Therefore, the modified Stirling Approximation is better suited as a crude estimate to show the dependency of the #unitaries with respect to  $N = 2^K$ .



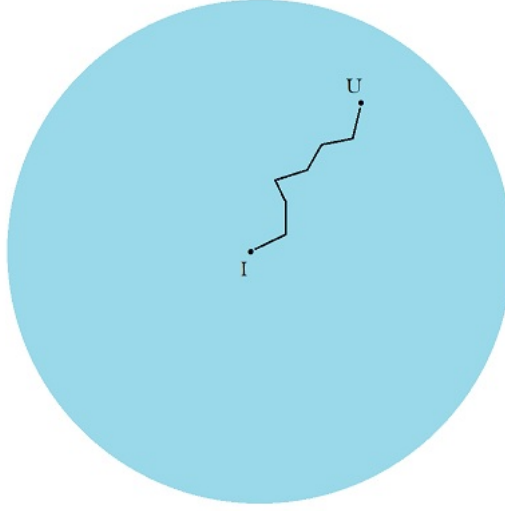


Figure 3.1: The relative complexity can be thought of in terms of a discrete curve from  $I$  to  $U$ .

### 3.4 Relative Complexity of Unitaries

Given two unitaries the relative complexity  $\mathcal{C}(U, V)$  is defined as the minimum number of gates<sup>4</sup> satisfying,

$$U = g_n g_{n-1} \dots g_1 V \quad (3.4.1)$$

to within tolerance  $\epsilon$ <sup>5</sup>. The relative complexity of  $U$  and the identity may be defined to be the complexity of  $U$ .

$$\mathcal{C}(I, U) \equiv \mathcal{C}(U) \quad (3.4.2)$$

The curve defining  $\mathcal{C}(I, U)$  is the shortest path i.e, the geodesic, but NOT the geodesic with respect to the inner product metric. It is a geodesic with respect to relative complexity. Figure 3.1 shows the path between the identity  $I$  and the unitary  $U$ <sup>6</sup>.

**Definition 3.2.4:** Relative complexity is a metric which satisfies:

<sup>4</sup>this set of gates is known as universal quantum gates.

<sup>5</sup>we shall explore this further in Chapter 4.

<sup>6</sup>A given circuit prepares a particular unitary operator  $U$ . Preparing  $U$  by a series of steps can be viewed as a discrete motion of a fictitious classical particle, called the auxiliary system  $\mathcal{A}$  in Brown-Susskind[12] The auxiliary system represents the motion of  $U(t)$  on  $SU(N)$  as time unfolds. The classical particle starts at the identity operator  $I$  and ends at  $U$ .

- $\mathcal{C} \geq 0$
- $\mathcal{C}(U, V) = 0$  iff  $U = V$
- $\mathcal{C}(U, V) = \mathcal{C}(V, U)$
- $\mathcal{C}(U, V) \leq \mathcal{C}(V, W) + \mathcal{C}(W, V)$  (Triangle Inequality)

This kind of metric is called right-invariant.

Suppose that,

$$U = (g_n g_{n-1} \dots g_1) V \quad (3.4.3)$$

Then for any  $W$  it follows that,

$$UW = (g_n g_{n-1} \dots g_1) VW \quad (3.4.4)$$

In other words the relative complexity of  $U$  and  $V$  is the same as that of  $UW$  and  $VW$ .

$$\mathcal{C}(U, V) = \mathcal{C}(UW, VW) \quad (3.4.5)$$

This is what it means for  $\mathcal{C}$  to be right-invariant.

As a counter-example, if we multiply  $W$  from left,

$$WU = (W g_n g_{n-1} \dots g_1 W^\dagger) WV \quad (3.4.6)$$

this would be called left-invariant. However,  $(W g_n g_{n-1} \dots g_1 W^\dagger)$  is generally not a product of  $n$  allowed gates. Therefore  $\mathcal{C}$  is not left-invariant.

From this definition, quantum complexity is a branch of right invariant geometry[3].

### 3.5 Complexity: A Graph Theory Perspective.

Quantum circuits can be described using graph theory[13] and a graph describing a circuit is called a *circuit graph*.

In this section, we will state some of the most important conjectured properties of circuit graphs. We have make the decision to omit most of the details of this section in the interest of not including too many pages and too much details in this dissertation. We highly encourage the reader to read up on this chapter on page 16 of the paper[1].

These are the properties of circuit graphs generated by iterating one-step circuits.

1. The number of possible choice  $d$  or the degree of the pairings of  $K$ -qubits in a circuit of a single step with  $K/2$  is given by,

$$d = \frac{K!}{K/2} \sim \left( \frac{2K}{e} \right)^{K/2} \quad (3.5.1)$$

Assuming *no-collisions*<sup>7</sup>, the number of unitaries that have been reached at depth  $D$  is,

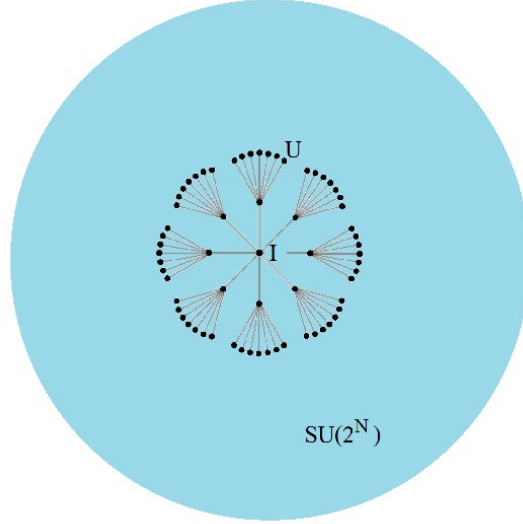


Figure 3.2: Diagrammatic notion of a circuit graph visualized by a decision tree in  $SU(2^N)$ . This is a regular tree with degree 8 and depth 2. The number of branches or edges  $d$  at a vertex is called the degree of that vertex. The black dots (endpoints) represents the unitary operators(gates) and is called leaves.

$$\#unitaries = d^D \approx \left( \frac{2K}{e} \right)^{DK/2} \quad (3.5.2)$$

The number of gates in a circuit of depth  $D$  is  $DK/2$ . Assuming no-collisions the path to each leaf is minimal path implying that the number of gates is the complexity.

We can then write,

$$\#unitaries = \left( \frac{2K}{e} \right)^{\mathcal{C}} \quad (3.5.3)$$

$$\sim e^{\mathcal{C} \log K} \quad (3.5.4)$$

This means that, *the sub-volume of  $SU(2^K)$  that corresponds to the unitaries of complexity  $\mathcal{C}$  grows exponentially with  $\mathcal{C}$ .*

2. The greatest distances between vertices (diameter) is  $\mathcal{C}_{max} \sim 4^K$ . The diameter is therefore logarithmic in the number of vertices.

---

<sup>7</sup>See [1] page 18 for further information

To show this, we argue that the number of  $\epsilon$ -regulated unitaries is finite and will eventually run out of room on  $SU(2^K)$ . This happens when the number of leaves as given in equation 3.3.29 is equal to the total number of  $\epsilon$ -balls as given in equation 3.3.16. This determines the maximum possible complexity.

$$\left(\frac{2^K}{e}\right)^{C_{max}} = \left(\frac{2^K}{e}\right)^{4^K/2} \quad (3.5.5)$$

$$C_{max} = 4^k \left[ \frac{1}{2} + \frac{|\log \epsilon|}{\log K} \right] \quad (3.5.6)$$

Again, strong dependence on  $K$  and weak dependence on  $\epsilon$ . Roughly,

$$C_{max} \sim 4^K \quad (3.5.7)$$

3. The number of vertices in the graph is of order  $e^{4^K}$

The point follows from the fact that the total number of unitaries is  $\sim e^{4^K}$ . We may identify this with the number of vertices in the graph. This implies that the diameter of the graph is logarithmic in the number of vertices.

4. Loops of length less than  $4^K$  are rare or absent.

When the graph reaches  $C_{max}$  it cannot continue to grow. Collisions occur and loops must form. The graph must double back on itself and revisit previously visited epsilon-balls. We show two possible loops that might form in

5. The graph is homogeneous and from any point looks tree-like out to distances of order the diameter.

## 3.6 The Second Law of Quantum Complexity.

In this section, we explore the similarity of quantum complexity and classical entropy, or as stated by Susskind and Brown, the existence of a second law of complexity. The reader is encouraged to read the paper[12] and section 9 on page 25 of [1]. Again, we present here a simplified version with omission of some details from the original paper in order to not exceed too many pages in this dissertation.

To show the relationship between quantum complexity and classical entropy, we consider a simplified version of the quantum evolution of a system at high temperature. We envision an ensemble of fictitious particles moving on  $SU(2^K)$ . We apply the concept from the circuit graph to simulate the discrete dynamics of the fictitious particles.

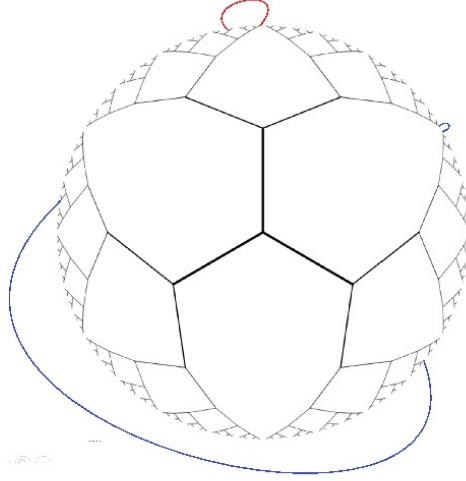


Figure 3.3: This is a regular tree with degree 3. The graph ceases being a tree and doubles back. Since a group space is homogeneous, the structure must look the same from every vertex. The shortest loops (girth of the graph) is of order  $4^K$ . This the red loop is too short and the loops must look like the blue loop.

After  $n$  steps the particle, with high probability, will be at the  $n$ th level, the auxiliary entropy will be  $S_{\mathcal{A}} = n \log d$ , and the complexity will be  $nK/2$ . Evidently the complexity and fictitious entropy of the auxiliary system are related and given by,

$$S_{\mathcal{A}} \approx \mathcal{C} \log K \quad (3.6.1)$$

We identify the entropy  $S_{\mathcal{A}}$  with the ensemble averaged complexity.

The second law of complexity is just the second law of thermodynamics—the overwhelming statistical likelihood that entropy will increase—applied to the ensemble average of complexity. The reason why complexity almost always increases when it is less than maximum is the same as why classical entropy almost always increases when it is less than maximum—the number of states exponentially increases with increasing entropy/complexity.

For a particular member of an ensemble, when  $\mathcal{C} \sim 4^K$ , the particle will have reached the maximum distance on the graph and the complexity will stop increasing.

Complexity equilibrium will have been achieved. The number of states with maximum complexity is so vast, that the particles will get lost among them and remain at maximum complexity for a recurrence time. The recurrence time for the classical system  $\mathcal{A}$  will be  $t_{recur} = e^{S_{\mathcal{A}}}$  which is doubly exponential in  $K$ ,

$$t_{recur} \sim e^{4^K} \quad (3.6.2)$$

Thus we expect a singly exponential time  $\sim 4^K$  during which complexity linearly grows, after which it remains approximately constant at its maximum. But then, on gigantically

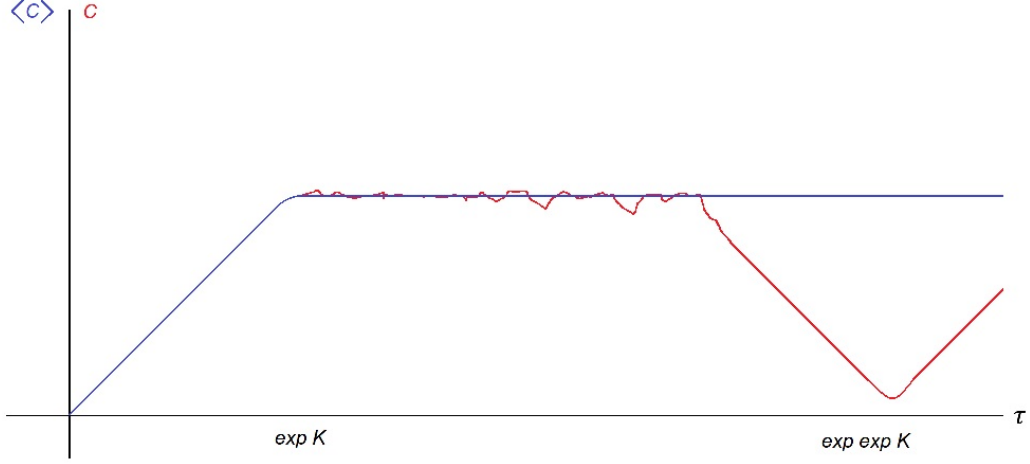


Figure 3.4: Evolution of complexity with time. The ragged red curve is the evolution for a specific instance of an ensemble. The smooth curve is the ensemble average.

long time scales  $\sim \exp \exp K$  it will recur to small values, and  $U(t)$  will return to the neighborhood of the identity. This is shown in figure 3.4.

