The Solovay-Kitaev Theorem

Dominik Roje

December 2019

Science Scholars project at University of Auckland,
supervised by André Nies

Contents

1 Preliminaries

1.1 Spectral Norm

We define the spectral norm in $SU(2)$ by

$$\|A\| = \sup_{v \in \mathbb{C}^2 \setminus \{0\}} \frac{\|Av\|}{\|v\|}$$

using the natural norm on $\mathbb{C}$. We note that equivalently, we obtain

$$\|A\| = \sigma_{\max} = \sqrt{\lambda_{\max}(A^\dagger A)}$$

where $\sigma_{\max}$ is the largest singular value of A. We also note that for a normal matrix N,

$$\|N\| = |\lambda_{\max}(N)|.$$

1.2 Universal Gate Sets

We say that a finite subset $\Gamma \subseteq SU(2)$ is a universal gate set if $\langle \Gamma \rangle$ is dense in $SU(2)$ and Γ is closed under inverses.

2 Shrinking Lemma

A key component of the proof is to be able to reduce the size of our net around the identity of $SU(2)$. By repeatedly taking the commutator of pairs in our original net, we find that we can construct arbitrarily close nets and hence approximate elements of $SU(2)$.

We note that any the Pauli matrices $\sigma_x, \sigma_y, \sigma_z$ form a basis for $\mathfrak{su}(2)$. Hence we can represent any element $A \in SU(2)$ by $A = e^{i\vec{a}\cdot\vec{\sigma}}$ where $\vec{a} \in \mathbb{R}^3$ and $\vec{\sigma}$ is a 3-vector of Pauli matrices. Without loss of generality, we can take $\|\vec{a}\| \leq \pi$.

2.1 Facts about $SU(2)$

Repeatedly, we will find ourselves reducing distances in $SU(2)$ to those in $\mathbb{R}^3$. This allows us to make use of concepts such as the cross-product and trigonometric functions. Hence we provide some facts define the relation between the two metrics.

Fact 1. If $\vec{a} \in \mathbb{R}$ then $\left\|e^{i\vec{a}\cdot\vec{\sigma}} - I\right\| = 2\sin\left(\frac{\|\vec{a}\|}{2}\right) = \|\vec{a}\| + O(\|\vec{a}\|^3)$.

Proof. Pick $\vec{a} \in \mathbb{R}$. We first wish to compute the eigenvalues of

$$\vec{a}\cdot\vec{\sigma} = \begin{bmatrix} a_z & a_z - ia_y \\ a_x + ia_y & -a_z \end{bmatrix}$$

which we can do in the traditional way.

$$\begin{aligned}
\begin{vmatrix} a_z - \lambda & a_z - ia_y \\ a_x + ia_y & -a_z - \lambda \end{vmatrix} &= -(a_z - \lambda)(a_z + \lambda) - (a_x - ia_y)(a_x + ia_y) \\
&= -(a_x^2 - \lambda^2) - (a_x^2 - a_y^2) \\
&= \lambda^2 - (a_x^2 + a_y^2 + a_z^2) \\
&= \lambda^2 - \|\vec{a}\|^2 .
\end{aligned}$$

Hence our eigenvalues are $\lambda = \pm \|\vec{a}\|$.

We now observe that any eigenvalue of $\vec{a} \cdot \vec{\sigma}$ is an eigenvalue of $e^{i\vec{a}\cdot\vec{\sigma}}$ with the same eigenvector which follows from the power series definition of matrix exponents. As $e^{i\vec{a}\cdot\vec{\sigma}}$ is also a 2×2 matrix, we have that all the eigenvalues of $e^{i\vec{a}\cdot\vec{\sigma}}$ are of the form $e^{i\lambda}$ for eigenvalues λ of $\vec{a} \cdot \vec{\sigma}$. Thus the eigenvalues of $e^{i\vec{a}\cdot\vec{\sigma}}$ are $e^{\pm i\|\vec{a}\|}$.

It then follows that the eigenvalues of $e^{i\vec{a}\cdot\vec{\sigma}} - I$ are $e^{\pm i\|\vec{a}\|} - 1$. We also know that because $e^{i\vec{a}\cdot\vec{\sigma}}$ and I commute and are both normal, that their sum is normal. Hence we have that the singular values are the absolute value of the eigenvalues. We then compute the singular values.

$$\begin{aligned}
|e^{\pm i\|\vec{a}\|} - 1| &= \sqrt{(\cos(\|\vec{a}\|) - 1)^2 + (\sin(\|\vec{a}\|))^2} \\
&= \sqrt{2 - 2\cos(\|\vec{a}\|)} \\
&= \sqrt{4\sin^2(\|\vec{a}\|/2)} \\
&= 2\sin\left(\frac{\|\vec{a}\|}{2}\right).
\end{aligned}$$

Thus $\left\|e^{i\vec{a}\cdot\vec{\sigma}} - I\right\| = 2\sin\left(\frac{\|\vec{a}\|}{2}\right)$. We proceed with the power series expansion of $\sin(x)$.

$$\begin{aligned}
2\sin\left(\frac{\|\vec{a}\|}{2}\right) &= 2\left(\frac{\|\vec{a}\|}{2} - \frac{(\|\vec{a}\|/2)^3}{3!} + \frac{(\|\vec{a}\|/2)^5}{5!} - \cdots\right) \\
&= \|\vec{a}\| - \frac{(2\|\vec{a}\|/2)^3}{3!} + \frac{(2\|\vec{a}\|/2)^5}{5!} - \cdots \\
&= \|\vec{a}\| + O(\|a\|^3).
\end{aligned}$$

□

Fact 2. If $\vec{b}, \vec{c} \in \mathbb{R}$ then $\left\|e^{i\vec{b}\cdot\vec{\sigma}} - e^{i\vec{c}\cdot\vec{\sigma}}\right\| \leq \left\|\vec{b} - \vec{c}\right\|$.

