The Problem of Constructing Efficient Universal Sets of Quantum Gates

Qingzhong Liang and Jessica Thompson

Abstract

The purpose of this report is threefold. First, we study the paper [Letter] in detail providing several new proofs and explanations there. Secondly we provide a new result on constructing efficient universal sets of quantum gates. Thirdly, we study two papers which give algorithms for the Solovay-Kitaev theorem. The paper [D-L] provides full details of parts 1 and 2 of this report.

Acknowledgement We would like thank Dr Steven Damelin for his help and encouragement. This research was supported via RTG grant 0943832 and the Department of Mathematics, University of Michigan.

REU webpage: http://www.ima.umn.edu/∼ damelin/REU.html

1 Introduction to Quantum Computing

1.1 Classical Bits and Quantum Bits

A *classical bit* is a basic unit of information in classical computing, with two possible states, 0 or 1.
The basic state in quantum computation is a single *quantum bit* (or *qubit* for short), which is a unit vector in $\mathbb{C}^2$.
Different from classical bits, a *quantum bit* (or *qubit* for short) can be in a linear combination of states, often called superpositions, denoted by $|\psi\rangle$:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle,$$

where $\alpha, \beta \in \mathbb{C}$, and $|\alpha|^2 + |\beta|^2 = 1$. For example, $|\psi\rangle = \frac{1}{\sqrt{3}}|0\rangle + \sqrt{\frac{2}{3}}|1\rangle$, when measuring the above qubit, $|0\rangle$ is given about 33% of the time and $|1\rangle$ about 66% of the time.

1.2 Quantum Gates

A *quantum gates* are the quantum computing's equivalent to a classical computing's logic gate. It is an operation that acts linearly on a qubit. Since qubits can be represented as unit vectors in $\mathbb{C}^2$, quantum gates can be thought of as linear maps from $\mathbb{C}^2$ to $\mathbb{C}^2$. Because it must be linear, a quantum gate can be represented by a matrix with complex coefficients. For a single qubits, it is a 2×2 matrix. Because qubits must be unit vectors, the results of applying a quantum gate to a qubit, must also be a unit vector. This implies that the matrix representing a quantum gate must be unitary. Thus, $G = PSU(2)$ (or $SU(2)$) represents all possible quantum gates that act on a single qubit.
An example of quantum gates is the *Hadamard gate*,

$$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix},$$

which sends $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ to $|\psi'\rangle = \frac{\alpha+\beta}{\sqrt{2}}|0\rangle + \frac{\alpha-\beta}{\sqrt{2}}|1\rangle$.

2 Special Unitary Group

A $n \times n$ matrix P is called *special unitary* if it is a unitary matrix with $\det P = 1$. All special unitary matrices form a group called the *special unitary group* denoted by $SU(n)$.

$$SU(n) = \{P \in U(n) | \det P = 1\} = \{P \in GL_n(\mathbb{C}) | P^\dagger P = I, \det P = 1\}$$

For example, to find the general form of a $n \times n$ matrix $P \in SU(2)$, let $P = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SU(2)$. Since $\det P = 1$, by Cramer's Rule, we have

$$P^{-1} = \frac{1}{\det P} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Since $P \in SU(2)$, we have $P^{-1} = P^\dagger$. Thus,

$$\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} \bar{a} & \bar{c} \\ \bar{b} & \bar{d} \end{bmatrix}$$
$$d = \bar{a}$$
$$c = -\bar{b}$$

Thus, for $P \in SU(2)$, P is in the form

$$\begin{bmatrix} a & b \\ -\bar{b} & \bar{a} \end{bmatrix}$$

where $a, b \in \mathbb{C}$ and $a\bar{a} + b\bar{b} = 1$.
If we set $a = x_1 + x_w i$ and $b = x_3 + x_4 i$, then the matrix P becomes,

$$\begin{bmatrix} x_1 + x_2 i & x_3 + x_4 i \\ -x_3 + x_4 i & x_1 - x_2 i \end{bmatrix}$$

Since $a, b \in \mathbb{C}$ and $a\bar{a} + b\bar{b} = 1, x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1$. Then, (x_1, x_2, x_3, x_4) is a point on the unit 3-sphere in $\mathbb{R}^4$ and each $P \in SU(2)$ corresponds to a point on S^3. This is a bijective and continuous map; hence, $SU(2)$ is homeomorphic to S^3.

3 Approximation Approach

3.1 Universal Subset

Let G be a group, $S = \{s_1, s_2, \cdots, s_n\}$ be a subset of G. Recall that S is universal if Γ, the group generated by s_i's is topologically dense in G.

3.2 Some Useful Functions

For $G = SU(2)$, let $S = \{s_1, s_2, \cdots, s_n\}$ be a finite universal subset of G, and Γ be the group generated by the $s_i's$. Then, Γ is topologically dense in G.

- Weight Function w

  The *weight function* is a non-negative function on S,

$$w : S \to \mathbb{R}_{\geqslant 0},$$

where the non-negative value $w(s_i)$ is the cost of s_i.

- Height h

  Since Γ is generated by S, for any $\gamma \in \Gamma$, there is at least one way to write γ in the form

  $$\gamma = s_{i1}s_{i2}\cdots s_{in},$$

  where $s_{i1}, s_{i2}, \cdots, s_{in} \in S$. Then, the *cost* of γ in this form is

  $$\sum_{k=1}^{n} w(s_{ik}).$$

  If $\gamma = s_{i1}s_{i2}\cdots s_{in} = s_{j1}s_{j2}\cdots s_{jm},$, we would like to compare the cost of each form. If

  $$\sum_{k=1}^{n} w(s_{ik}) \leqslant \sum_{k=1}^{m} w(s_{jk}),$$

  we will prefer to use $\gamma = s_{i1}s_{i2}\cdots s_{in}$ to construct γ. In other words, for each $\gamma \in \gamma$, we want to construct it in a way with minimal cost. Hence the *height* of each $\gamma \in \Gamma$, denoted by $h(\gamma)$, is defined as:

  $$h(\gamma) = \min\left\{\sum_{k=1}^{n} w(s_{in}) : \gamma = s_{i1}s_{i2}\cdots s_{in}\right\},$$

  which is the minimal cost to construct γ.

  For example, let $S = \{s_1, s_1^{-1}, s_2, s_2^{-1}, s_3, s_3^{-1}\}$, where

  $$s_1 = \frac{1}{\sqrt{5}}\begin{pmatrix} 1+2i & 0 \\ 0 & 1-2i \end{pmatrix},\ s_2 = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 & 2i \\ 2i & 1 \end{pmatrix},\ s_3 = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$$

  Recall that from [B-G-S], S is a universal subset of $SU(2)$. Let Γ be the group generated by S. If the weights $w(s_i)$ are all taken to be 1, then for $\gamma \in \Gamma$ and $\gamma = \begin{pmatrix} 1+2i & -4+2i \\ 4+2i & 1-2i \end{pmatrix}$, we have

  $$\gamma = s_1s_2 = s_1s_1s_1^{-1}s_2.$$

  For $\gamma = s_1s_2$, the cost is

  $$w(s_1) + w(s_2) = 1 + 1 = 2,$$

  and for $\gamma = s_1s_1s_1^{-1}s_2$, the cost is

  $$w(s_1) + w(s_1) + w(s_1^{-1}) + w(s_2) = 1 + 1 + 1 + 1 = 4.$$

  Since the cost of the first way is less, we prefer to construct γ by $\gamma = s_1s_2$, instead of $\gamma = s_1s_1s_1^{-1}s_2$. According to the definition of height, $h(\gamma) \leqslant w(s_1)+w(s_2) = 2$. Also, since $s_1, s_2, s_3, s_1^{-1}, s_2^{-1}, s_3^{-1} \neq \gamma$, we have $h(\gamma) > 1$. Hence, $h(\gamma) = 2$.