### 3.6.1 Hamiltonian Evolution.

The constraint of energy conservation restricts the motion of  $U(t)$  to lie on a relatively low dimensional subspace of  $SU(2^K)$ . Let's write  $U(t)$  in the form,

$$U(t) = e^{iHt} \quad (3.6.3)$$

$$= \sum_{i=1}^{2^K} |E_i\rangle \langle E_i| e^{-iE_i t} \quad (3.6.4)$$

For a given Hamiltonian the motion is restricted to a torus defined by the set of  $2^K$  phases,

$$e^{i\theta_i} = e^{-iE_i t} \quad (3.6.5)$$

In other words,  $U(t)$  moves on a  $2^K$ -dimensional torus embedded in the  $(4^K - 1)$ -dimensional group  $SU(2^K)$ . Although in itself a very large space, the torus can only cover a tiny fraction of the full  $SU(2^K)$  space.

The question is whether  $U(t)$  fills the torus? Typically the answer is yes. If the energy levels are incommensurate, as will be the case if the system is chaotic, then the motion will be ergodic on the torus. The recurrence time for  $U$  to return to the neighborhood of the identity is the time for all the phases to get simultaneously close to 1, and that takes time,

$$t_{recur} \sim e^{2^K} \quad (3.6.6)$$

This is in contrast to the recurrence time  $\sim e^{4^K}$  random (Brownian) circuits which fill the entire  $4^K$ -dimensional unitary group. Therefore it is not possible for  $U(t)$  to visit any but an infinitesimal fraction of the maximally complex unitaries with  $\mathcal{C} \sim 4^K$ .

On the other hand the unitaries of complexity  $\leq 2^K$  comprise a set with volume comparable to the torus. It is therefore likely that for a Hamiltonian system the growth of complexity in Figure 3.4 reaches  $2^K$ .

In equation 3.3.29 we estimated the volume of the portion of  $SU(2^K)$  with complexity  $\mathcal{C}$  to be  $e^{\mathcal{C} \log K}$

That was the basis for the connection between complexity and the entropy of the auxiliary system, 3.3.33,

$$S_{\mathcal{A}} = \mathcal{C} \log K \tag{3.6.7}$$

A similar question can be formulated for motion restricted to the torus: What is the volume of the torus occupied by unitaries of complexity  $\mathcal{C}$ ? We expect that the answer is exponential,  $e^{\alpha \mathcal{C}}$ , but with  $\alpha$  parametrically smaller than  $\log K$ . We infer that  $\alpha$  is independent of  $K$ . In this case equation 3.3.33 would be replaced by,

$$S_{\mathcal{A}} \approx \mathcal{C} \tag{3.6.8}$$

# Chapter 4

## Efficient Unitary Approximation: The Solovay-Kitaev Theorem

In Chapter 2, we presented the concept of universal set of quantum gates in order to approximate any given unitary operation. However, we did not give a complete argument about how many gates we would need from our set to achieve the desired approximation.

The quantum complexity of a unitary transformation or quantum state is defined as the size of the circuit, i.e the number of universal that executes the unitary  $U$  or prepares the state. As mentioned previously, the unitary  $U$  is required to solve a given problem in a quantum computer. The difficulty of performing the computation is characterized by the number of gates used by the algorithm, which is said to be efficient if the number of gates required grows only polynomially with the size of problem. Therefore it is natural for us to be able to quantify or to say the least approximate the number of gates required from a finite set of universal gates in preparing  $U$ .

The Solovay-Kitaev Theorem demonstrates that it is possible to approximate any unitary operation to precision of  $\mathcal{O}(\log^c(1/\epsilon))$  gates from a given set, where  $c$  is a constant between 1 and 4 that depends on the implementation. The best value of  $c$  that we can achieve is 1[14], but achieving this value seems to be very difficult.

Note that the big  $\mathcal{O}$  notation is used in this chapter to classify functions according to their growth rates.

### 4.1 The Basic Idea

In the most basic idea, the Solovay-Kitaev Theorem shows that a set of single-qubit quantum universal gates generates a *dense* subset of  $SU(2)$  and the generating *finite* set fills  $SU(2)$  up sufficiently quick. This simply means that it is possible to obtain good approximation to any desired gate using a rather short sequence of gates from the generating set. This notion can be extended to qudits, which will be explored briefly in the last section of the chapter.

We shall attempt to present a non-exhaustive version of the proof of Solovay-Kitaev Theorem, using the ideas and notions as espoused by Nielsen and Chuang[5] and Ozols[15].



In order to harness the power of quantum computation, *fault-tolerant*<sup>1</sup> gates are essential in building the circuits and such gates are only limited to a few types such as the Clifford group and the  $\pi/8$  gates.

**Definition 4.1.1:** An instruction set  $\mathcal{G}$  for a  $d$ -dimensional qudit is a finite set of quantum gates satisfying:

- All gates  $g \in \mathcal{G}$  are in  $SU(d)$ , that is they are unitary and have determinant 1.
- For each  $g \in \mathcal{G}$  and its respective inverse operation  $g^\dagger$  is also in  $\mathcal{G}$ .
- $\mathcal{G}$  is a universal set for  $SU(d)$ , i.e the group generated by  $\mathcal{G}$  is dense in  $SU(d)$ . This means that given any quantum gate  $U \in SU(d)$  and any accuracy  $\epsilon > 0$  there exists a product  $S \equiv g_1 \dots g_m$  of gates from  $\mathcal{G}$  which is an  $\epsilon$ -approximation to  $U$ .

#### 4.1.1 Matrix Norm

The Solovay-Kitaev theorem allows us to calculate the distance between all elements of the group induced by the matrix norm of  $SU(2)$ . To prove the solovay-Kitaev theorem, we will use the trace norm.

**Definition 4.1.2:** Let  $A, B \in \mathcal{M}_n(\mathbb{C})$ , where  $\mathcal{M}_n(\mathbb{C})$  denotes the set of all  $n \times n$  matrices with complex entries. The trace norm is defined as,

$$\|A\| := \text{Tr}|A| = \text{Tr}\sqrt{A^\dagger A} \quad (4.1.1)$$

If  $A$  is normal s.t  $AA^\dagger = A^\dagger A$  and  $\lambda_1, \dots, \lambda_n$  are the eigenvalues of  $A$ , this norm can be computed as  $\|A\| = \sum_{i=1}^n |\lambda_i|$ .

The trace norm satisfies the following properties:

- Unitary invariance:  $\|UAV\| = \|A\|$  for any unitary operators  $U$  and  $V$ .
- Triangle inequality:  $\|A + B\| \leq \|A\| + \|B\|$ .
- Submultiplicativity:  $\|AB\| \leq \|A\| \cdot \|B\|$ .

The metric induced by the trace norm is given by  $d(A, B) := \|A - B\|$ .

**Definition 4.1.3:** A distance in  $\mathcal{M}_n(\mathbb{C})$ ,

$$d(\cdot, \cdot) : \mathcal{M}_n(\mathbb{C}) \rightarrow \mathbb{R} \quad (4.1.2)$$

is called unitary bi-invariant if and only if

---

<sup>1</sup>Quantum computer with a physical error rate below a certain threshold.

$$d(A, B) = d(UA, UB) = d(AU, BU) \quad (4.1.3)$$

for any unitary operator  $U$ .

**Remark 4.1.4:** Some texts use a different norm called the operator norm. This norm is defined as,

$$\|A\|_{op} = \sup_{v \neq 0} \frac{|Av|}{|v|} \quad (4.1.4)$$

Equivalently,  $\|A\|_{op}^2$  is the largest eigenvalue of the operator  $A^*A$ .

## 4.2 Distance measure in $SU(2)$ .

### 4.2.1 The $\epsilon$ -net

**Definition 4.2.1:** A subset  $S$  of  $SU(2)$  is said to be *dense* in  $SU(2)$  if for any element  $U$  of  $SU(2)$  and  $\epsilon > 0$  there is an element  $s \in S$  such that  $D(s, U) < \epsilon$ . Suppose  $S$  and  $W$  are subsets of  $SU(2)$ . Then  $S$  is said to form an  **$\epsilon$ -net** for  $w_n \in W$ , where  $w_n$  are a set of points in  $W$  where  $\epsilon > 0$ , if every point in  $W$  is within a distance  $d(s, W_n) \leq \epsilon$  of some point in  $S$ .

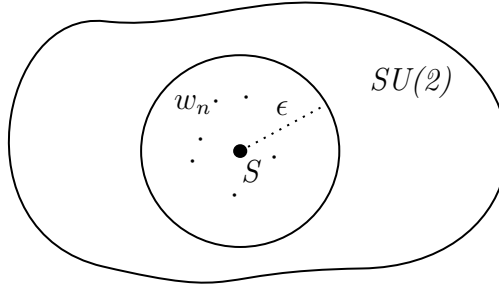


Figure 4.1: Diagrammatic notion of an  $\epsilon$ -net

**Remark 4.2:** Let  $S_\epsilon := \{U \in SU(2) : \|U - I\| < \epsilon\}$  be an closed  $\epsilon$ -ball in  $SU(2)$  centered around the identity.

### 4.2.2 The Lie Algebra of $SU(2)$

Let  $\mathfrak{su}(n)$  be the Lie Algebra of the Lie group  $SU(n)$  be the set of all  $n \times n$  traceless Hermitian matrices. The following are the properties of  $\mathfrak{su}(n)$ ,

1. If  $H \in \mathfrak{su}(n)$ , then  $e^{-iH} \in SU(n)$  satisfying  $\det(\exp^{-iH}) = \exp(-i\text{Tr}H) = 1$
2. If  $A, B \in \mathfrak{su}(n)$ , since  $(i[A, B])^\dagger = -i(BA - AB) = i[A, B]$  and  $\text{Tr}(i[A, B]) = i\text{Tr}(AB - BA) = 0$
3. The Pauli matrices are the infinitesimal generators of  $SU(2)$ .

**Definition 4.2.2:** The Lie Bracket for  $A, B \in \mathfrak{su}(n)$  is the binary operation of the Lie algebra  $:= [A, B] = AB - BA$ .

**Definition 4.2.3:** The group commutator for  $U, V \in SU(n)$  is given by  $[[U, V]] := UVU^\dagger V^\dagger$ .

**Proposition 4.2.4** Let  $\epsilon \geq 0$  and let  $A, B \in \mathfrak{su}(n)$  s.t  $\|A\|, \|B\| \leq \epsilon$ . As such  $\|e^{-[A, B]} - [[e^{-iA}, e^{-iB}]]\| \leq c\epsilon^3$ , for  $c$  is an arbitrary constant.

*Proof.* The Taylor expansion for the first term is,

$$e^{-[A, B]} = I - [A, B] + \frac{1}{2}[A, B]^2 - \dots \quad (4.2.1)$$

Next expand the second term,

$$[[e^{-iA}, e^{-iB}]] = e^{-iA}e^{-iB}e^{iA}e^{iB} \quad (4.2.2)$$

$$= I - [A, B] + \frac{i}{2}[A + B, A^2 + 2AB + B^2] + \mathcal{O}(\epsilon^4) + \dots \quad (4.2.3)$$

Note that all terms are equal up to second order, thus using the triangle inequality and the submultiplicativity of the norm the result follows.  $\square$

The Pauli matrices  $\sigma_i$  forms the basis of the generator for the Lie algebra  $\mathfrak{su}(n)$ . The inner product  $\vec{r} \cdot \vec{\sigma}$  of a linear combination of the Pauli matrices with coefficient  $\vec{r} \in \mathbb{R}^3$  is given by,

$$\vec{r} \cdot \vec{\sigma} = r_x \sigma_x + r_y \sigma_y + r_z \sigma_z \quad (4.2.4)$$

We will require some commutation relation and identities of Pauli matrices for the subsequent steps.

The following are some identities which will be used alongside with simple proof where required.

$$[\sigma_i, \sigma_j] + \{\sigma_i, \sigma_j\} = (\sigma_i \sigma_j - \sigma_j \sigma_i) + (\sigma_i \sigma_j + \sigma_j \sigma_i) \quad (4.2.5)$$

$$2\sigma_i \sigma_j = 2i\epsilon_{ijk}\sigma_k + 2\delta_{ij}I \quad (4.2.6)$$

$$\sigma_i \sigma_j = \delta_{ij}I + i\epsilon_{ijk}\sigma_k \quad (4.2.7)$$

**Lemma 4.2.4:** Let  $\vec{y}, \vec{z} \in \mathbb{R}^3$ , then  $[\vec{y} \cdot \vec{\sigma}, \vec{z} \cdot \vec{\sigma}] = 2i(\vec{y} \times \vec{z}) \cdot \vec{\sigma}$ .