- The Set with Bounded Height $V(t)$

  With the definition of height, the set $V(t)$ is defined as

$$V(t) = \{\gamma \in \Gamma : h(\gamma) \leqslant t\},$$

which is the subset, in which every element can be constructed with cost less or equal to t. It is a set with restriction to height. We will use subsets in this form to approximate any element $x \in G$.

4 Invariant Metric in $SU(2)$

Definition. *Let G be a group and d_G be a metric on G. d_G is called* left invariant *if $d_G(gx, gy) = d_G(x, y)$ for any $x, y, g \in G$,* right invariant *if $d_G(xg, yg) = d_G(x, y)$ for any $x, y, g \in G$, and* invariant *if d_G is both left invariant and right invariant.*

For $G = SU(2)$, define the metric on G by $d_G : G \times G \to R_{\geqslant 0}$,

$$d_G(x, y) = \sqrt{1 - \frac{|\text{trace}(x^* y)|}{2}},$$

for $x, y \in SU(2)$. To prove d_G is invariant, notice that for any $h \in SU(2)$, we have $h^* h = I$. Hence,

$$(hx)^* hy = x^* h^* hy = x^* (h^* h) y = xy.$$

Also, since $h^* = h^{-1}$, we have $(xh)^* yh = h^* x^* yh = h^{-1}(x * y)h$. Also, since $\text{trace}(h^{-1}(x * y)h) = \text{trace}(x^* y)$, we have

$$\text{trace}((xh)^* yh) = \text{trace}(h^{-1}(x * y)h) = \text{trace}(x^* y).$$

Then,

$$d_G^2(x, y) = d_G^2(hx, hy) = d_G^2(xh, yh)$$

Hence, d_G is a invariant metric on $SU(2)$.

5 Approach

Now, our goal is using elements of a subset V of G to approximate each point $x \in G$. To guarantee the accuracy, the error should not be greater than a given value $\varepsilon > 0$.
As we have mentioned, we want to use an appropriate subset $V(t)$ to approximate G. That is, for any $x \in G$, there exists $\gamma \in V(t)$ that $d_G(x, \gamma) < \varepsilon$, and we use γ to approximate x. In other words, for any $x \in G$, there exists $\gamma \in V(t)$ that x is within the ball centered at γ with radius ε, i.e. $x \in B_G(\gamma, \varepsilon)$. Actually, for a given $\varepsilon > 0$, there is a least $t = t_\varepsilon > 0$ that $V(t_\varepsilon)$ can be used to approximate $G = SU(2)$. That is, G is covered by the balls:

$$G \subset \bigcup_{\gamma \in V(t_\varepsilon)} B_G(\gamma, \varepsilon).$$

6 Haar Measure

6.1 Definition

Let G be a compact group, a *Haar measure* on G is a measure $\mu : \Sigma \to \mathbb{R}_{\geqslant 0}$, with Σ a σ-algebra containing all Borel subsets of G, that

i) $\mu(G) = 1$

ii) For all $g \in G$ and $S \in \Sigma$, $\mu(gS) = \mu(S)$, where $gS = \{gs | s \in S\}$.

6.2 Haar Measure on $G = SU(2)$

Since $G = SU(2)$ is a compact group, there is a normalized Haar measure μ on $G = SU(2)$. Since for given $\varepsilon > 0$, we have

$$G \subset \bigcup_{\gamma \in V(t_\varepsilon)} B_G(\gamma, \varepsilon).$$

Then,

$$\mu(\bigcup_{\gamma \in V(t_\varepsilon)} B_G(\gamma, \varepsilon)) = \mu(G) = 1.$$

Then consider any two elements $\gamma_1, \gamma_2 \in V(t)$, and suppose $B_G(\gamma_1, \varepsilon)$ and $B_G(\gamma_2, \varepsilon)$ are the two balls centered at them. Let $g = \gamma_2 \gamma_1^{-1} \in G$, then $g\gamma_1 = \gamma_2$ and $\gamma_1 = g^{-1}\gamma_2$. Then, for any $h \in B_G(\gamma_1, \varepsilon)$, we have

$$d_G(\gamma_1, h) < \varepsilon.$$

Since $d_G(x, y)$ is an invariant metric on $G = SU(2)$, we have

$$d_G(\gamma_2, gh) = d_G(g\gamma_1, gh) = d_G(\gamma_1, h) < \varepsilon.$$

Hence, $gh \in B_G(\gamma_2, \varepsilon)$.
Similarly, if $h' \in B_G(\gamma_2, \varepsilon)$, we have

$$d_G(\gamma_2, h') < \varepsilon.$$

Then,

$$d_G(\gamma_1, g^{-1}h') = d_G(g^{-1}\gamma_2, g^{-1}h') = d_G(\gamma_2, h') < \varepsilon.$$

Hence, $g^{-1}h' \in B_G(\gamma_1, \varepsilon)$. Hence,

$$B_G(\gamma_2, \varepsilon) = gB_G(\gamma_1, \varepsilon).$$

Since the Haar measure on μ on G is left invariant

$$\mu(B_G(\gamma_2, \varepsilon)) = \mu(gB_G(\gamma_1, \varepsilon)) = \mu(B_G(\gamma_1, \varepsilon)).$$

Hence, the Haar measure of any two ball in $SU(2)$ satisfies:

$$\mu(B_G(\gamma_1, \varepsilon)) = \mu(B_G(\gamma_2, \varepsilon)) = \mu(B_G(\varepsilon)),$$

for any $\gamma_1, \gamma_2 \in SU(2)$. Here, $\mu(B_G(\varepsilon))$ is any ball with radius ε in G.
Also, since there are totally $|V(t_\varepsilon)|$ balls, and any two of the balls may have non-empty intersection. Hence

$$|V(t_\varepsilon)| \cdot \mu(B_G(\varepsilon)) = \sum_{\gamma \in V(t_\varepsilon)} \mu(B_G(\gamma, \varepsilon)) \geqslant \mu(\bigcup_{\gamma \in V(t_\varepsilon)} B_G(\gamma, \varepsilon)) = \mu(G) = 1,$$

then,

$$|V(t_\varepsilon)| \cdot \mu(B_G(\varepsilon)) \geqslant 1.$$

Moreover, the Haar measure μ can be thought of as the volume of a ball with radius ε on the surface of S^3. Hence,

$$\mu(B_G(\varepsilon)) \sim \varepsilon^3,$$

when ε is small.

7 Efficiency of Approximations

The *efficiency* of a universal gate set S is measured by its ability to approximate any $x \in G$ by γ's with small height. Also, for a given $\varepsilon > 0$, we have proved that $|V(t)| \cdot \mu(B_G(\varepsilon)) \geqslant 1$. Hence,

$$|V(t)| \geqslant \frac{1}{\mu(B_G(\varepsilon))}$$

for any ε and universal subset S. When the values of $|V(t)|$ and $\frac{1}{\mu(B_G(\varepsilon))}$ become closer, the measure of the intersections among the balls is decreased, which means the approximation is more efficient. Mathematically, we express the efficiency of an approximation with universal subset S by

$$K(S) := \overline{\lim_{\varepsilon \to 0}} \frac{\log |V(t_\varepsilon)|}{\log(\frac{1}{\mu(B_G(\varepsilon))})}.$$

Since $|V(t_\varepsilon)| \geqslant \frac{1}{\mu(B_G(\varepsilon))}$, we have

$$K(S) \geqslant 1.$$

And, if $K(S) = 1$, the generating gates S are optimally asymptotically efficient. Hence, the goal now is to find how small $K(S)$ we can make, by choosing a universal subset S suitably.

8 Approximation Example for $q \equiv 1(\mod 4)$

8.1 The Universal Subset

For convenience, we will sometimes use $q = 5$ as an example, but the results are true for $q \equiv 1(\mod 4)$. Consider the Pauli matrices I, X, Y, Z, where

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

Then, let $S = \{s_1, s_1^{-1}, s_2, s_2^{-1}, s_3, s_3^{-1}\}$, where

$$s_1 = \frac{1}{\sqrt{5}}(I+2iX) = \frac{1}{\sqrt{5}}\begin{pmatrix} 1+2i & 0 \\ 0 & 1-2i \end{pmatrix},\ s_1^{-1} = \frac{1}{\sqrt{5}}(I-2iX) = \frac{1}{\sqrt{5}}\begin{pmatrix} 1-2i & 0 \\ 0 & 1+2i \end{pmatrix}$$
$$s_2 = \frac{1}{\sqrt{5}}(I+2iY) = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 & 2i \\ 2i & 1 \end{pmatrix},\ s_2^{-1} = \frac{1}{\sqrt{5}}(I-2iY) = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 & -2i \\ -2i & 1 \end{pmatrix}$$
$$s_3 = \frac{1}{\sqrt{5}}(I+2iZ) = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix},\ s_3^{-1} = \frac{1}{\sqrt{5}}(I-2iZ) = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}.$$

Since $s_1, s_2, s_3 \in SU(2)$, S is a subset of $SU(2)$. Also, from [B-G-S], Γ, the subgroup generated by S, is dense in $SU(2)$. Hence, $S = \{s_1, s_1^{-1}, s_2, s_2^{-1}, s_3, s_3^{-1}\}$ is universal.
If the weights $w(s_i)$ are all taken to be 1, then for $\gamma \in \Gamma$, the height $h(\gamma)$ is the reduced word length of γ.
Define

$$U(t) = \{\gamma \in \Gamma | h(\gamma) = t\}.$$

There are 6 elements with reduced word length 1: $s_1, s_1^{-1}, s_2, s_2^{-1}, s_3, s_3^{-1}$, hence,

$$|V(1)| = |U(1)| = 6.$$

Also, if $\gamma \in \Gamma$ has reduced word length t with the last symbol $a \in S$, then we can add any symbols in S except a^{-1} to obtain a reduced word with length $t+1$. Hence,

$$|U(t+1)| = |U(t)| \times 5, \text{ for } t \geqslant 1.$$

Hence,

$$|U(t)| = \begin{cases} 1, & t = 0 \\ 6 \cdot 5^{t-1}, & t \geqslant 1 \end{cases}.$$

Hence, for $t \geqslant 1$,

$$|V(t)| = \sum_{k=1}^{t} |U(k)| = 1 + 6 + 6 \cdot 5 + \cdots + 6 \cdot 5^{t-1} = \frac{3}{2}5^t - \frac{1}{2}.$$

Hence,

$$|V(t)| = \begin{cases} 1 & t = 0 \\ \frac{3}{2}5^t - \frac{1}{2} & t \geqslant 1 \end{cases}.$$

In general,

$$|U(t)| = \begin{cases} 1, & t = 0 \\ (q+1) \cdot q^{t-1}, & t \geqslant 1 \end{cases},$$

and,

$$|V(t)| = \begin{cases} 1 & t = 0 \\ \frac{q+1}{q-1}q^t - \frac{2}{q-1} & t \geqslant 1 \end{cases}$$

8.2 Representation

From [B-G-S], we can represent the elements of the group as projections of integer solutions of spherical functions on S^3.

Theorem. *In the setting of this example* ($q = 5$)*, each point of* $V(t)$ *actually corresponds to an integer solution of*

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = 5^h,\ h \leqslant t.$$

The theorem is based on the following lemma and the following isomorphism.

Lemma. $U = \left\{p = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} | N(p) = 5^l,\ l \geqslant 0\right\}$.

Let

$$U_1 = \left\{ \frac{p}{\sqrt{N(p)}} |\ p \in U \right\}.$$

Then U_1 is generated by $\{\mathbf{i}, \mathbf{j}, \mathbf{k}, \frac{1}{\sqrt{5}}(\mathbf{1} + 2\mathbf{i}), \frac{1}{\sqrt{5}}(\mathbf{1} + 2\mathbf{j}), \frac{1}{\sqrt{5}}(\mathbf{1} + 2\mathbf{k})\}$. By the mapping property, there exists a homomorphism $\varphi : U_1 \to W$ with

$$\begin{aligned}
&\pm\mathbf{i} \mapsto \pm iX,\ \ \pm\mathbf{j} \mapsto \pm iY,\ \ \pm\mathbf{k} \mapsto \pm iZ,\\
&\frac{1}{\sqrt{5}}(1 \pm 2\mathbf{i}) \mapsto \frac{1}{\sqrt{5}}(1 \pm 2iX) = s_1, s_1^{-1}\\
&\frac{1}{\sqrt{5}}(1 \pm 2\mathbf{j}) \mapsto \frac{1}{\sqrt{5}}(1 \pm 2iY) = s_2, s_2^{-1}\\
&\frac{1}{\sqrt{5}}(1 \pm 2\mathbf{k}) \mapsto \frac{1}{\sqrt{5}}(1 \pm 2iZ) = s_3, s_3^{-1}.
\end{aligned}$$