*Proof.*

$$(\vec{y} \cdot \vec{\sigma})(\vec{z} \cdot \vec{\sigma}) = y_i \sigma_i z_j \sigma_j \quad (4.2.8)$$

$$= y_i z_j \sigma_i \sigma_j \quad (4.2.9)$$

$$= y_i z_j (\delta_{ij} I + i \epsilon_{ijk} \sigma_k) \quad (4.2.10)$$

$$= y_i z_j \delta_{ij} I + i \epsilon_{ijk} y_i z_j \sigma_k \quad (4.2.11)$$

$$= \vec{y} \cdot \vec{z} \cdot I + i(\vec{y} \times \vec{z}) \cdot \sigma \quad (4.2.12)$$

Using this identity,

$$[\vec{y} \cdot \vec{\sigma}, \vec{z} \cdot \vec{\sigma}] = (\vec{y} \cdot \vec{\sigma})(\vec{z} \cdot \vec{\sigma}) - (\vec{z} \cdot \vec{\sigma})(\vec{y} \cdot \vec{\sigma}) \quad (4.2.13)$$

$$= i(\vec{y} \times \vec{z}) \cdot \sigma - i(\vec{z} \times \vec{y}) \cdot \sigma \quad (4.2.14)$$

$$= 2i(\vec{y} \times \vec{z}) \cdot \sigma \quad (4.2.15)$$

for  $i \neq j$ . □

**Axiom 4.2.6:** Let  $u : \mathbb{R}^3 \rightarrow SU(2)$  be the map  $u(\vec{r}) := \exp(-\frac{i}{2}\vec{r} \cdot \vec{\sigma})$ , where  $\vec{r} \in \mathbb{R}^3$ . This provides a correspondence between the Lie algebra  $\mathfrak{su}(2)$  and the Lie group  $SU(2)$

### 4.2.3 Distance relations of the $SU(2)$

On this section, we show the notion of distance between the elements of  $SU(2)$  corresponding to vectors in  $\mathbb{R}^3$ ,

**Proposition 4.3.1:** Let  $\vec{r} \in \mathbb{R}^3$ , then  $\|u(\vec{r}) - I\| = 4 \sin(\frac{\|\vec{r}\|}{4}) = \|\vec{r}\| + \mathcal{O}(\|\vec{r}\|^3)$

*Proof.* Compute the eigenvalues of of,

$$\vec{r} \cdot \vec{\sigma} = \begin{pmatrix} r_z & r_x - ir_y \\ r_x + ir_y & -r_z \end{pmatrix} \quad (4.2.16)$$

$$\det(\vec{r} \cdot \vec{\sigma} - \lambda I) = \begin{vmatrix} r_z - \lambda & r_x - ir_y \\ r_x + ir_y & -r_z - \lambda \end{vmatrix} \quad (4.2.17)$$

$$= -(r_z^2 - \lambda^2) - (r_x^2 - i^2 r_y^2) \quad (4.2.18)$$

$$= \lambda^2 - r_x^2 - r_y^2 - r_z^2 = 0 \quad (4.2.19)$$

$$\lambda^2 = \|\vec{r}\|^2 \quad (4.2.20)$$

Hence the eigenvalues are  $\lambda = \pm \|\vec{r}\|$

It follows that the eigenvalues of  $\|e^{-\frac{i}{2}\vec{r} \cdot \vec{\sigma}} - I\|$  are  $|e^{\pm \frac{i}{2}\|\vec{r}\|} - I|$

Computing with the eigenvalues,

$$|e^{\pm \frac{i}{2} \|\vec{r}\|} - I| = \sqrt{(e^{\frac{i}{2} \|\vec{r}\|} - I)(e^{-\frac{i}{2} \|\vec{r}\|} - I)} \quad (4.2.21)$$

$$= \sqrt{1 - e^{\frac{i}{2} \|\vec{r}\|} - e^{-\frac{i}{2} \|\vec{r}\|} + 1} \quad (4.2.22)$$

$$= \sqrt{2 - 2 \cos(\|\frac{\vec{r}}{2}\|)} \quad (4.2.23)$$

$$= \sqrt{4 \left( \frac{1 - \cos(\|\vec{r}/2\|)}{2} \right)} \quad (4.2.24)$$

$$= 2 \sin\left(\frac{\|\vec{r}\|}{4}\right) \quad (4.2.25)$$

Hence,  $\|u(\vec{r}) - I\| = |e^{+\frac{i}{2} \|\vec{r}\|} - 1| + |e^{-\frac{i}{2} \|\vec{r}\|} - 1| = 4 \sin(\frac{\|\vec{r}\|}{4})$ . Next, write down the power series expansion for the sine term,

$$4 \sin\left(\frac{\|\vec{r}\|}{4}\right) = 4 \left( \frac{\|\vec{r}\|}{4} - \frac{(\|\vec{r}\|/4)^3}{3!} + \frac{(\|\vec{r}\|/4)^5}{5!} - \dots \right) \quad (4.2.26)$$

$$= \|\vec{r}\| + \mathcal{O}(\|\vec{r}\|^3) \quad (4.2.27)$$

□

**Remark 4.3.2:** Given an  $\epsilon \geq 0$ , we will denote by  $S_\epsilon$  the closed ball of radius  $\epsilon$  and center the identity matrix in  $SU(d)$  :

$$S_\epsilon := \{s \in SU(2); d(s, I) < \epsilon\} \quad (4.2.28)$$

The following proposition relates the distance of an element of  $SU(2)$  to the identity and the length of a vector in  $\mathbb{R}^3$ .

**Proposition 4.3.3:** Let  $\vec{r} \in \mathbb{R}^3$ . If  $u(\vec{r}) \in S_\epsilon$  then  $|\vec{r}| < \epsilon + \mathcal{O}(\epsilon^3)$ .

*Proof.* Using Remark 4.2 and Proposition 4.3.1, we have:  $\|u(\vec{r}) - I\| = 4 \sin \frac{|\vec{r}|}{4} < \epsilon$ . Thus  $|\vec{r}| < 4 \arcsin \frac{\epsilon}{4}$

Showing the Taylor expansion of  $\arcsin x = x + \frac{1}{2} \cdot \frac{x^3}{3} + \frac{1.3}{2.4} \cdot \frac{x^5}{5} + \dots$  □

**Proposition 4.3.4:** If  $\vec{y}, \vec{z} \in \mathbb{R}^3$  and  $|\vec{y}|, |\vec{z}| < \epsilon$  then  $\|u(\vec{y}) - u(\vec{z})\| = \|\vec{y} - \vec{z}\| + \mathcal{O}(\epsilon^3)$ .

*Proof.* By using unitary invariance<sup>2</sup>, Proposition 4.3.1 and triangle inequality,

$$\|u(\vec{y}) - u(\vec{z})\| = \|u(\vec{y})u(\vec{z})^\dagger - I\| \quad (4.2.29)$$

$$= \|u(\vec{y})u(\vec{z})^\dagger - u(\vec{y} - \vec{z})\| + \|u(\vec{y} - \vec{z}) - I\| \quad (4.2.30)$$

---

<sup>2</sup>According to Definition 4.1.1, for  $U, V \in SU(2)$  s.t  $\|U - V\| \Rightarrow \|UV^\dagger V - V\| = \|UV^\dagger - I\|$

Invoking Axiom 4.2.6 , the first term on 4.2.31 vanishes and with the second term, we get  $\|u(\vec{y} - \vec{z}) - I\| = 4 \sin\left(\frac{\|\vec{y} - \vec{z}\|}{4}\right)$ . Using the Taylor expansion for  $\sin x = x - \frac{x^3}{3!} + \dots$ , thus completing the proof.  $\square$

### 4.3 The "Shrinking" Lemma.

In this section, we will write down the an important lemma which is required for the proof of the Solovay Kitaev Theorem with reference from the paper by Zarapico. [16]

**Lemma 4.3** ("Shrinking" Lemma): Let  $\mathcal{G}$  be a finite set of elements in  $SU(2)$  containing its own inverses such that  $\langle G \rangle$  is dense in  $SU(2)$ . There exist constants  $\epsilon_0$  independent of  $\mathcal{G}$  such that  $\epsilon \leq \epsilon_0$  if  $\mathcal{G}_l$  is a  $(\epsilon, \epsilon^2)$ -net i.e  $\mathcal{G}_l$  is an  $\epsilon^2$ -net for  $S_\epsilon$ , then  $G_{5l}$  is a  $(\sqrt{C}\epsilon^{3/2}, C\epsilon^3)$ -net for some constant  $C$ .

We shall prove Lemma 4.3 shortly, but let us see how it implies the Solovay-Kitaev theorem. There are two steps to the proof. The first step is to apply Lemma 4.3 iteratively to show that the neighbourhood of the origin fills in very quickly as the word length  $l$  is increased. Since  $\mathcal{G}$  is dense in  $SU(2)$  we can find an  $l_0$  such that  $G_{l_0}$  is an  $\epsilon_0^2$ -net for  $SU(2)$  and thus also for  $S_{\epsilon_0}$ . Applying Lemma 4.3 with  $\epsilon = \epsilon_0$  and  $l = l_0$  implies that  $G_{5l_0}$  is a  $C(\sqrt{C}\epsilon_0^{3/2})^3$ -net for  $S_{\sqrt{C}(\sqrt{C}\epsilon_0^{3/2})^{3/2}}$ . Iterating this procedure  $k$  times, we find that  $\mathcal{G}_{5^k l_0}$  is an  $\epsilon(k)^2$ -net for  $S_{\epsilon(k)}$ , where,

$$\epsilon(k) = \frac{(C\epsilon_0)^{(3/2)^k}}{C} \quad (4.3.1)$$

The main idea in the proof of the Lemma is taking group commutators of elements in  $S_\epsilon$  and proving that these commutators fills  $S_{\epsilon^2}$  much more densely. The proof is shown using the unitary invariance of the distance which allows us to apply a translation step in order to get good approximations for any element of  $S_{\sqrt{C}\epsilon^{3/2}}$ .

*Proof.* Assume that  $\mathcal{G}_l$  is a  $(\epsilon, \epsilon^2)$ -net for  $\epsilon > 0$ . We first prove that there is a constant  $C$  such that  $\mathcal{G}_{4l}$  is a  $(\epsilon^2, C\epsilon^3)$ -net. Let  $U \in S_{\epsilon^2}$  and by Axiom 4.2.5 and Remark 4.3.3, for  $x \in \mathbb{R}^3$  s.t,

$$U = u(x) = e^{ix \cdot \sigma} \quad (4.3.2)$$

By applying Proposition 4.3.4, we get  $\|u(x) - I\| < \epsilon^2$  and expanding  $\sin(x)$  with the Taylor expansion,

$$|x| < \epsilon^2 + \mathcal{O}(\epsilon^6) \quad (4.3.3)$$

Now choose  $y, z \in \mathbb{R}^3$  such that  $x = y \times z$  and  $|y|, |z| < \epsilon$ . Recall that by Proposition 4.3.1 and Remark 4.3.3 one can verify that  $u(y), u(z) \in S_\epsilon$ .

Since  $\mathcal{G}_l$  is a  $(\epsilon, \epsilon^2)$ -net, there exists  $Y, Z \in SU(2)$  such that  $Y, Z \in \mathcal{G}_l \cap S_\epsilon$ ,

$$\|Y - u(y)\| \leq \epsilon^2, \quad \|Z - u(z)\| \leq \epsilon^2 \quad (4.3.4)$$

From Proposition 4.3.4 and Proposition 4.3.5, the following inequalities can be written,

$$\|u(y_0) - I\| < \epsilon \quad (4.3.5)$$

$$\|u(z_0) - I\| < \epsilon \quad (4.3.6)$$

$$|y_0|, |z_0| < \epsilon + \mathcal{O}(\epsilon^3) \quad (4.3.7)$$

$$|y_0 - y| < \epsilon^2, |z_0 - z| < \epsilon^2 \quad (4.3.8)$$

We can now prove that,

$$\|U - [[u(y_0), u(z_0)]]\| < C\epsilon^3 \quad (4.3.9)$$

The triangle inequality and  $U = u(y \times z)$  are used here to show,

$$\|U - [[u(y_0), u(z_0)]]\| \leq \|u(y \times z) - u(y_0 \times z_0)\| + \|u(y_0 \times z_0) - [[u(y_0), u(z_0)]]\| \quad (4.3.10)$$

For the first term, following Proposition 4.3.5,

$$\|u(y \times z) - u(y_0 \times z_0)\| = |y \times z - y_0 \times z_0| + \mathcal{O}(\epsilon^6) \quad (4.3.11)$$

$$= |(y - y_0) \times (z - z_0) + y_0 \times (z - z_0) + (y - y_0) \times z_0| + \mathcal{O}(\epsilon^6) \quad (4.3.12)$$

$$\leq |y - y_0||z - z_0| + |y_0||z - z_0| + |y - y_0||z_0| + \mathcal{O}(\epsilon^6) \quad (4.3.13)$$

$$\leq 2\epsilon^3 + \mathcal{O}(\epsilon^4) + \dots \quad (4.3.14)$$

For the second term we use Proposition 4.2.3 and Lemma 4.2.4,

$$\|u(y_0 \times z_0) - [[u(y_0), u(z_0)]]\| \leq \quad (4.3.15)$$

$$\leq \|\exp(\frac{1}{2}y_0 \cdot \sigma, \frac{1}{2}z_0 \cdot \sigma) - [[\exp(\frac{1}{2}y_0 \cdot \sigma), \exp(\frac{1}{2}z_0 \cdot \sigma)]]\| \leq c\epsilon^3 \quad (4.3.16)$$

Since  $u(y_0) = Y \in \mathcal{G}_l$  and  $u(z_0) = Z \in \mathcal{G}_l$ , we conclude that any  $U \in S_{\epsilon^2}$  can be approximated with a sequence of  $4l$  elements of  $\mathcal{G}$ . This proves that  $G_{4l}$  is  $(\epsilon^2, C\epsilon^3)$ -net.

Finally, we prove that there is a constant  $\epsilon'$  such that for any  $\epsilon \leq \epsilon'$ ,  $\mathcal{G}_{5l}$  is a  $(\sqrt{C}\epsilon^{\frac{3}{2}}, C\epsilon^3)$ -net.

Now let  $U \in S_{\sqrt{C}\epsilon^3}$ . The condition that defines  $\epsilon'$  is that  $S_{\sqrt{C}\epsilon^3} \subseteq S_{\epsilon'}$  or equivalently  $C\epsilon' < 1$ . Since  $G_l$  is a  $(\epsilon, \epsilon^2)$ -net, it follows that for any  $\epsilon \leq \epsilon'$  we can find  $V \in G_l$  such that,

$$\|U - V\| = \|UV^\dagger - I\| < \epsilon^2 \quad (4.3.17)$$

This result shows that  $UV^\dagger \in S_{\epsilon^2}$ . Since  $\mathcal{G}_{4l}$  is  $(\epsilon^2, C\epsilon^3)$ -net we can find  $y_0, z_0 \in \mathbb{R}^3$  such that,

$$\|UV^\dagger - YZY^\dagger Z^\dagger\| < C\epsilon^3 \quad (4.3.18)$$

where we defined  $Y := u(y_0)$  and  $Z := u(z_0)$ . Using the property of unitary invariance again,

$$\|UV^\dagger - YZY^\dagger Z^\dagger\| = \|U - YZY^\dagger Z^\dagger V\| < C\epsilon^3 \quad (4.3.19)$$

Since  $Y, Z, V \in \mathcal{G}_l$ , it implies that  $YZY^\dagger Z^\dagger V \in \mathcal{G}_{5l}$ , hence concluding that  $\mathcal{G}_{5l}$  is a  $(\sqrt{C}\epsilon^{\frac{3}{2}}, C\epsilon^3)$ -net.

□

## 4.4 The Solovay-Kitaev Theorem

If  $\mathcal{G}$  is a universal gate set that is closed under inverses, then we can approximate any  $U \in SU(2)$  to any accuracy  $\epsilon > 0$  by a sequence of gates in  $\mathcal{G}$ .

**Theorem 4.4:** There is a constant  $c$  such that for any  $\mathcal{G}$  and  $\epsilon > 0$  one can choose  $l = \mathcal{O}(\log^c(1/\epsilon))$  so that  $\mathcal{G}_l$  is an  $\epsilon$ -net for  $SU(2)$

The iterated "shrinking" lemma allows to obtain a good approximation for any element of  $SU(2)$  that is sufficiently close to identity. To prove the Solovay-Kitaev theorem, it is imperative to obtain good approximation of *any* elements of  $SU(2)$ . Firstly, start with a rough approximation and then refining it by invoking the iterated "shrinking" lemma for different values of  $k$  (starting with a smaller one), for  $k \in \mathbb{Z}^+$ .

Intuitively this corresponds to approaching the desired element by performing steps whose size decreases the closer we get.

*Proof.* (the **Solovay-Kitaev theorem:** Without the loss of generality we may suppose  $\epsilon_0$  has been chosen such that  $C\epsilon_0 < 1$  and therefore  $\epsilon(k)$  gets small very quick as  $k$  increases. It will also be useful to note that, provided  $\epsilon_0$  is chosen small enough,  $\epsilon^2(k) < \epsilon(k+1)$ ).



Since  $\langle G \rangle$  is dense in  $SU(2)$ , we can choose  $U_0 \in \mathcal{G}_{l_0}$  such that  $\|U - U_0\| < \epsilon^2(0)$ . Let  $\Delta_1 := UU_0^\dagger$  and then,

$$\|\Delta_1 - I\| = \|(U - U_0)U^\dagger\| = \|U - U_0\| < \epsilon^2(0) < \epsilon(1) \quad (4.4.1)$$

Hence  $\Delta_1 \in S_{\epsilon_1}$ . By invoking the iterated shrinking lemma with  $k = 1$ ,  $\exists U_1 \in \mathcal{G}_{l_1}$  such that  $\|\Delta_1 - U_1\| = \|UU_0^\dagger - U_1\| = \|U - U_1U_0\| < \epsilon^2(1)$

Similarly, let  $\Delta_2 := \Delta_1 U_1^\dagger = UU_0^\dagger U_1^\dagger$ . Then,

$$\|\Delta_2 - I\| = \|(U - U_1U_0)U_0^\dagger U_1^\dagger\| = \|U - U_1U_0\| < \epsilon^2(1) < \epsilon(2) \quad (4.4.2)$$

Hence  $\Delta_2 \in S_{\epsilon_2}$ . Invoking the shrinking lemma for  $k = 2$ , we get  $U_2 \in \mathcal{G}_{l_2}$  such that,

$$\|\Delta_2 - U_2\| = \|UU_0^\dagger U_1^\dagger - U_2\| = \|U - U_2U_1U_0\| < \epsilon^2(2) \quad (4.4.3)$$

Continuing in this way after  $k$ -steps we get  $U_k \in \mathcal{G}_{l_k}$  such that,

$$\|U - U_k U_{k-1} \dots U_0\| < \epsilon^2(k) \quad (4.4.4)$$

In doing so, we obtained a sequence of,

$$L = \sum_{m=0}^k 5^m l_0 = \frac{5^{k+1} - 1}{4} l_0 < \frac{5}{4} 5^k l_0 \quad (4.4.5)$$

gates that approximates  $U$  to accuracy  $\epsilon^2(k)$ . To approximate to some desired accuracy  $\epsilon$ , we must choose  $k$  such that,

$$\epsilon^2(k) < \epsilon \quad (4.4.6)$$

Substituting 4.3.1, taking the log on both sides, rearranging the equation and this can be expressed as,

$$\left(\frac{3}{2}\right)^k = \frac{\log(1/C^2\epsilon)}{2\log(1/C\epsilon)} \quad (4.4.7)$$

It follows that the number of gates required to approximate to within  $\epsilon$  satisfies ( $c = \log 5 / \log(3/2) \approx 4$ ) so that  $5^k = (\frac{3}{2})^{kc}$ . Then,

$$\# \text{gates} < \frac{5}{4} 5^k l_0 = \frac{5}{4} \left( \frac{3}{2} \right)^{kc} l_0 < \frac{5}{4} \left( \frac{\log(1/C^2 \epsilon)}{2 \log(1/C \epsilon_0)} \right)^c l_0 \quad (4.4.8)$$

That is, the number of gates required to approximate to within  $\epsilon$  is  $\mathcal{O}(\log^c(1/\epsilon))$ , completing the proof of the Solovay-Kitaev theorem.  $\square$

## 4.5 Generalization to $SU(d)$ .

In the previous section, we showed the Solovay-Kitaev theorem for the most basic case in  $SU(2)$ . The theorem can be extended to the general case of  $SU(d)$ ,  $d \leq 2$ . The definitions of  $\epsilon$ -net,  $(\epsilon, \epsilon_0)$ -net and  $S_\epsilon$  extend to arbitrary dimension. As in classical computation, the concept of a gate can also be generalized.

A system of  $n$ -qubits can be represented as a normalized vector in  $\mathbb{C}^{2^n}$  or more precisely as a point in  $\mathbb{CP}^{2^n-1}$  considering the equivalence between global phases. Thus we can generalize the definition of a quantum gate on a single qubit to a quantum circuit acting on  $n$ -qubits as,

**Definition 4.5.1:** A quantum gate acting on  $n$ -qubits is a unitary transformation  $U : \mathbb{CP}^{n-1} \rightarrow \mathbb{CP}^{n-1}$  such that  $U \in U(2^n)$ .

The proof of the Solovay-Kitaev theorem does not depend on the dimension, it depends uniquely on the Shrinking Lemma. Thus if we generalize the Shrinking lemma of the Solovay-Kitaev theorem will follow. To do this, it is sufficient to prove Lemma 4.3 for  $n$ -qubits. Therefore,

**Proposition 4.5.2:** Given a finite set of elements in  $SU(d)$ ,  $d \leq 2$  containing its own inverses such that  $\langle G \rangle$  is dense in  $SU(d)$ , then:

$$\text{If } G_l \text{ is } (\epsilon, \epsilon^2)\text{-net, then } G_{4l} \text{ is a } (\epsilon^2, C\epsilon^3) \quad (4.5.1)$$

for some constant  $C$ .

The proof of this result relies on Proposition 4.2.3 and the following proposition,

**Proposition 4.5.3:** Let  $H$  be a  $n \times n$  traceless Hermitian matrix. Then we can find  $F$  and  $G$  Hermitian matrices such that,

$$[F, G] = iH \quad (4.5.2)$$

$$\|F\|, \|G\| \leq n^{1/4} \left( \frac{n-1}{2} \right)^{1/2} \sqrt{\|H\|} \quad (4.5.3)$$

The proof of this Proposition and the detailed construction of the general Shrinking Lemma can be found in Section 5 of [17]. Next we will review the main idea to proof Proposition 4.5.2 using Proposition 4.3.4 i.e., given  $H$  is a Hermitian matrix, then  $d(I, e^{iH}) = \|H\| + \mathcal{O}(\|H\|^3)$ .

*Proof.* We sketch the main idea of the proof and refer to [17] for details.

Given any  $U \in SU(d)$  we can find a Hermitian matrix  $H$  such that  $U = e^{iH}$ . If  $U \in S_\epsilon^2$ , using  $d(I, U) = \|H\| + \mathcal{O}(\|H\|^3)$ , we can apply Proposition 4.5.3 to find  $F$  and  $G$  such that  $[F, G] = iH$  and  $\|F\|, \|G\| \leq c'\epsilon$  where  $c'$  is a constant term given by the proposition. Setting  $Y = e^{iF}$  and  $Z = e^{iG}$  and using Proposition 4.2.3 it follows that,

$$d(U, YZY^\dagger Z^\dagger) < C\epsilon^3 \quad (4.5.4)$$

for some constant  $C$ . Moreover  $d(I, Y), d(I, Z) < c'\epsilon$  and since  $\mathcal{G}_l$  is a  $(\epsilon, \epsilon^2)$ -net the result follows. The details for the constants can be found in the paper [17].

□

With the result, the second part of the Shriking Lemma follows with the same construction. Therefore the Solovay-Kitaev Theorem can be generalized for any arbitrary dimension.

**Theorem 4.5.4:** (solovay-Kitaev Theorem) If  $\mathcal{G} \subseteq SU(d)$  is a universal family of gates (where  $SU(d)$  is the group of unitary operators in a  $d$ -dimensional Hilbert space),  $\mathcal{G}$  is closed under inverse (i.e  $g \in \mathcal{G} \leftrightarrow g^{-1} \in \mathcal{G}$ ) and  $\mathcal{G}$  generates a dense subset of  $SU(d)$ , then  $\forall U \in SU(d), \epsilon > 0, \exists g_1, g_2, \dots, g_l \in \mathcal{G} : \|U - U_{g_1} U_{g_2} \dots U_{g_l}\| \leq \epsilon$ . Then  $\mathcal{G}_l$  is an  $\epsilon$ -net in  $SU(d)$  for

$$l = \mathcal{O}(\log^c(1/\epsilon)) \text{ with } c = \frac{\log 5}{\log 3/2} \approx 4 \quad (4.5.5)$$

For further details of this section, we recommend the reader to read up on [17] chapter 5 page 10.

# Chapter 5

## The Geometry of Quantum Complexity

We have so far established that quantum complexity can be defined as a discrete finite set of unitary gates in the  $SU(n)$  space.

In this chapter, we shall investigate quantum complexity in the notion and perspective of a smooth continuous distance between two unitaries in  $SU(n)$ . The analogue to such distance with respect to the shortest and most efficient algorithm discussed in the previous chapter will naturally be the geodesic or the shortest path between any two unitary operator.

We shall use and discuss the ideas from the papers by Nielsen et al in the ascending order [3], [2], [4] and [6] in this chapter.

We show that the optimal Hamiltonian evolution for synthesis of a desired unitary necessarily obeys a simple universal geodesic equation. We also show that finding optimal quantum circuits is essentially equivalent to finding the shortest path between two points in a certain curved geometry.

On the last part of this chapter, we discuss upper and lower bounds relating the quantum gate complexity of a unitary operation,  $U$ , to the optimal control cost associated to the synthesis of  $U$ . These bounds apply for any optimal control problem, and can be used to show that the quantum gate complexity is essentially equivalent to the optimal control cost for a wide range of problems, including time-optimal control and finding minimal distances on certain Riemannian, subriemannian, and Finslerian manifolds.

In order to measure distances in the Lie group manifold, we will use the Riemannian geometry as a tool to extrapolate these ideas. The Riemannian geometry is chosen due to its properties which can be used to calculate lengths, curves and distances in the manifold of  $SU(2^N)$  of  $N$ -qubit unitary operators. We shall start off with some important definitions and properties of topological space and manifold.

## 5.1 Preliminaries

In order to proceed with this chapter, we will require some tools and terms from differential geometry and manifold. We will keep it brief and succinct just to give the reader a basic idea of these terms which we will use it in this chapter.

**Definition 5.1.1:** A **topological space** is a set  $X$  together with a collection of open subsets  $T$  that satisfies the four conditions[18],

- The empty set  $\emptyset$  is in  $T$ .
- $X$  is in  $T$
- The intersection of a finite number of sets in  $T$  is also in  $T$ .
- The union of an arbitrary number of sets in  $T$  is also in  $T$

**Definition 5.1.2:** A **manifold** is a topological space that is locally Euclidean (i.e., around every point, there is a neighbourhood that is topologically the same as the open unit ball in  $\mathbb{R}^n$ [19]. To illustrate this idea, the Earth which is a round sphere when zoomed upon to a small scale appears at "flat". In general any object that is nearly "flat" on small scales is a manifold. More concisely, any object that can be *charted* is a manifold.

As a topological space, a manifold can be compact or noncompact, and connected or disconnected. If a manifold contains its own boundary, it is called a manifold with boundary. The closed unit ball in  $\mathbb{R}^n$  is a manifold with boundary and its boundary is the unit sphere. The concept can be generalized to manifolds with corners. By definition, every point on a manifold has a neighbourhood together with a homeomorphism<sup>1</sup> of that neighbourhood with an open ball in  $\mathbb{R}^n$ .