Also, since W is generated by $iX, iY, iZ, 1 + 2iX, 1 + 2iY, 1 + 2iZ$, φ is surjective. And since $\ker\varphi = \{\mathbf{1}\}$, φ is injective, and hence is an isomorphism. Then, since Γ is a subgroup of W, Γ is isomorphic to the subgroup of U_1, $\langle \frac{1}{\sqrt{5}}(1 \pm 2\mathbf{i}), \frac{1}{\sqrt{5}}(1 \pm 2\mathbf{j}), \frac{1}{\sqrt{5}}(1 \pm 2\mathbf{k})\rangle$.
Hence, for any $\gamma \in V(t) \subset \Gamma$, we have $h(\gamma) \leqslant t$, and

$$\varphi^{-1}(\gamma) = \frac{1}{\sqrt{5}}q_1 \cdot \frac{1}{\sqrt{5}}q_2 \cdots \frac{1}{\sqrt{5}}q_{h(\gamma)} = \frac{1}{5^{\frac{h(\gamma)}{2}}}q_1 q_2 \cdots q_{h(\gamma)},$$

where $q_i \in \{1 \pm 2\mathbf{i}, 1 \pm 2\mathbf{j}, 1 \pm 2\mathbf{k}\}$. Since, $N(\varphi^{-1}(\gamma)) = 1$, we have

$$\varphi^{-1}(\gamma) = \frac{a}{5^{\frac{h(\gamma)}{2}}} + \frac{b}{5^{\frac{h(\gamma)}{2}}}\mathbf{i} + \frac{c}{5^{\frac{h(\gamma)}{2}}}\mathbf{j} + \frac{d}{5^{\frac{h(\gamma)}{2}}}\mathbf{k},$$

and

$$\frac{a^2}{5^{h(\gamma)}} + \frac{b^2}{5^{h(\gamma)}} + \frac{c^2}{5^{h(\gamma)}} + \frac{d^2}{5^{h(\gamma)}} = 1,\ h(\gamma) \leqslant t.$$

Hence, $\gamma \in V(t)$ corresponds to an integer solution of

$$a^2 + b^2 + c^2 + d^2 = 5^l,$$

where $l = h(\gamma) \leqslant t$.
Example. Let

$$\gamma = s_1 s_2 = \frac{1}{5}\begin{pmatrix} 1 + 2i & -4 + 2i \\ 4 + 2i & 1 - 2i \end{pmatrix} \in V(t),\ \text{for } t \geqslant 2.$$

We can check that

$$1^2 + 2^2 + (-4)^2 + 2^2 = 1 + 4 + 16 + 4 = 25 = 5^2 = 5^{h(\gamma)}.$$

Hence, γ corresponds to the integer solution $(1, 2, -4, 2)$ of $x_1^2 + x_2^2 + x_3^2 + x_4^2 = 5^2$.

8.3 The Range of $K(S)$

8.3.1 A Useful Fact

From [Letter], we have the following useful fact for the lower bound of $K(S)$. It is based on the fact that $SU(2)$ is homeomorphic to S^3. See [D-L].

Fact. *If $y \in \mathbb{Z}^4$, $y \neq 0$, then the ball $B_{S^3}(\frac{y}{|y|}, \frac{1}{10(|y|5^{\frac{t}{2}})^{\frac{1}{2}}})$ contains a ball in S^3 with radius $r = \frac{1}{20t(|y|5^{\frac{t}{2}})^{\frac{1}{2}}}$, which has no points corresponds to elements in $V(t)$.*

8.3.2 Lower Bound of $K(S)$

Then, we can calculate the lower bound of $K(S)$. By definition, for a given ε, t_ε is the least t for which

$$G \subset \bigcup_{\gamma \in V(t_\varepsilon)} B_G(\gamma, \varepsilon).$$

Since there is $y \in \mathbb{Z}^4$ and a ball in S^3 of radius

$$\frac{1}{20t_\varepsilon(|y|5^{\frac{t_\varepsilon}{2}})^{\frac{1}{2}}}$$

which has no points of $V(t)$, we have

$$\varepsilon > \frac{1}{20t_\varepsilon(|y|5^{\frac{t_\varepsilon}{2}})^{\frac{1}{2}}},$$

$$\frac{1}{\varepsilon} < 20t_\varepsilon(|y|5^{\frac{t_\varepsilon}{2}})^{\frac{1}{2}} = 20t_\varepsilon|y|^{\frac{1}{2}}5^{\frac{t_\varepsilon}{4}}.$$

Hence,

$$\begin{aligned}
K(S) &= \overline{\lim_{\varepsilon \to 0}} \frac{\log|V(t_\varepsilon)|}{\log(\frac{1}{\mu(B_G(\varepsilon))})} \\
&= \overline{\lim_{\varepsilon \to 0}} \frac{\log_5(\frac{3}{2}(5^{t_\varepsilon} - 1))}{\log_5(\frac{1}{\varepsilon^3})} \\
&\geqslant \lim_{\varepsilon \to 0} \frac{\log_5(\frac{3}{2}(5^{t_\varepsilon} - 1))}{3\log_5(20t_\varepsilon|y|^{\frac{1}{2}}5^{\frac{t_\varepsilon}{4}})} \\
&= \lim_{\varepsilon \to 0} \frac{\log_5 5^{t_\varepsilon} + \log_5(\frac{3}{2}(\frac{5^{t_\varepsilon}-1}{5^{t_\varepsilon}}))}{3\log_5(5^{\frac{t_\varepsilon}{4}}) + 3\log_5(20t_\varepsilon|y|^{\frac{1}{2}})} \\
&= \lim_{t_\varepsilon \to \infty} \frac{t_\varepsilon + \log_5(\frac{3}{2}(\frac{5^{t_\varepsilon}-1}{5^{t_\varepsilon}}))}{\frac{3t_\varepsilon}{4} + 3\log_5(20t_\varepsilon|y|^{\frac{1}{2}})} \\
&= \frac{4}{3}.
\end{aligned}$$

8.3.3 An Equivalent Expression of $K(S)$

Now, we calculate the upper bound of $K(S)$. Recall that in order to approximate any $x \in G = SU(2)$, for given accuracy ε, we choose t_ε to be the least t for which

$$G \subset \bigcup_{\gamma \in V(t)} B_G(\gamma, \varepsilon).$$

Note that $SU(2)$ is a double cover of $SO(3)$, which is the rotation group of $\mathbb{R}^3$. That is

$$SU(2)/\{I, -I\} \cong SO(3).$$

Since $-I \notin \Gamma$, if $P \in V(t)$,then $-P \notin V(t)$. Hence, if we use $V_0(t)$ to denote the corresponding elements of $V(t)$ in $SO(3)$, then $|V_0(t)| = |V(t)| = \frac{3}{2}(5^t - 1)$.
Then, approximating $SU(2)$ can be done by approximating $SO(3)$. Again, let t_ε be the smallest t that $V_0(t)$ can be used to approximate every element of $SO(3)$. Then,

$$K(S) = \overline{\lim_{\varepsilon \to 0}} \frac{\log |V_0(t_\varepsilon)|}{\log(\frac{1}{\mu(B_{S^2}(\varepsilon))})},$$

where

$$\mu(B_{S^2}(\varepsilon)) \sim c\varepsilon^2.$$

8.3.4 Hecke Operator

Recall that

$$|U_0(t)| = |U(t)| = \begin{cases} 1, & t = 0 \\ (q+1) \cdot q^{t-1}, & t \geqslant 1 \end{cases},$$

and

$$\begin{aligned} |V_0(t)| = |V(t)| &= \sum_{i=0}^{t} |U(t)| \\ &= \frac{q+1}{q-1} q^t - \frac{2}{q-1}. \end{aligned}$$