Smooth manifolds (also called differentiable manifolds) are manifolds for which overlapping charts "relate smoothly" to each other, meaning that the inverse of one followed by the other is an infinitely differentiable map from Euclidean space to itself.

A manifold may be endowed with more structure than a locally Euclidean topology such that it could be smooth, complex or even algebraic. A smooth manifold with a metric is called a Riemannian manifold.

**Definition 5.1.3:** A **chart** or coordinate chart is a way of expressing the points of a small neighbourhood usually on a manifold  $M$  as coordinates in Euclidean space. The pair  $(\mathcal{U}_p, \varphi_p)$  is known as a chart (with chart map  $\varphi$ ) for all points  $p$  in  $M$  and the collection of all the charts is called the atlas.

The union of all the charts cover the manifold  $M$ , i.e,

---

<sup>1</sup>A homeomorphism also called a continuous transformation is an equivalence relation and a one-to-one correspondence between points in two geometric figures or topological spaces that is continuous in both directions. A homeomorphism which also preserves distances is called an isometry.

$$\bigcup_p \mathcal{U}_p = M \quad (5.1.1)$$

**Definition 5.1.4: Tangent space.** Let  $x$  be a point in an  $n$ -dimensional compact manifold  $M$  and attach at  $x$  a copy of  $\mathbb{R}^n$  tangential to  $M$ . The resulting structure is called tangent space of  $M$  at  $x$  and is denoted  $T_x M$ . If  $\gamma$  is a smooth curve passing through  $x$  then the derivative of  $\gamma$  at  $x$  is a vector in  $T_x M$ . Figure 5.1 provides a diagrammatic representation of a tangent space  $T_x M$  on a point  $x$  on a manifold  $M$ .

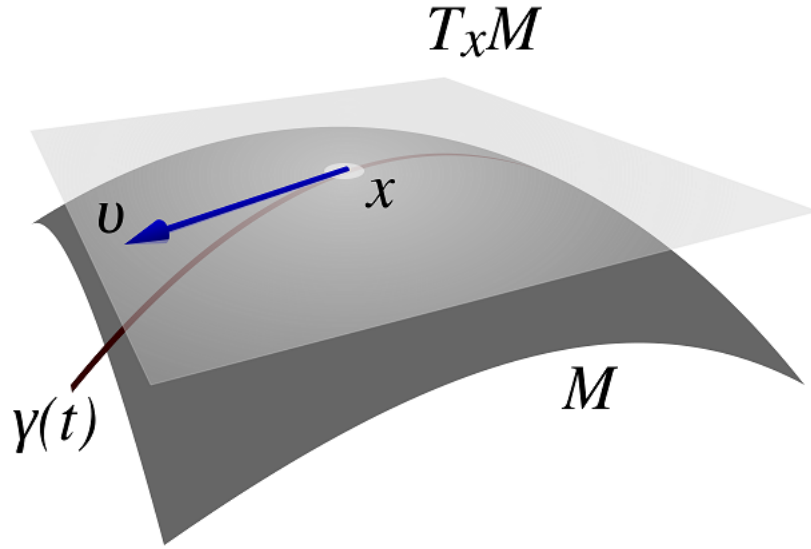


Figure 5.1: Diagrammatic illustration of tangent space  $T_x M$  at point  $x$  on  $M$  where the  $v$  is the tangent vector of  $\gamma(t)$  the parametrised curve. [20]

**Definition 5.1.5: (Linear and bilinear functionals)** If  $V$  is a finite dimensional vector space, a *linear functional* on  $V$  is a linear transformation from  $V$  into  $\mathbb{R}$ . The collection of all linear functionals is denoted  $V^*$  and is itself a vector space and its dimension is the same as  $V$ 's.

A *bilinear functional* on  $V$  is a function  $\phi : V \times V \rightarrow \mathbb{R}$  such that  $\mathbf{u} \mapsto \phi(\mathbf{v}, \mathbf{u})$  and  $\mathbf{v} \mapsto \phi(\mathbf{v}, \mathbf{u})$  are both linear functionals. We say that  $\phi$  is symmetric if  $\phi(\mathbf{u}, \mathbf{v}) = \phi(\mathbf{v}, \mathbf{u})$ . If  $\phi, \psi \in V^*$ , then we define  $\phi \odot^2 \psi : V \times V \rightarrow \mathbb{R}$  by,

$$\phi \odot \psi(\mathbf{u}, \mathbf{v}) := \frac{1}{2}(\phi(\mathbf{u})\psi(\mathbf{v}) + \phi(\mathbf{v})\phi(\mathbf{u})) \quad (5.1.2)$$

This is always a symmetric bilinear functional.

---

<sup>2</sup> $\odot$  denotes symmetric tensor product[21]

**Definition 5.1.6: 1-forms.** Suppose now that  $M \subset \mathbb{R}^n$  is a smooth  $k$ -dimensional manifold and  $\sigma$  is a regular chart for  $M$ . If  $i \leq d$ , then  $dx_i$  is the function which send  $\mathbf{p} \in M$  to the linear functional defined by,

$$(dx_i)_{\mathbf{p}}\left(\sum_{j=1}^k a_j \sigma_{x_j}\right) := a_i \quad (5.1.3)$$

A 1-*form* on  $M$  is a function  $\omega$  defined on  $M$  such that for each  $\mathbf{p} \in M$ ,  $\omega_{\mathbf{p}} \in T_{\mathbf{p}}M^*$  and for each chart  $\sigma$ ,  $\omega = \sum_{i=1}^k f_i dx_i$  where  $f_i$  is a smooth function on the range of  $\sigma$ .

**Definition 5.1.7: Bilinear forms.** Let  $M$  and  $\sigma$  be as above. If  $i, j \leq k$  then  $dx_i dx_j$  is a function defined on the range of  $\sigma$  such that  $(dx_i dx_j)_{\mathbf{p}} := (dx_i)_{\mathbf{p}} \odot (dx_j)_{\mathbf{p}}$ .

A *symmetric bilinear form* is a function  $\langle \cdot, \cdot \rangle$  such that for each  $\mathbf{p} \in M$ ,  $\langle \cdot, \cdot \rangle_{\mathbf{p}}$  is a symmetric bilinear functional on  $T_{\mathbf{p}}M$  and such that for each chart, there are smooth functions  $f_{i,j}$  such that  $\langle \cdot, \cdot \rangle = \sum_{i,j=1}^k f_{i,j} dx_i dx_j$ . A symmetric bilinear form  $\langle \cdot, \cdot \rangle$  is *non-degenerate* if for each  $\mathbf{v} \in TM$ ,  $\langle \mathbf{v}, \mathbf{v} \rangle > 0$ .

**Definition 5.1.8: The Riemannian Metric.**[22] Suppose that  $M \subseteq \mathbb{R}^n$  is a smooth  $k$ -dimensional manifold. A *Riemannian metric* is a symmetric, non-degenerate bilinear form on  $M$ .

A smooth manifold equipped with a Riemannian metric is called a *Riemannian manifold*. As noted above in the definition of bilinear functions, a Riemannian metric  $\langle \cdot, \cdot \rangle$  defines a norm  $\|\cdot\|_{\mathbf{p}}$  on  $T_{\mathbf{p}}M$  for each  $\mathbf{p} \in M$  by  $\|\mathbf{v}\|_{\mathbf{p}} := \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle_{\mathbf{p}}}$ . Moreover this norm "remembers" the bilinear function  $\langle \cdot, \cdot \rangle_{\mathbf{p}}$ :

$$\langle \mathbf{u}, \mathbf{v} \rangle_{\mathbf{p}} = \frac{1}{2}(\|\mathbf{u} - \mathbf{v}\|_{\mathbf{p}}^2 - \|\mathbf{u}\|_{\mathbf{p}}^2 - \|\mathbf{v}\|_{\mathbf{p}}^2) \quad (5.1.4)$$

**The first fundamental form.** If  $M \subseteq \mathbb{R}^n$  is a smooth manifold, there is natural Riemannian metric which  $M$  inherits from  $\mathbb{R}^n$ :

$$\langle \mathbf{u}_{\mathbf{p}}, \mathbf{v}_{\mathbf{p}} \rangle_{\mathbf{p}} = \mathbf{u} \cdot \mathbf{v} \quad (5.1.5)$$

This is called the *first fundamental form* of  $M$  and is the default Riemannian metric on  $M$  unless another is specified. If  $S \subseteq \mathbb{R}^2$  is a smooth surface with chart  $\sigma(u, v)$ , then the first fundamental form can be expressed in local coordinates as  $E du^2 + 2F du dv + G dv^2$  where,

$$E := \sigma_u \cdot \sigma_u = \|\sigma_u\|^2 \quad (5.1.6)$$

$$F := \sigma_u \cdot \sigma_v \quad (5.1.7)$$

$$G := \sigma_v \cdot \sigma_v = \|\sigma_v\|^2 \quad (5.1.8)$$

**Lengths of curves.** If  $M$  is a Riemannian manifold with a Riemannian metric  $\langle \cdot, \cdot \rangle^M$  and  $\gamma : [a, b] \rightarrow M$  is a smooth curve, then the length of  $\gamma$  with respect to the metric is,

$$\ell^M(\gamma) := \int_a^b \|\dot{\gamma}(t)\|_{\gamma(t)}^M dt \quad (5.1.9)$$

If  $M$  is clear from the context, we will write  $\ell(\gamma)$  for  $\ell^M(\gamma)$ .

**Pulling back metrics.** If  $M$  and  $N$  are smooth manifolds,  $f : M \rightarrow N$  is a smooth function and  $\langle \cdot, \cdot \rangle^N$  is a Riemannian metric on  $N$ , then  $f$  allows use to define a Riemannian metric  $f^*\langle \cdot, \cdot \rangle^N$  on  $M$  by

$$f^*\langle \mathbf{u}, \mathbf{v} \rangle^N := \langle Df\mathbf{u}, Df\mathbf{v} \rangle^N \quad (5.1.10)$$

where  $\mathbf{u}, \mathbf{v} \in TM$ . Note that  $f^*\langle \cdot, \cdot \rangle^N$  should be regarded as a single symbol - we are applying  $f^*$  to the metric itself, not composing the metric with  $f^*$ . Similarly we define  $f^*\|\cdot\|^N$  by

$$f^*\|\mathbf{v}\|^N := \|Df\mathbf{v}\|^N \quad (5.1.11)$$

**Local isometries.** If  $M$  and  $N$  are Riemannian manifolds and  $f : M \rightarrow N$  is a smooth function, we say that  $f$  is a (local) isometry if  $f$  is a (local) diffeomorphism and whenever  $\gamma : [a, b] \rightarrow M$  is a smooth curve,  $\ell^M(\gamma) = \ell^N(f \circ \gamma)$ . That is,  $f$  preserves the lengths of curves.

**Theorem 5.1.9:** Suppose that  $M$  and  $N$  are Riemannian manifolds and  $f : M \rightarrow N$  is a local diffeomorphism. The following are equivalent:

- $f$  is a local isometry.
- $f^*\|\cdot\|^N = \|\cdot\|^M$ .
- $f^*\langle \cdot, \cdot \rangle^N = \langle \cdot, \cdot \rangle^M$ .

An alternate and succinct way to define the Riemannian metric is as follows:

**Definition 5.1.9:** Suppose for every point  $x$  in a manifold  $M$ , an inner product  $\langle \cdot, \cdot \rangle_x$  is defined on a tangent space  $T_x M$  of  $M$  at  $x$ . Then the collection of all these inner products is called the Riemannian metric



## 5.2 The Riemannian Metric of the Lie Group $SU(2^N)$

The Lie group  $SU(2^N)$  has associated Lie algebra  $\mathfrak{su}(2^N)$  which consists of traceless skew hermitian matrices. A Lie algebra can be thought of as a vector space (which fits in the place of a tangent space) with additional structure called the Lie bracket.

A tangent vector at  $U \in M$ , where  $M$  is a manifold can be associated with the Hamiltonian  $H \in \mathfrak{su}(2^N)$  to the curve  $e^{-iHt}U$  at  $t = 0$ . We shall call  $H$  the *Hamiltonian representation* of this tangent vector. With these identifications, the Riemannian metric  $\langle \cdot, \cdot \rangle_U$  at point  $U$  is a positive-definite bilinear form  $\langle H, J \rangle_U$  defined on traceless Hamiltonian  $H$  and  $J$ . Throughout this paper we assume that this bilinear form is constant as a function of  $U$  and so write  $\langle \cdot, \cdot \rangle_U = \langle \cdot, \cdot \rangle$ . A metric which is constant in this way is known as right-invariant metric.

For a curve  $U(t)$  on  $M$  generated by Hamiltonian  $H(t)$  evolving according to Schrödinger's equation, the length is given by,

$$\int dt \langle H(t), H(t) \rangle^{1/2} \quad (5.2.1)$$

integrating the length element along the curve and the distance between  $U, V \in M$ ,  $d(U, V)$  is the infimum<sup>3</sup> over all curves from  $U$  to  $V$ .

We decompose  $\mathfrak{su}(2^N)$  into subspaces  $\mathcal{P}$  for one and two-qubit unitaries (one and two body Hamiltonians) and  $\mathcal{Q}$  the subspace for three or more qubit unitaries such that  $\mathcal{P} + \mathcal{Q} = \mathfrak{su}(2^N)$ . Any Hamiltonian can be uniquely decomposed as a sum  $H = H_P + H_Q$  of one and two body terms  $H_P$  and three or more body terms  $H_Q$ . Using these notations we define maps  $\mathcal{P}$  and  $\mathcal{Q}$  by  $\mathcal{P}(H) \equiv H_P$  and  $\mathcal{Q}(H) \equiv H_Q$ . We define the right-invariant Riemannian metric which we call the *standard metric*,

$$\langle H, J \rangle \equiv \frac{\text{tr}(H\mathcal{P}(J)) + p\text{tr}(H\mathcal{Q}(J))}{2^n} \quad (5.2.2)$$

where  $\mathcal{P}$  and  $\mathcal{Q}$  are projections operators on subspaces  $P, Q$  and  $p$  is a penalty parameter.

The equation above has a general form,

$$\langle H, J \rangle = \frac{\text{tr}(H\mathcal{S}(J))}{2^n} \quad (5.2.3)$$

where  $\mathcal{S}$  is a strictly positive superoperator<sup>4</sup> In equation 5.2.2,  $\mathcal{S} = \mathcal{P} + p\mathcal{Q}$ . We call a metric of this form a metric of standard form.

<sup>3</sup>A lower bound of subset  $S$  of a partially ordered set  $(P, \leq)$  is an element  $a$  of  $P$  such that  $a \leq x$  for all  $x \in S$ . A lower bound of  $a$  of  $S$  is called an *infimum* or greatest lower bound of  $S$  if for all lower bounds  $y$  of  $S$  in  $P$ ,  $y \leq a$  ( $a$  is larger than or equal to any other lower bound).

<sup>4</sup>A linear operator on vector space of linear operators satisfying  $\text{tr}(H\mathcal{S}(H)) > 0$  for  $H \neq 0$  is defined as a superoperator.

This induces a metric on  $SU(2^n)$  which satisfies the following inequalities in terms of the approximate  $G(U, \epsilon)$  and exact gate complexities  $G(U)$ , where  $G(U, \epsilon)$  is the number of one and two qubit required to approximate  $U$  within accuracy  $\epsilon$  with respect to the trace norm and  $G(U)$  is the number of the one and two qubit gates required to synthesize it exactly:

$$\frac{b_0 G(U, \epsilon)^{b_1} \epsilon^{b_2}}{n^{b_3}} \leq d(I, U) \leq G(U) \quad (5.2.4)$$

where  $b_1, b_2, b_3$  are some positive constants. These two inequalities may be summarized by saying that the distance  $d(I, U)$  gives us both a lower bound on the exact gate complexity and an upper bound on the approximate gate complexity of synthesizing  $U$ . We shall see that in the next chapter  $d(I, U)$  is also known as the cost function of synthesizing  $U$ .

This inequalities holds if  $P$  is replaced by a space generated by different universal gate set of dimension  $poly(n)$  say  $P'$ ,

$$poly(n, \epsilon, G'(U, \epsilon)) \leq d'(I, U) \leq G'(U) \quad (5.2.5)$$

where  $d', G'$  are now with respect to  $P'$ .

Note that we use "metric" interchangeably to mean the Riemannian metric tensor on  $SU(2^n)$  (that is the inner product manifold) and the metric  $d$  which turns  $SU(2^n)$  into a metric space when it is clear from context.

### 5.2.1 The Geodesic Equation

In analogy with Euclidean space where the shortest path between two points is a straight line which is determined by the endpoints and constant directional derivatives, a geodesic on a Riemannian manifold is defined as a curve for which the covariant derivative of the velocity vector field is zero[23].

For a curve passing through origin with tangent  $Y$ , the covariant derivative of a vector field  $Z$  in a fixed coordinate system  $\{x^k\}$  is given by,

$$(\nabla_Y Z)^j = \frac{\partial z^j}{\partial x^k} y^k + \Gamma_{kl}^j y^k z^l \quad (5.2.6)$$

where  $\nabla_Y Z$  is the connection,  $y^k, z^k$  are natural coordinate representations for the vector fields  $Y, Z$  with respect to the coordinate system  $\{x^k\}$ , summation is implied over repeated indices and  $\Gamma_{kl}^j$  are the Christoffel coefficients.

Now the geodesic equation is in local coordinates - for a point  $p$  in the chart  $(U, \phi)$ , local coordinates are just the identification of the open set  $U$  via  $\phi$  with the Euclidean space for us to work in.

However the metric we defined in equation 5.2.3 was in terms of Hamiltonian  $H$  which is an element of the tangent space. There is no such coordinate representation for the Hamiltonian which can be identified in a natural set of coordinates such as the  $x_j$  above. To remedy this, we introduce a fixed system of coordinates which we shall call the Pauli coordinates.

The way to get a nice set of local coordinates is by using the Lie structure about  $\mathbb{I} \in SU(2^n)$ . First note that about the origin we can associate a coordinate vector  $\hat{x}$  with a tangent vector  $H = \hat{x} \cdot \hat{\sigma} \sum x_\sigma \sigma$  (recall the tangent space at identity is the  $\mathfrak{su}(2^N)$ : the space of the skew Hermitian matrices with trace zero,  $H$  can be written in terms of generalized Pauli matrices, which can be taken as tensor products of the single-qubit Pauli matrices).  $\hat{x}$  is the Pauli representation of  $H$  corresponding to local coordinates.

Let  $U = e^{iX}$  be in a small neighbourhood about the identity and  $H$  be a Hamiltonian representing a tangent vector. We are interested in the Pauli representation of  $H$  in the tangent space at  $U$ . This follows from a computation using the Baker-Campbell-Hausdorff formula. For full details please refer to [3] page 4. In this case we are only concerned about a small neighbourhood of  $U$  as we just need to be able to know how tangent vectors change at origin in terms of local coordinates to evaluate the covariant derivative.

From applying the Baker-Campbell-Hausdorff formula as outlined in the paper, we have superoperators  $\varepsilon_X, \mathcal{D}_X$  connecting the Pauli coordinates  $J$  at  $X$  with the Hamiltonian  $H$  where  $H = \varepsilon_X(J)$  with inverse  $\varepsilon_X^{-1} = \mathcal{D}_X$  (the inverse only exists in a small neighbourhood about the origin),  $\varepsilon_X^\dagger = \varepsilon_{-X}$  and  $\mathcal{D}_X^\dagger = \mathcal{D}_{-X}$ . Extending these results we can compute the metric, Christoffel symbols and the covariant derivative (all in the local coordinates)[3] (page 5). By using these tools and objects, we define the geodesic equation given the Riemannian metric of equation 5.2.2 with  $\mathcal{S} = \mathcal{P} + q\mathcal{Q}$  on  $SU(2^n)$ , let  $M = (1 - q^{-1})\mathcal{S}(H)$ , then the geodesic equation assuming  $q > 1$  is,

$$\dot{M} = i[M, \mathcal{P}(M)] \quad (5.2.7)$$

This equation is known as Lax Equation and represents the geodesic of the Riemannian manifold in  $SU(2^n)$ .

The Lax equation is a pair of matrices or operators  $M(t), \mathcal{P}(t)$  dependent on time and acting on a fixed Hilbert space which satisfy the above equation where  $[M, \mathcal{P}] = M\mathcal{P} - \mathcal{P}M$  is the commutator.

### 5.3 The Geometry of Quantum Circuits

On this chapter, we shall discuss the application on the Riemannian geometry in the manifold of the Lie Group  $SU(2^N)$  based on the paper by Dowling and Nielsen[3]

A tangent vector to a point on the  $SU(2^N)$  manifold can be thought of as a traceless Hamiltonian  $H$ , i.e an element of the Lie algebra  $\mathfrak{su}(2^N)$  of traceless  $2^N \times 2^N$  Hermitian matrices.

In  $SU(2^N)$  the time evolution unitary operator  $U(t)$  can be thought of as a curve in the manifold between the identity operator  $I$  to the desired unitary operation  $U$ . The curve  $U$  is a *smooth* function :  $[0, t_f] \rightarrow SU(2^N)$  such that  $U(0) = I$  and  $U(t_f) = U$ . The Hamiltonian  $H$  is called the Hamiltonian representation of this tangent vector on this manifold.

The Riemannian metric  $\langle \cdot, \cdot \rangle_U$  at a point  $U$  is a positive-definite bilinear form  $\langle H, J \rangle_U$  defined on traceless Hamiltonians  $H$  and  $J$ . A metric which is constant this in this way is known as a right-invariant metric.

Suppose  $U(t)$  is a curve in  $SU(2^N)$  is generated by the Hamiltonian  $H(t)$  according to the Schrödinger equation  $\dot{U} = iHU$ . Then the length of the curve is,

$$\int \langle H(t), H(t) \rangle^{1/2} dt \quad (5.3.1)$$

The length of this curve can be defined as the total *cost* of synthesizing the Hamiltonian that generates the evolution along the curve:

$$d([U]) \equiv \int_0^{t_f} F(H(t)) dt \quad (5.3.2)$$

The cost function  $F(H(t))$  defines a Riemannian geometry on the space of unitary operations. The cost function  $F(H(t))$  imposed on the control Hamiltonian  $H(t)$  characterizes the difficulty of the computation, i.e the complexity.

Finding the optimal control function  $H(t)$  for synthesizing a desired unitary  $U$  then corresponds to finding minimal geodesics of the Riemannian geometry[2].

In order to choose a cost function on the control Hamiltonian, firstly express  $H(t)$  in terms of the Pauli operator expansion,

$$H = \sum_{\sigma}^I h_{\sigma} \sigma + \sum_{\sigma}^{II} h_{\sigma} \sigma \quad (5.3.3)$$

The first sum  $\sigma$  ranges over all possible one and two body interactions, that is all products of either one or two Pauli matrices acting on  $n$  qubits. In the second sum  $\sigma$  ranges over all other tensor products of Pauli matrices and the identity and the  $h_{\sigma}$  are real coefficients.

A measure of the cost of applying a particular Hamiltonian during the synthesis of a desired unitary operation is,

$$F(H) \equiv \sqrt{\sum_{\sigma}^I h_{\sigma}^2 + p^2 \sum_{\sigma}^{II} h_{\sigma}^2} \quad (5.3.4)$$

The definition of control cost leads to a natural notion of distance in  $SU(2^N)$ . We refer back to 5.3.2, the length of the curve  $d[U]$  is invariant with respect to different parameterizations  $[U]$ , we can always rescale the Hamiltonian  $H(t)$  such that  $F(H(t)) = 1$  and the desired unitary  $U$  is generated at time  $t_f = d([U])$ . We assume from now on that we are working with such normalized curves.

Referring to the paper by Nielsen, Downling, et al[2], we can construct a quantum circuit containing a number of gates polynomial in  $d(I, U)$  which approximates  $U$  closely, by using the optimal control Hamiltonian  $H(t)$ . The idea behind such construction is expressed through three separate lemmas before combining them to obtain the result (Figure 1)[2]. The proofs for these lemma can be found in the reference paper.

**Lemma 5.3.1:** Let  $H_P(t)$  be the projected Hamiltonian obtained from a Hamiltonian  $H(t)$  generating a unitary  $U$ . Let  $U_P$  be the corresponding unitary generated by  $H_P(t)$ . Then,

$$\|U - U_P\| \leq \frac{2^n d([U])}{p} \quad (5.3.5)$$

where  $\|\cdot\|$  is the operator norm<sup>5</sup> and  $p$  is the penalty parameter appearing in the definition of the metric. Thus, by choosing  $p$  sufficiently large, say  $p = 4^n$ , we can ensure that  $\|U - U_P\| \leq d([U])/2^n$ .

Motivated by the preceding lemma, we change our aim from accurately synthesizing  $U$  to accurately synthesize  $U_P$ . To do this, break the evolution according to  $H_P(t)$  up into many small intervals, each of the length  $\Delta$ . The next lemma shows the evolution according to the time-dependent Hamiltonian  $H_P(t)$  over such a small time interval can always be accurately simulated by a constant mean Hamiltonian, which we denote  $\bar{H}_P^\Delta$ .

**Lemma 5.3.2:** Let  $U$  be an  $n$ -qubit unitary generated by applying a time-dependent Hamiltonian  $H(t)$  satisfying  $\|H(t)\| \leq c$ , for an real constant  $c$  over a time interval  $[0, \Delta]$ . Then defining the mean Hamiltonian  $\bar{H} \equiv \frac{1}{\Delta} \int_0^\Delta dt H(t)$  we have,

$$\|U - \exp(-i\bar{H}\Delta)\| \leq 2(e^{c\Delta} - 1 - c\Delta) = \mathcal{O}(c^2\Delta^2) \quad (5.3.6)$$

To apply this lemma to  $H_P(t)$ , note that elementary norm inequalities and the observation  $F(H_P(t)) \leq 1$  imply that  $\|H_P(t)\| \leq \frac{3}{\sqrt{2}} n F(H_P(t)) \leq \frac{3}{\sqrt{2}} n$ . Lemma 5.2.2 implies that over a time interval  $\Delta$  we have,

$$\|U_P^\Delta - \exp(-i\bar{H}_P^\Delta\Delta)\| \leq 2(e^{3/\sqrt{2}n\Delta} - (1 + \frac{3}{\sqrt{2}}n\Delta)) = \mathcal{O}(n^2\Delta^2) \quad (5.3.7)$$

---

<sup>5</sup>The operator norm of  $X$  is defined as  $\|X\| = \max_{|\psi\rangle} |\langle\psi|X|\psi\rangle|$  where the maximization is over all normalized vectors,  $|\langle\psi|\psi\rangle|^2 = 1$

where  $U_P^\Delta$  is the evolution generated by  $H_P(t)$  over the time interval  $\Delta$  and  $\bar{H}_P^\Delta$  is the corresponding mean Hamiltonian.