Hence,

$$q^{\frac{t}{2}} = \sqrt{\frac{q-1}{q+1}|V(t)| + \frac{2}{q+1}}.$$

Denote

$$U_0(t) = \left\{A_1, A_2, \cdots, A_N, A_1^{-1}, A_2^{-1}, \cdots, A_N^{-1}\right\} \subset SO(3),$$

where $2N = (q+1) \cdot q^{t-1}$. Then the Hecke operator is $T : L^2(S^2) \to L^2(S^2)$,

$$T_t f(x) = \sum_{i=1}^{N} (f(A_i x) + f(A_i^{-1} x)) = \sum_{\gamma \in U_0(t)} f(\gamma x),$$

and

$$T_{V_0(t)} f(x) = \sum_{\gamma \in V_0(t)} f(\gamma x) = \sum_{i=0}^{t} T_i f(x),$$

where $L^2(S^2) = \left\{f : S^2 \to \mathbb{C} | \int_{S^2} |f(x)|^2 \mathrm{d}x < \infty\right\}$. For example, when $q = 5$,

$$T_1 f(x) = \sum_{s \in U_0(1)} f(sx)$$
$$= f(S_1 x) + f(S_1^{-1} x) + f(S_2 x) + f(S_2^{-1} x) + f(S_3 x) + f(S_3^{-1} x).$$

Also, from [L-P-S 1], $L^2(S^2)$ decomposes under the Laplace operator Δ as

$$L^2(S^2) = \bigoplus_{n=0}^{\infty} H_n,$$

where H_n is the space of spherical harmonics of degree n. Each H_n is an eigenspace for Δ with eigenvalue $n(n+1)$ and the dimension of H_n is $2n+1$.
Since the Hecke operator is self-adjoint, $S^2 \subset \mathbb{R}^3$, and $T(H_n) = H_n$. By the Spectral theorem, there exists a sequence of eigenvalues

$$\lambda_0(t), \lambda_1(t), \lambda_2(t), \cdots \in \mathbb{R},$$

and an orthonormal basis of $L^2(S^2)$

$$\phi_0, \phi_1, \phi_2 \cdots,$$

consisting of eigenfunctions of the Hecke operator such that

$$T_t \phi_j = \lambda_j \phi_j,$$

and each ϕ_j is an orthonormal basis of some H_n. That is, $\{\phi_0\}$ is the orthonormal basis of H_0, $\{\phi_1, \phi_2, \phi_3\}$ is the orthonormal basis of H_1, $\{\phi_4, \phi_5, \phi_6, \phi_7, \phi_8\}$ is the orthonormal basis of H_2, $\cdots$

8.3.5 Eigenvalues of Hecke Operator

Let $x_0 \in S^2$ be a fixed point, then the Hecke orbit of x_0 in S^2 is $O_{x_0} = \{sx_0 | s \in V_0(t)\}$. By the definition of eigenfunction, we have

$$T_t \phi_j(x_0) = \sum_{s \in U_0(t)} \phi_j(sx_0) = \lambda_j(t) \phi_j(x_0).$$

In particular, H_0 is 1-dimensional with eigenvalue 0, and the orthonormal basis of H_0 is the constant spherical function

$$\phi_0(x) = \frac{1}{\sqrt{4\pi}}.$$

Hence, we have

$$T_t \phi_0(x_0) = \sum_{s \in U_0(t)} \phi_0(sx_0) = |U_0(t)| \frac{1}{\sqrt{4\pi}} = \lambda_0(t) \phi_0(x_0) = \lambda_0(t) \frac{1}{\sqrt{4\pi}}.$$

Then,

$$\lambda_0(t) = |U_0(t)| = (q+1) \cdot q^{t-1}.$$

We have already seen that $\lambda_0(T_t) = |U(t)|$ with constant eigenfunction. We want to bound the other eigenvalues. From [L-P-S 2], or from the Ramanujan Conjectures, we have the following theorem

Theorem. *For a prime $q \equiv 1(\mod 4)$, the eigenvalues of T_t satisfy that*

$$|\lambda_j(T_1)| \leqslant 2\sqrt{q}, \text{ for } j \geqslant 1.$$

Now, let us find more relations between the eigenvalues. From the properties of the Hecke Operator, it can be verified that we have the following

Theorem 8.1. *For a prime $q \equiv 1(\mod 4)$, the Hecke operator T_t satisfies that*

$$T_{t+1}f(x) = T_1(T_t f(x)) - qT_{t-1}f(x), \text{ for } t \geqslant 2, \tag{1}$$

$$T_2 f(x) = T_1(T_1 f(x)) - (q+1)f(x), \tag{2}$$

$$T_0 f(x) = f(x). \tag{3}$$

Also, the eigenfunction corresponding to $\lambda_j(T_{t+1})$ is the nonconstant spherical harmonic function ϕ_j. Since $L^2(S^2) = \bigoplus_{n=0}^{\infty} H_n$ and the orthonormal bases of H_k do not depend on t, ϕ_j is also the eigenfunction of T_n, for $n \leqslant t$ with eigenvalue $\lambda_j(T_n)$.

Proposition 1. *For $t \in \mathbb{N}^*$, the Hecke operators $T_t (\geqslant 1)$ have the same orthonormal eigenfunctions.*

Moreover, an eigenfunction is not a zero function. Combining with the proof of Proposition 1, we have the following:

Theorem 8.2. *Let ϕ_j be a eigenfunction, then the corresponding eigenvalues satisfy*

$$\lambda_j(T_{t+1}) = \lambda_j(T_1)\lambda_j(T_t) - q\lambda_j(T_{t-1}), \text{ for } t \geqslant 2, \tag{4}$$

$$\lambda_j(T_2) = (\lambda_j(T_1))^2 - (q+1), \tag{5}$$

$$\lambda_j(T_0) = 1. \tag{6}$$

Then, we consider $T_{V_0(t)}$, the Hecke operator of elements with height less or equal to t. From the definition, $T_{V_0(t)}f(x) = \sum_{\gamma \in V_0(t)} f(\gamma x) = \sum_{i=0}^{t} T_i f(x)$, hence

$$\lambda_j(T_{V_0(t)}) = \sum_{i=0}^{t} \lambda_j(T_i).$$

On the other hand, from (4) we have

$$\begin{aligned}\sum_{i=2}^{t} \lambda_j(T_{i+1}) &= \sum_{i=2}^{t} (\lambda_j(T_1)\lambda_j(T_i) - q\lambda_j(T_{i-1})) \\ &= \lambda_j(T_1)\sum_{i=2}^{t}\lambda_j(T_i) - q\sum_{i=2}^{t}\lambda_j(T_{i-1}).\end{aligned}$$

Combining with (5) and first few terms, we have the following recurrence relation

Proposition 2. *The eigenvalues of $T_{V_0(t)}$ satisfy the recurrence relation*

$$\lambda_j(T_{V_0(t+1)}) = \lambda_j(T_1)\lambda_j(T_{V_0(t)}) - q\lambda_j(T_{V_0(t-1)}), \text{ for } t \geqslant 1 \tag{7}$$

and $\lambda_j(T_{V_0(0)}) = 1$.