The third and final lemma shows that evolution according to a time-independent Hamiltonian  $H$  containing only one- and two-body terms can be very accurately simulated using number of quantum gates that is not too large.

**Lemma 5.3.3 :** Suppose  $H$  is an  $n$ -qubit two-body Hamiltonian whose Pauli expansion coefficients satisfy  $|h_\sigma| \leq 1$ . There exists a unitary  $U_A$ , satisfying,

$$\|e^{-iH\Delta} - U_A\| \leq c_2 n^4 \Delta^3 \quad (5.3.8)$$

that can be synthesized using at most  $c_1 n^2 / \Delta$  one and two-qubit gates where  $c_1$  and  $c_2$  are constants.

This results follows from standard procedures for simulating quantum evaluations using quantum gates (see, e.g, Chapter 4 of [5]) and is proved in the appendix. Note that the average Hamiltonian  $\bar{H}_P^\Delta$  provided by Lemma 5.3.2 satisfies the assumptions of Lemma 5.3.3, since the Pauli expansion coefficients of  $H_P(t)$  satisfy  $|h_\sigma| \leq 1$  for all times.

To integrate Lemma 5.3.1-3, suppose  $H(t)$  is the time-dependent normalized Hamiltonian generating the minimal geodesic of length  $d(I, U)$ . Let  $H_P(t)$  be the corresponding projected Hamiltonian, which generates  $U_P$  and satisfies  $\|U - U_P\| \leq d(I, U)/2^n$ , as guaranteed by Lemma 5.3.1 where we have chosen  $p = 4^n$  as the penalty. Now divide the time interval  $[0, d(I, U)]$  up into a large number  $N$  of time intervals each of length  $\Delta = d(I, U)/N$ .

Let  $U_P^j$  be the unitary operation generated by  $H_P(t)$  over the  $j$ th time interval. Let  $U_M^j$  be the unitary operation generated by the corresponding mean Hamiltonian. Then Lemma 2 implies that,

$$\|U_P^j - U_M^j\| \leq 2(e^{(3/\sqrt{2})n\Delta} - (1 + \frac{3}{\sqrt{2}}n\Delta)) \quad (5.3.9)$$

Lemma 5.3.3 implies that we can synthesize a unitary operation  $U_A^j$  using at most  $c_1 n^2 / \Delta$  one and two-qubit gates and satisfying  $\|U_M^j - U_A^j\| \leq c_2 n^4 \Delta^3$ .

Putting all these results together and using the triangle inequality repeatedly,

$$\|U - U_A\| \leq \|U - U_P\| + \|U_P - U_A\| \quad (5.3.10)$$

$$\leq \frac{d(I, U)}{2^n} + \sum_{j=1}^N \|U_P^j - U_A^j\| \quad (5.3.11)$$

$$\leq \frac{d(I, U)}{2^n} + \sum_{j=1}^N (\|U_P^j - U_M^j\| + \|U_M^j - U_A^j\|) \quad (5.3.12)$$

$$\leq \frac{d(I, U)}{2^n} + 2 \frac{d(I, U)}{\Delta} (e^{(3/\sqrt{2})n\Delta} - (1 + \frac{3}{\sqrt{2}}n\Delta)) + c_2 d(I, U) n^4 \Delta^2 \quad (5.3.13)$$

Provided we choose  $\Delta$  to scale at most as  $1/(n^2 d(I, U))$ , we can ensure that the error in our approximation  $U_A$  to  $U$  is small while the number of gates scales as  $n^6 d(I, U)^3$ . Summing up and we have the following theorem:

**Theorem 5.3.4:** Using  $\mathcal{O}(n^6 d(I, U)^3)$  one and two-qubit gates, it is possible to synthesize a unitary  $U_A$  satisfying  $\|U - U_A\| \leq c$ , where  $c$  is any constant.

This results demonstrate that up to polynomial factors, the optimal way of generating a unitary operation is to move along a minimal geodesic curve connecting  $I$  and  $U$ . Since the length of such geodesics also provides a lower bound on the minimal number of quantum gates required to generate  $U$  which will be discussed in the next chapter, the geometric formulation offers an alternative approach which may suggest efficient quantum algorithms or provide a way of proving that a given algorithm is indeed optimal.

It would of course be highly desirable to completely classify the geodesics of the metric we construct. An infinite class of such geodesics has been constructed in a paper by Nielsen[6], and shown to have an intriguing connection to the problem of finding the closest vector in a lattice.

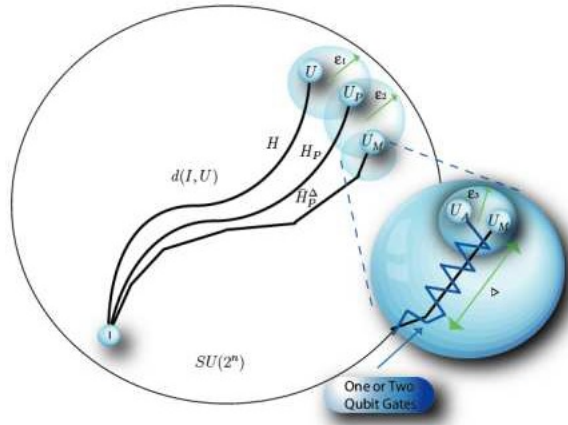


Figure 5.2: Constructing a quantum circuit to approximate  $U$

#### Comments on Figure 5.2:

Schematic of the three steps used to construct a quantum circuit approximating the

unitary operation  $U$ . The circuit is of size polynomial in the distance  $d(I, U)$  between the identity and  $U$ .

First we project the Hamiltonian  $H(t)$  for the minimal geodesic path onto one and two-qubit terms, giving  $H_P(t)$ . By choosing the penalty  $p$  large enough ( $p = 4^n$ ) we ensure the error in this approximation is small,  $\epsilon \leq d(I, U)/2^n$ .

Next we break up the evolution according to  $H_P(t)$  into  $N$  small time steps of size  $\Delta = d(I, U)/N$  and approximate with a constant mean Hamiltonian  $H_P^j$  over each step.

Finally we approximate evolution according to the constant mean Hamiltonian over each step by a sequence of one and two-qubit quantum gates. The total errors  $\epsilon_2$  and  $\epsilon_3$  introduced by these approximations can be made smaller than any desired constant by choosing the step size  $\Delta$  sufficiently small,  $\Delta = \mathcal{O}(1/(n^2 d(I, U)))$ .

In total, we need  $\mathcal{O}(n^6 d(I, U)^3)$  quantum gates to approximate  $U$  to within some constant error which can be made arbitrarily small.

## 5.4 Complexity of the upper and lower bounds

Recall that equation 5.2.2 induces a distance  $d(\cdot, \cdot)$  which is related to quantum gate complexity by the inequalities,

$$\text{poly}(n, \epsilon, G(U, \epsilon)) \leq d([U]) \leq G(U) \quad (5.4.1)$$

where  $G(U)$  is the complexity of implementing  $U$  exactly,  $G(U, \epsilon)$  is the complexity of computing an  $\epsilon$  approximation with respect to a set of unitaries  $\mathcal{U}$  and  $d$  is a metric of standard form (dependent on  $\epsilon$ ),  $d([U])$  is the distance between  $U$  and the identity with respect of  $d$ .

The general ideas for the proof of this inequalities are proven in [2][6] where the goal of these paper were to find a time-dependent Hamiltonian  $H(t)$  synthesizing  $U$ . Note that a time dependent Hamiltonian can be written as  $H(t) = \sum \gamma_\sigma(t) \sigma$  in Pauli representation.

**Definition 5.4.1:** For a Hamiltonian  $H(t) = \sum \gamma_\sigma(t) \sigma$ , the curve  $\gamma(t)$  defined as the vector of curves corresponding to each  $\sigma$ ,  $(\gamma_\sigma(t))$  is called control curves[6] which controls the synthesis according to the Schrödinger's equation with boundary conditions:

$$\dot{V} = iH(t)V \quad (5.4.2)$$

where  $V(0) = \mathbb{I}$ ,  $V(1) = U$

The natural setting for this problem is of Finsler geometry<sup>6</sup> which is a generalization of Riemannian geometry where instead of an inner product, we have a cost function on the tangent space. Next we define a *local metric*.

---

<sup>6</sup>Finsler geometry is Riemannian geometry without the restriction that the line element be quadratic and of the form:  $F^2 = g_{ij}(x) dx^i dx^j$ .



**Definition 5.4.2:** A manifold  $M$  with *local metric* is a manifold with a function  $F : TM \rightarrow [0, \infty)$  (where  $TM$  denotes the tangent bundle) such that for each  $x \in M$  and all  $y \in T_x M$ ,  $F(x, y) \leq 0$  with  $F(x, y) = 0$  iff  $y = 0$ .  $F$  is positively homogeneous in second coordinate and  $F(x, \cdot)$  satisfies the triangle inequality for each  $x$ .

So  $F$  gives a notion of length at each point on the manifold: it gives an asymmetric norm.

**Definition 5.4.3:** A manifold with a local metric  $F$  that is smooth with the Hessian of  $F(x, \cdot)^2$  at each tangent vector  $v \neq 0$  is positive definite, is a Finsler manifold. We call such  $F$  a Finsler function.

We consider  $F$  as the cost of applying  $H$  and we want two operations to be strictly higher cost than single operation.

Given a cost function on the tangent space  $c$ , we can associate a cost of applying a Hamiltonian  $H(t)$  for time  $[0, T]$  by,

$$C(H(t)) = \int_0^T dt \, c(H(t)) \quad (5.4.3)$$

We define the cost of a unitary  $U$  as the cost infimum over all permissible Hamiltonians synthesizing  $U$ ,

$$C(U) = \inf_{T, H} C(H(t)) \quad (5.4.4)$$

Note how  $C(U)$  relates to the lengths of curves when the cost function defines a metric.

We restrict to the setting of where the cost function is given by a metric of form  $\mathcal{S} = \mathcal{P} + q\mathcal{Q}$  (ignoring  $2^{-n}$  overall for simplicity) with  $\mathcal{P}$ ,  $\mathcal{Q}$  the same projection operators as equation 5.2.2,

$$c(H) = \sqrt{\langle H, H \rangle} = \sqrt{\text{tr}(\mathcal{S}(H)H)} \quad (5.4.5)$$

That is, all permissible Hamiltonians, weighted differently. For more general results on this, please refer to [4]. For a Hamiltonian  $H$ , let  $H_P$  be the Hamiltonian obtained by projecting onto the subspace  $P$ .

In order to find the  $\text{poly}(n, \epsilon, G(U, \epsilon) \leq d([U]))$  bound where  $d$  is the metric or equivalently the cost function of form 5.3.5, we begin by using the following lemmas[4]:

**Lemma 5.4.4:** For  $H$ , an  $n$ -qubit one or two body Hamiltonian (that is,  $H = \mathcal{P}(H) \equiv H_P$ ), let  $V$  be the unitary generated by  $H$  over  $[s, s + \Delta]$ . Then for the  $\Delta$ -averaged Hamiltonian  $\bar{H}(t)$ ,

$$\bar{H}(t) = \frac{1}{\Delta} \int_0^\Delta dt H(t) \quad (5.4.6)$$

we have

$$\|V - e^{-i\bar{H}\Delta}\| = \mathcal{O}(N_P^2\Delta^2) \quad (5.4.7)$$

where  $N_P$  is the maximum norm over one or two body Hamiltonians:

$$N_P = \sup_{H \in P} \|H\| \quad (5.4.8)$$

**5.4.5:** Let  $g(\Delta, \delta)$  be the complexity of approximating an arbitrary  $\Delta$ -averaged  $n$ -qubit unitary (i.e one that comes from a  $\Delta$ -averaged Hamiltonian) to an accuracy better than  $\delta$  in matrix norm using one and two qubit gates. Then  $g(\Delta, \delta) = \mathcal{O}(p(n)\Delta^2/\delta)$ , for some polynomial  $p(n)$ .

Starting with a unitary  $U$  given in terms of a Hamiltonian  $H(t)$  satisfying the conditions of equation 5.4.2, we divide an interval  $[0, T]$  into sub-intervals of length  $\Delta = T/N$  and we have,

$$\|U_P^j - U_M^j\| \leq \mathcal{O}(N_P^2\Delta^2) \quad (5.4.9)$$

where  $U_P^j$  is the unitary generated by  $H_P$  over  $j$ -th interval,  $U_M^j$  is the unitary generated by the mean Hamiltonian. We define the final approximating unitary  $U_A$  as  $U_M^j$  applied in sequence.

Subsequently, following[4] page 4, we have  $T \leq C(U)/c_A$  where  $c_A$  is the minimal cost associated to any unitary:  $c_A = \inf_H c(H)$ .

So by construction  $U_A$  has complexity at most  $Ng(\Delta, \delta) = Tg(\Delta, \delta)/\Delta$ .

Then using  $\|U - U_A\| \leq \|U - U_P\| + \|U_P - U_A\|$  and by repeated application of Lemma 5.4.4, we obtain

$$\|U - U_A\| \leq RC(U) + \mathcal{O}\left(\frac{N_P^2 C(U) \Delta}{c_A}\right) + \frac{C(U) \delta}{c_A \Delta} \quad (5.4.10)$$

where the parameter  $R = \sup_H \|H - H_P\|/c(H)$  measures the quality of approximation with respect to the cost function.

For a metric of form  $\mathcal{S} = \mathcal{P} + p\mathcal{Q}$ ,

$$c(H) = \sqrt{\sum_P h_\sigma^2 + p \sum_Q h_\sigma^2} \quad (5.4.11)$$

we have  $R \leq 2^n/p$  since  $R$  is maximized when the target Hamiltonian only has three or higher body interactions :  $c_A = 1$ ,  $N_P = \mathcal{O}(n)$  and  $g(\Delta, \mathcal{O}(n^4\Delta^3)) \leq \mathcal{O}(n^2/\Delta)$

Applying equation 5.4.10 we deduce that we can synthesize an operation  $U_A$  satisfying,

$$\|U - U_A\| \leq \frac{2^n C(U)}{p} + \mathcal{O}(\epsilon) \quad (5.4.12)$$

for  $U_A$  synthesized from  $\mathcal{O}(C(U)^3 n^6 \epsilon^4)$ . Standard results of universality imply that  $C(U) \leq \mathcal{O}(4^n)$  for all unitaries  $U$ , so by choosing  $p = 8^n/\epsilon$  we obtain,

$$G(U, \epsilon) \leq \mathcal{O}(C(U)^3 n^6 / \epsilon^2) \quad (5.4.13)$$

which is the desired polynomial scaling. This establishes the lower bound for this setting since  $C(U)$  is the length of curve from  $I$  to  $U$ .

The upper bound  $G(U)$  follows by integrating a smooth control function  $\gamma$  satisfying 5.4 on  $SU(2^n)$  equipped with a Finsler manifold structure via a Finsler function  $F$ .

We begin by looking at the minimal sequence of gates that synthesize  $U$ , say  $U = \prod_1^{G(U)} U_k$  for  $U_k = e^{-iH_k}$  where  $U_k$  come from a set  $\mathcal{A}$ . Then considering the control curve (scaled to run on  $[0, 1]$ )  $\gamma$  that applies  $H_k$  one after the other for the fixed size interval each,

$$\gamma(t) \cdot \sigma = G(U) H_i \quad (5.4.14)$$

for  $t \in [i-1)/G(U), i/G(U)]$ .

Now we require that  $\mathcal{A}$  be identified with a subset  $\mathcal{H}$  of  $\mathfrak{su}(2^N)$  such that the mapping  $\mathcal{H} \rightarrow \mathbf{A}$  is given by  $H \rightarrow e^{-iH}$  is bijective and that  $F(V, H) \leq 1$  for all  $V \in SU(2^n)$  and  $H \in \mathcal{H}$ . We call such  $\mathcal{F}$   $\mathcal{A}$ -bounding. We further require that  $\mathcal{A}$  can generate exactly  $U^7$ .

For technical reasons, we need to regularize  $\gamma(t)$  to make it smooth and we can do this by multiplying it with a positive function  $r(t)$  which has unit integral ( $r(t)$  is essentially a *modifier*). Now from positive homogeneity of  $F$  along the  $\mathcal{A}$ -boundedness, it follows that,

$$d([U]) \leq \int_0^1 F(V(t), r(t)\gamma(t)\sigma) dt \quad (5.4.15)$$

$$= \int_0^1 r(t) F(V(t), \gamma(t)\sigma) dt \leq \int_0^1 r(t) G(U) dt = G(U) \quad (5.4.16)$$

The standard metric from equation 5.2.2 relates to the Finsler functions  $F_2, F_q$  introduced in [6], hence the upper bound holds.

---

<sup>7</sup>We can always add an  $\epsilon$  and get a bound on  $G(U, \epsilon)$  by making the argument with  $U'$  which is  $\epsilon$  close to  $U$  for  $\epsilon$  depends on  $\mathcal{A}$

# Chapter 6

## Discussion and Conclusion

### 6.1 Overview and Summary of Dissertation.

This paper covers some popular ideas of quantum complexity primarily in the field of quantum computation. We begin by introducing the notion of quantum circuit complexity, mainly in the language of quantum computation and quantum information in Chapter 2.

Afterwards, we discuss the notion of complexity growth and its relationship with unitary operators and entropy of particular systems in Chapter 3. This chapter shows that the complexity of a system is exponentially proportional to the number of qubits of the system. Additionally, this gives us a wider picture of the behaviour of complexity which can find applications in different fields such as blackhole physics.

In Chapter 4, we explored the Solovay-Kitaev theorem which allows us to approximate any unitary operations of a set of finite gates efficiently up to a polylogarithmic order. Staying true to the main title of quantifying complexity, this theorem gives us the notion of quantifying the *size* of a quantum circuit for a countable amount of universal gates. The theorem is crucial in helping scientists to design quantum circuits and has fundamental importance in defining quantum class complexity such as the BQP (bounded-error quantum polynomial).

Lastly, we give a brief overview of complexity as geometry in the manifold of  $SU(2^N)$  using Riemannian geometry and the geodesic equation. This geometric reformulation of quantum complexity suggests that the tools of Riemannian geometry may be useful in analyzing quantum circuit complexity. We also review the upper and lower bounds relating the quantum gate complexity of a unitary operation,  $U$  to the optimal control cost associated to the synthesis of  $U$ . These bounds apply for any optimal control problem, and can be used to show that the quantum gate complexity is essentially equivalent to the optimal control cost for a wide range of problems, including time-optimal control and finding minimal distances on certain Riemannian, subriemannian, and Finslerian manifolds[4]. This will allow us to better understand the optimal cost for specific choices of control problem, and what it implies for quantum gate complexity.

## 6.2 Future Areas for Research.

The ideas of quantum complexity hold great potential for a wide variety of applications in modern physics. The paper outlined by Susskind offers glimpses of many applications of complexity theory such as in blackholes, quantum gravity and AdS/CFT correspondence. These are fairly new areas of physics which are an active field of research at the present time.

The Solovay-Kitaev theorem and algorithm are fundamental results in the theory of quantum computation, and it seems likely that variants of these results will be used in future implementations of quantum computers to compile quantum algorithms, such as Shor's into a fault-tolerant form. There are still some current open problems concerning the Solovay-Kitaev theorem namely, to find a lower than  $c = 1$  for any gates set, inverse-free version of the theorem and extension of the theorem to other Lie Groups.

The lower bound of quantum gate complexity  $G(U, \epsilon)$  as discussed in Chapter 5 :  $\frac{b_0 G(U, \epsilon)^{b_1} \epsilon^{b_2}}{n^{b_3}} \leq d(I, U) \leq G(U)$  is the approximate gate complexity of  $U$ , defined to be the minimal number of one- and two-qubit gates required to synthesize some  $n$ -qubit unitary operation  $V$  such that  $\|U - V\| \leq \epsilon$ [3]. This lower bound shares very similar properties of the approximated polylogarithmic number of gates of the Solovay-Kitaev theorem. Attempting to find any connection between these two objects would be the subject of our interest next.

Lastly, we have not mentioned an important aspect of quantum mechanics and complexity; quantum entanglement. It would be logical to focus our attention on the properties of complexity of entangled particles since quantum entanglement is ever-pervasive and holds great power in quantum information. The study of quantum complexity of entangled states/systems gives rise to new tools such as tensor networks[24]. A tensor network is a collection of tensors with indices connected according to a network pattern. It can be used to efficiently represent a many-body wave-function in an otherwise exponentially large Hilbert space. Therefore this offers a different angle at *quantifying* complexity. With more time and access to advanced tools and resources, tensor networks would be the main subject of interest for further research.

# References and Bibliography

- [1] L. Susskind, *Three lectures on complexity and black holes*, 2018. DOI: [10.48550/ARXIV.1810.11563](https://doi.org/10.48550/ARXIV.1810.11563). [Online]. Available: <https://arxiv.org/abs/1810.11563>.
- [2] Nielsen-Dowling-Gu-Doherty. “Quantum computation as geometry.” (Mar. 2006), [Online]. Available: <https://arxiv.org/abs/quant-ph/0603161>.
- [3] M. a. M. Dowling {and} Nielsen. “The geometry of quantum computation.” (Dec. 31, 2006), [Online]. Available: <https://arxiv.org/abs/quant-ph/0701004>.
- [4] Nielsen-Dowling-Gu-Dohert. “Optimal control, geometry, and quantum computing.” (Mar. 19, 2006), [Online]. Available: <https://arxiv.org/abs/quant-ph/0603160>.
- [5] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010. DOI: [10.1017/CBO9780511976667](https://doi.org/10.1017/CBO9780511976667).
- [6] M. Nielsen. “A geometric approach to quantum circuit lower bounds.” (Feb. 11, 2005), [Online]. Available: <https://arxiv.org/abs/quant-ph/0502070>.
- [7] R. Cleve, A. Ekert, C. Macchiavello, and M. Mosca, “Quantum algorithms revisited,” *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1969, pp. 339–354, 1998. DOI: [10.1098/rspa.1998.0164](https://doi.org/10.1098/rspa.1998.0164). [Online]. Available: <https://doi.org/10.1098/rspa.1998.0164>.
- [8] N. D. Mermin. “Breaking rsa encryption with a quantum computer: Shor’s factoring algorithm.” (2006), [Online]. Available: <https://web.archive.org/web/20121115112940/http://people.ccmr.cornell.edu/~mermin/qcomp/chap3.pdf>.
- [9] J. Simon. “Lecture notes on quantum mechanics and quantum information.” (), [Online]. Available: [https://www.learn.ed.ac.uk/ultra/courses/\\_90007\\_1/cl/outline](https://www.learn.ed.ac.uk/ultra/courses/_90007_1/cl/outline).
- [10] S. Lloyd, “Almost any quantum logic gate is universal,” *Phys. Rev. Lett.*, vol. 75, pp. 346–349, 2 1995. DOI: [10.1103/PhysRevLett.75.346](https://doi.org/10.1103/PhysRevLett.75.346). [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.75.346>.
- [11] L. J. Boya, E. Sudarshan, and T. Tilma, “Volumes of compact manifolds,” *Reports on Mathematical Physics*, vol. 52, no. 3, pp. 401–422, 2003. DOI: [10.1016/S0034-4877\(03\)80038-1](https://doi.org/10.1016/S0034-4877(03)80038-1). [Online]. Available: [https://doi.org/10.1016/S0034-4877\(03\)80038-1](https://doi.org/10.1016/S0034-4877(03)80038-1).
- [12] A. R. Brown and L. Susskind, “Second law of quantum complexity,” *Physical Review D*, vol. 97, no. 8, 2018. DOI: [10.1103/PhysRevD.97.086015](https://doi.org/10.1103/PhysRevD.97.086015). [Online]. Available: <https://doi.org/10.1103/PhysRevD.97.086015>.

- [13] H. W. Lin, “Cayley graphs and complexity geometry,” *Journal of High Energy Physics*, vol. 2019, no. 2, 2019. DOI: [10.1007/jhep02\(2019\)063](https://doi.org/10.1007/jhep02(2019)063). [Online]. Available: <https://doi.org/10.1007%2Fjhep02%282019%29063>.
- [14] A. W. Harrow, B. Recht, and I. L. Chuang, “Efficient discrete approximations of quantum gates,” *Journal of Mathematical Physics*, vol. 43, no. 9, pp. 4445–4451, 2002. DOI: [10.1063/1.1495899](https://doi.org/10.1063/1.1495899). [Online]. Available: <https://doi.org/10.1063%2F1.1495899>.
- [15] M. Ozols. “The solovay-kitaev theorem.” (Dec. 10, 2009), [Online]. Available: <http://home.lu.lv/~sd20008/papers/essays/Solovay-Kitaev.pdf>.
- [16] J. L. Zarapico, *Efficient unitary approximations in quantum computing: The Solovay-Kitaev Theorem*, 2018. [Online]. Available: <http://diposit.ub.edu/dspace/bitstream/2445/140398/1/TFG-Lumbreras-Zarapico-Josep.pdf>.
- [17] N. M. Dawson Christopher. “The solovay-kitaev algorithm.” (May 2005), [Online]. Available: <https://arxiv.org/abs/quant-ph/0505030>.
- [18] “Topological space.” (), [Online]. Available: <https://mathworld.wolfram.com/TopologicalSpace.html>.
- [19] “Manifolds.” (), [Online]. Available: <https://mathworld.wolfram.com/Manifold.html>.
- [20] “Tangent space.” (), [Online]. Available: [https://en.wikipedia.org/wiki/Tangent\\_space#/media/File:Tangentialvektor.svg](https://en.wikipedia.org/wiki/Tangent_space#/media/File:Tangentialvektor.svg).
- [21] P. K. Suetin, A. I. Kostrikin, and Y. I. Manin, *Linear Algebra and Geometry*. CRC Press, Oct. 1, 1997, 324 pp., Google-Books-ID: r28nV\_sIoxwC, ISBN: 9789056990497.
- [22] “Riemannian metric.” (), [Online]. Available: <http://pi.math.cornell.edu/~justin/4540/metric.html>.
- [23] W. T. Loring, *Differential Geometry*. Springer International Publishing, 2017.
- [24] J. Eisert, “Entanglement and tensor network states,” 2013. DOI: [10.48550/ARXIV.1308.3318](https://arxiv.org/abs/1308.3318). [Online]. Available: <https://arxiv.org/abs/1308.3318>.