Since $\lambda_j(T_1)$ is a constant, and from Theorem, $|\lambda_j(T_1)| \leqslant 2\sqrt{q}$ for $j \geqslant 1$. Hence the characteristic function of the sequence $\{\lambda_j(T_{V_0(t)})$

$$x^2 - \lambda_j(T_1)x + q = 0$$

has two equal real solutions(when $|\lambda_j(T_1)| = 2\sqrt{q}$) or two conjugative solutions (when $|\lambda_j(T_1)| < 2\sqrt{q}$). Then, we can check that

- If $|\lambda_j(T_1)| < 2\sqrt{q}$, we have

$$\lambda_j(T_{V_0(t)}) = q^{\frac{t}{2}} \left(\frac{\sin(t+1)\theta_j}{\sin\theta_j} + \frac{\sin t\theta_j}{q^{\frac{1}{2}}\sin\theta_j} \right) \text{ for } j \geqslant 1. \tag{8}$$

- If $|\lambda_j(T_1)| = 2\sqrt{q}$, we have

$$\lambda_j(T_{V_0(t)}) = q^{\frac{t}{2}}[1 + (1 + \frac{1}{q^{\frac{1}{2}}})t], \text{ for } j \geqslant 1. \tag{9}$$

Hence, we have the following:

Theorem. *For a prime $q \equiv 1(\mod 4)$, the eigenvalues of T_t satisfy that*

$$|\lambda_j(T_{V_0(t)})| \leqslant 2q^{\frac{t}{2}}t, \text{ for } j \geqslant 1.$$

From this theorem and $q^{\frac{t}{2}} = \sqrt{\frac{q-1}{q+1}|V(t)| + \frac{2}{q+1}}$, we have

$$|\lambda_j(T_{V_0(t)})| \leqslant 2t\sqrt{\frac{q-1}{q+1}|V(t)| + \frac{2}{q+1}}, \text{ for } j \geqslant 1.$$

8.3.6 Point-Pair Invariant

Let $k_\varepsilon(x, y)$ be a point-pair invariant on S^2 such that

$$k_\varepsilon(\sigma x, \sigma y) = k_\varepsilon(x, y), \text{ for } \sigma \in O_f,$$

and with the properties,

$$k_\varepsilon(x, y) \geqslant 0,$$

$$\int_{S^2} k_\varepsilon(x, y)\mathrm{d}y = 1,$$

$$k_\varepsilon(x, y) = 0 \text{ if and only if } d_{S^2}(x, y) \geqslant \varepsilon.$$

From the last two properties, we have

$$\int_{S^2} k_\varepsilon(x, y)\mathrm{d}y = \int_{B_{S^2}(x,\varepsilon)} k_\varepsilon(x, y)\mathrm{d}y = 1.$$

And since

$$\int_{B_{S^2}(x,\varepsilon)} \mathrm{d}y = \mu(B_{S^2}(x, \varepsilon)),$$

and

$$\mu(B_{S^2}(x,\varepsilon)) \sim c_1\varepsilon^2, \text{ when } \varepsilon \to 0,$$

we can choose

$$k_\varepsilon(x,x) \leqslant \frac{c}{\varepsilon^2}, \text{ for a fixed constant } c.$$

For example, we can choose

$$k_\varepsilon = \begin{cases} c_\varepsilon = \frac{1}{\mu(B_{S^2}(x,\varepsilon))}, \ d_{S^2}(x,y) < \varepsilon \\ 0, \ d_{S^2}(x,y) \geqslant \varepsilon \end{cases}.$$

8.3.7 A Useful Fact

From [Letter], [L-P-S 1], and the above discussion, we have the following fact now.

Fact.

$$|V(t_\varepsilon)| \leqslant \frac{4\pi c t_\varepsilon^2}{\varepsilon^4}.$$

8.3.8 Upper Bound of $K(S)$

From [D-L], we can prove the following:

$$\frac{1}{\varepsilon} \geqslant (\frac{|V_0(t_\varepsilon)|}{4\pi c t_\varepsilon^2})^{\frac{1}{4}}$$

and

$$K(S) \leqslant 2.$$

8.4 Discussion

- Now, we have

$$\frac{4}{3} \leqslant K(S) = \overline{\lim_{\varepsilon\to 0}} \frac{\log|V(t_\varepsilon)|}{\log(\frac{1}{\mu(B_G(\varepsilon))})} \leqslant 2.$$

  Since $|V(t_\varepsilon)| = \frac{q+1}{q-1}q^{t_\varepsilon} - \frac{2}{q-1}$, we have

$$\log_q(|V(t_\varepsilon)|) = \log_q(\frac{q+1}{q-1}q^{t_\varepsilon} - \frac{2}{q-1}) = t_\varepsilon + \log_q(\frac{q+1}{q-1} - \frac{2}{(q-1)q^{t_\varepsilon}}).$$

  Hence, for $\varepsilon > 0$ and $x \in SU(2)$, there exists $\gamma \in \Gamma$ such that $d_G(x,\gamma) < \varepsilon$ and

$$h(\gamma) \leqslant t_\varepsilon \leqslant t_\varepsilon + \log_q(\frac{q+1}{q-1} - \frac{2}{(q-1)q^{t_\varepsilon}}) \leqslant 6\log_q(\frac{1}{\varepsilon}).$$

  Also, there exists some $y \in SU(2)$ such that if $\gamma \in \Gamma$ and $d_G(y,\gamma) < \varepsilon$, then

$$\frac{\log_q(V(h(\gamma)))}{\log(\frac{1}{\mu(B_G(\varepsilon))})} \geqslant \frac{4}{3},$$

$$h(\gamma) \geqslant 4\log_q(\frac{1}{\varepsilon}) - \log(\frac{1}{\mu(B_G(\varepsilon))})\log_q(\frac{q+1}{q-1} - \frac{2}{(q-1)q^{t_\varepsilon}}).$$

- From [Letter], we have the following

Theorem. *Most points y with $y \in SU(2)$ have optimally good approximation. Or equivalently, most of ρ with $\rho \in SO(3)$ have good approximation. That is, for $\varepsilon > 0$ there exists an explicit t'_ε with*

$$\frac{\log |V_0(t'_\varepsilon)|}{\log(\frac{1}{\varepsilon^2})} \to 1 \text{ as } \varepsilon \to 0,$$

and for most points $y \in SU(2)$ with respect to Haar measure, there is $\gamma \in \Gamma$ such that

$$d_G(\gamma, y) < \varepsilon \text{ and } h(\gamma) \leqslant t'_\varepsilon.$$

Remark. *In the proof of the theorem, we construct t'_ε explicitly as*

$$t'_\varepsilon = \log_q(\frac{1}{\varepsilon^{2+\varepsilon^\varepsilon}}) = (2+\varepsilon^\varepsilon)\log_q(\frac{1}{\varepsilon}),$$

The proof is in [D-L], and the following diagram gives the measure of B, which is the set of points that can not be approximated well in this case.

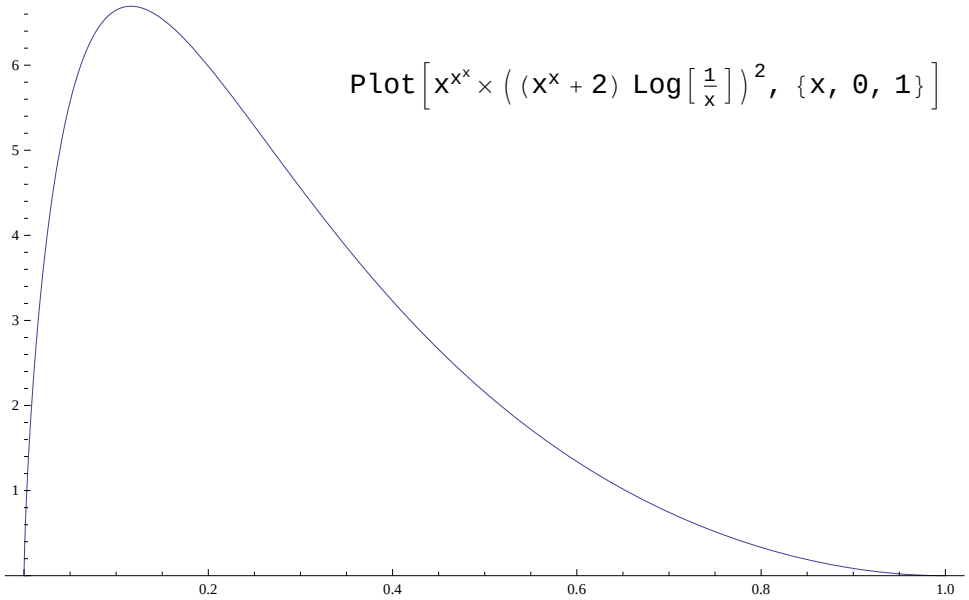

Figure 1:
Measure of B

9 An Arbitrary Set of Generators

In the last example, the universal subset we use is $S = \{s_1, s_1^{-1}, s_2, s_2^{-1}, s_3, s_3^{-1}\}$, with efficiency $\frac{4}{3} \leqslant K(S) \leqslant 2$. In order to improve this estimation, we use a larger universal subset:

$$T = \{s_1, s_1^{-1}, s_2, s_2^{-1}, s_3, s_3^{-1}, iX, iY, iZ\}.$$

Together with T and certain assumptions of covering radius on S^3 we have:

Theorem 9.1.

$$K(T) = 4/3$$

and so evidently the strict $K(T) < 2$.

See [D-L].

10 Solovay-Kitaev Alogrithm

One of the biggest challenges to creating a functional quantum computer is designing a working fault-tolerance system. A quantum computer must be able to correct environmental and hardware errors. An efficient method to control these errors is by using an algorithm to correct them. However, the computer is limited to only a finite set of gates, a library gate set. A simple example might be the Hadamard gate, the $\pi/8$-gate, and its inverse. Thus, the gates in the algorithm would need to be decomposed into these library gates. The Solvay-Kitaev (S-K) theorem and algorithm provide an effective way to decompose the gates from the correcting algorithm into those from the library set.

10.1 Solvay-Kitaev Theorem

The S-K theorem states that for a given $\epsilon > 0$ and subset G of $SU(2)$, such that the group generated by G is dense in $SU(2)$, there is a constant c such that for any $U \in SU(2)$ there exists a finite sequence S of length $O(log^c(1/\epsilon))$ of gates from G with $d(U, S) < \epsilon$.
Remark: The value of c depends on the proof of the theorem. The smallest possible value is $c = 1$; however, the proof is typically non-constructive. Thus, currently there is no algorithm that gives $c = 1$. The algorithm below from [D-N] gives a value of $c = \ln 5/\ln(3/2) \approx 3.97$. The calculations for this c are shown in section 2.3. The S-K theorem does not often provide much information on K but it is useful to construct polynomial time algorithms.

10.2 S-K Alogrithm for qubits

Here is the SK algorithm in pseudocode, See [K-M-M]:

function Solvay-Kitaev(Gate U, depth n)
if $(n == 0)$
    Return Basic Approximation to U
else
    Set U_{n-1} = Solvay-Kitaev$(U, n-1)$
    Set V, W = GC-Decompose$(UU_{n-1}^{\dagger})$
    Set V_{n-1} = Solvay-Kitaev$(V, n-1)$
    Set W_{n-1} = Solvay-Kitaev$W, n-1)$
    Return $U_n = V_{n-1}W_{n-1}V_{n-1}^{\dagger}W_{n-1}^{\dagger}U_{n-1}$

The first line

function Solvay-Kitaev(Gate U, depth n)

shows that the algorithm depends on two inputs: the single-qubit quantum gate U the algorithm is approximating and an non-negative integer n which relates to the accuracy of the approximation. The function returns a sequence of instructions which approximates U to an accuracy of ϵ_n, where ϵ_n is a decreasing function of n, i.e. $\epsilon_n \to 0$ as $n \to \infty$.

The function is recursive and terminates when n = 0,

if $(n == 0)$
    Return Basic Approximation to U

Calculations must be done before using the algorithm, so that one can find a basic ϵ_0-approximation to any $U \in SU(2)$. Since ϵ_0 is a constant, the calculation can be done by enumerating and storing a large number of sequences of gates in G of length l_0, and creating a search algorithm to find the closet approximation to U.

If n is not equal to 0, the first step is to find an ϵ_{n-1}-approximation to U.

else

Set U_{n-1} = Solvay-Kitaev$(U, n-1)$

In the next line,

Set V, W = GC-Decompose$(UU_{n-1}^\dagger)$

one must find unitary gates V and W such that $UU_{n-1}^\dagger = VWV^\dagger W^\dagger$ and $d(I,V), D(I,W) < c_{gc}\sqrt{\epsilon_{n-1}}$, where $c_{gc} \approx 1/\sqrt{2}$. This decomposition is called a balanced group commutator.
Let $M := UU_{n-1}^\dagger$. M is a rotation, by an angle θ about some axis $\hat{n}$ on the Bloch sphere. Consider ϕ satisfying,

$$\sin(\theta/2) = 2\sin^2(\phi/2)\sqrt{1-\sin^4(\phi/2)}$$
$$\phi = 2\sin^{-1}\left[\frac{\sqrt[4]{1-\cos(\theta/2)}}{\sqrt[4]{2}}\right].$$

Let $\tilde{V}$ be a rotation by ϕ about the $\hat{x}$ axis of the Bloch sphere, and $\tilde{W}$ be a rotation by ϕ about the $\hat{y}$ axis. Then, $\tilde{V}\tilde{W}\tilde{V}^\dagger\tilde{W}^\dagger$ is conjugate to M, i.e. $M = S(\tilde{V}\tilde{W}\tilde{V}^\dagger\tilde{W}^\dagger)S^\dagger$ for some unitary S.
Let $N := \tilde{V}\tilde{W}\tilde{V}^\dagger\tilde{W}^\dagger$. Because M and N are unitary, they are both diagonalizable. Furthermore, because they are conjugates, they have the same eigenvalues. Thus, there are a diagonal matrix D, and two unitary matrices S_M and S_N; such that,

$$M = S_M D S_M^\dagger$$
$$N = S_N D S_N^\dagger.$$

This implies,

$$M = S_M S_N^\dagger N S_N S_M^\dagger$$

Letting $S = S_M S_N^\dagger$, $V = S\tilde{V}S^\dagger$ and $W = S\tilde{W}S^\dagger$, V and W satisfy $UU_{n-1}^\dagger = VWV^\dagger W^\dagger$.

In the next line of the algorithm, a ϵ_{n-1}-approximation is found for V and W.

Set V_{n-1} = Solvay-Kitaev$(V, n-1)$
Set W_{n-1} = Solvay-Kitaev$W, n-1)$

In the final line,

Return $U_n = V_{n-1}W_{n-1}V_{n-1}^\dagger W_{n-1}^\dagger U_{n-1}$

Because $V_{n-1}W_{n-1}V_{n-1}^\dagger W_{n-1}^\dagger$ is an ϵ_n-approximation of $U_n U_{n-1}^\dagger$,
$V_{n-1}W_{n-1}V_{n-1}^\dagger W_{n-1}^\dagger U_{n-1} = U_n U_{n-1}^\dagger U_{n-1} = U_n$ is an ϵ_n-approximation of U_n.

10.3 Analysis

Let ϵ_n be as above, l_n be the length of the sequence returned by the algorithm, and t_n be the corresponding runtime. From the pseudocode, there are the recurrences,

$$\epsilon_n = c_{approx}\epsilon_{n-1}^{3/2}$$

$$l_n = 5l_{n-1}$$

$$t_n \leq 3t_{n-1} + const,$$

The first equation comes from the fact that $V_{n-1}W_{n-1}V_{n-1}^{\dagger}W_{n-1}^{\dagger}$ is an $c_{approx}\epsilon_{n-1}^{3/2}$-approximation of $U_nU_{n-1}^{\dagger}$. Because one wants this to be an ϵ_n-approximation, one sets $\epsilon_n = c_{approx}\epsilon_{n-1}^{3/2}$.
For the second equation, the returned sequence is made up of 5 sequences of accuracy $n-1$; thus, $l_n = 5l_{n-1}$.
Finally, because the function calls itself 3 times with an accuracy $n-1$, $t_n \leq 3t_{n-1} + const$ The constant in t_n comes from the need to compute the balanced g group commutators, etc. It does not depend on n.

The above recursions imply:

$$\begin{aligned}
\epsilon_n &= c_{approx}\epsilon_{n-1}^{3/2} \\
&= c_{approx}(c_{approx}\epsilon_{n-2}^{3/2})^{3/2} \\
&= c_{approx}(c_{approx}\dots(c_{approx}\epsilon_0^{3/2})^{3/2}\dots)^{3/2} \\
&= c_{approx}^{\sum_{j=0}^{n-1}(3/2)^n}(\epsilon)^{(3/2)^n}
\end{aligned}$$

Since $\sum_{j=0}^{n-2}(3/2)^n = \frac{1-(3/2)^2}{1-3/2} = 2((3/2)^n - 1)$,

$$\begin{aligned}
\epsilon_n &= (c_{approx}^2)^{3/2-1}(\epsilon_0)^{(\frac{3}{2})^n} \\
&= \frac{1}{c_{approx}^2}(\epsilon_0 c_{approx}^2)^{(\frac{3}{2})^n}
\end{aligned}$$

$$l_n = 5l_{n-1} = 5^2 l_{n-2} = 5^n l_0 = O(5^n)$$

$$t_n = 3t_{n-1} + const = 3^2 t_{n-2} + const' = 3^n t_0 + const'' = O(3^n)$$

To obtain a given accuracy ϵ,

$$\epsilon = \frac{1}{c_{approx}^2}(\epsilon_0 c_{approx}^2)^{(\frac{3}{2})^n}$$
$$c_{approx}^2 \epsilon = (\epsilon_0 c_{approx}^2)^{(\frac{3}{2})^n}$$
$$\ln\left(c_{approx}^2 \epsilon\right) = (\frac{3}{2})^n \ln\left(\epsilon_0 c_{approx}^2\right)$$
$$(\frac{3}{2})^n = \frac{\ln\left(c_{approx}^2 \epsilon\right)}{\ln\left(\epsilon_0 c_{approx}^2\right)}$$
$$(\frac{3}{2})^n = \frac{\ln\left(1/c_{approx}^2 \epsilon\right)}{\ln\left(1/\epsilon_0 c_{approx}^2\right)}$$
$$n \ln\frac{3}{2} = \ln\left(\frac{\ln\left(1/c_{approx}^2 \epsilon\right)}{\ln\left(1/\epsilon_0 c_{approx}^2\right)}\right)$$
$$n = \frac{\ln\left(\frac{\ln\left(1/c_{approx}^2 \epsilon\right)}{\ln\left(1/\epsilon_0 c_{approx}^2\right)}\right)}{\ln\frac{3}{2}}$$

Thus, n must satisfy,

$$n = ceil\left[\frac{\ln\left[\frac{\ln(1/\epsilon c_{approx}^2)}{\ln(1/\epsilon_0 c_{approx}^2)}\right]}{\ln(3/2)}\right].$$

Note: *ceil* means to round up the decimal number to the nearest integer.
Substituting this value into the equation of l_n gives,

$$\begin{aligned} l_n &= O\left(5^n\right) \\ &= O\left(e^{ln(5)n}\right) \\ &= O\left(\left[\frac{\ln(1/\epsilon c_{approx}^2)}{\ln(1/\epsilon_0 c_{approx}^2)}\right]^{ln5/ln(3/2)}\right) \\ &= O\left(\ln^{\ln 5/\ln(3/2)}(1/\epsilon)\right) \end{aligned}$$

Similarly,

$$t_n = O\left(\ln^{\ln 3/\ln(3/2)}(1/\epsilon)\right)$$

where $\ln 5/\ln(3/2) \approx 3.97$.

References

[B-G-S] A.Bocharov, Y.Gurevich, and K.Svore. "Efficient decomposition of single-qubit gates into V basis circuits." Physical Review A 88.1 (2013): 012313.

[D-N] C.M. Dawson, M.A. Nielsen, "The Solovay-Kitaev Algorithm", Quantum Information and Information, Vol 6, No 1 (2006), pp 081-095.

[K-M-M] V.Kliuchnikov, M.Dmitri, and M.Michele. "Practical approximation of single-qubit unitaries by single-qubit quantum Clifford and T circuits." arXiv preprint arXiv:1212.6964 (2012).

[D-L] Q. Liang and S.B. Damelin. "On a Problem of Sarnak Related to Constructing Efficient Universal Sets of Quantum Gates", arXiv 1283431.

[L-P-S 1] A.Lubotzky, R.Phillips, and P.Sarnak, "Hecke operators and distributing points on the sphere I." *Communications on Pure and Applied Mathematics* 39.S1 (1986): S149-S186.

[L-P-S 2] A.Lubotzky, R.Phillips, and P.Sarnak, "Hecke operators and distributing points on S^2 II." *Communications on Pure and Applied Mathematics* 40.4 (1987): 401-420.

[Letter] P.Sarnak, "Letter to Scott Aaronson and Andy Parlington on the Solovay-Kitaes Theorem and Golden Gates."

[T] S.Tornier. "Haar Measures." (2014).

[W] D.B.Westra. "The Haar Measure on $SU(2)$." (2008).