Proof. By applying Fact 1 we obtain

$$\left\|e^{i\vec{b}\cdot\vec{\sigma}} - e^{i\vec{c}\cdot\vec{\sigma}}\right\| = \left\|e^{i(\vec{b}-\vec{c})\cdot\vec{\sigma}} - I\right\| = 2\sin\left(\frac{\left\|\vec{b} - \vec{c}\right\|}{2}\right) \leq \left\|\vec{b} - \vec{c}\right\|.$$

□

Fact 3. If $\vec{b}, \vec{c} \in \mathbb{R}$ then $\left[\vec{b} \cdot \vec{\sigma}, \vec{c} \cdot \vec{\sigma}\right] = 2i(\vec{b} \times \vec{c}) \cdot \vec{\sigma}$.

Proof. We note the following properties of Pauli matrices and the cross product.

$$\begin{aligned} [\sigma_j, \sigma_k] &= 2i\varepsilon_{jkl}\sigma_l, \\ (e_j \times e_k) &= \varepsilon_{jkl} e_l. \end{aligned}$$

Consider an arbitrary component of our commutator:

$$\begin{aligned} \left[\vec{b} \cdot \vec{\sigma}, \vec{c} \cdot \vec{\sigma}\right]_l &= b_j c_k \left[\sigma_j, \sigma_k\right] + b_k c_j \left[\sigma_k, \sigma_j\right] \\ &= b_j c_k \left[\sigma_j, \sigma_k\right] - b_k c_j \left[\sigma_j, \sigma_k\right] \\ &= (b_j c_k - b_k c_j) \left[\sigma_k, \sigma_j\right] \\ &= 2i(b_j c_k - b_k c_j)\varepsilon_{jkl}\sigma_l \\ &= 2i(\vec{b} \times \vec{c})_l \sigma_l. \end{aligned}$$

Thus we have our result.

□

Fact 4. If $\vec{b}, \vec{c} \in \mathbb{R}$ with $\left\|\vec{b}\right\| = O(\varepsilon)$ and $\|\vec{c}\| = O(\varepsilon)$ then $\left\|[\![e^{i\vec{b}\cdot\vec{\sigma}}, e^{i\vec{c}\cdot\vec{\sigma}}]\!] - e^{-\left[\vec{b}\cdot\vec{\sigma}, \vec{c}\cdot\vec{\sigma}\right]}\right\| = O(\epsilon^3)$.

Proof. Follows from Ozols' proof [1].

□

Lemma 1. (*Shrinking Lemma*) If Γ is an ε^2-net for S_ε then $[\![\Gamma, \Gamma]\!]$ is an $O(\varepsilon^3)$-net for S_{ε^2}.

Proof. Suppose Γ is an ε^2 net for S_ε. We want to show for any $A \in S_{\varepsilon^2}$, there exist $U, V \in \Gamma$ such that $\|A - [\![U, V]\!]\| = O(\varepsilon^3)$.

So let us pick $A \in S_{\varepsilon^2}$ and corresponding $\vec{a} \in \mathbb{R}^3$ such that $\|\vec{a}\| \leq \pi$ and $A = e^{i\vec{a}\cdot\vec{\sigma}}$. Now we choose $\vec{b}, \vec{c} \in \mathbb{R}^3$ such that $-2\vec{b} \times \vec{c} = \vec{a}$ and $\left\|\vec{b}\right\| = \|\vec{c}\| = \sqrt{\|a\|/2}$. Define $B = e^{i\vec{b}\cdot\vec{\sigma}}$ and $C = e^{i\vec{c}\cdot\vec{\sigma}}$.

Note that because Γ is an ε^2-net, that $\|A - I\| = \|\vec{a}\| + O(\|\vec{a}\|^2) \leq \varepsilon^2$ and hence $\|\vec{a}\| = O(\varepsilon^2)$. By construction we then obtain $\left\|\vec{b}\right\| = O(\varepsilon)$ and $\|\vec{c}\| = O(\varepsilon)$. We then invoke Fact 4 to obtain

$$\|[\![B, C]\!] - A\| = \left\|[\![e^{i\vec{b}\cdot\vec{\sigma}}, e^{i\vec{c}\cdot\vec{\sigma}}]\!] - e^{-\left[\vec{b}\cdot\vec{\sigma}, \vec{c}\cdot\vec{\sigma}\right]}\right\| = O(\epsilon^3).$$

We now take $U = e^{i\vec{u}\cdot\vec{\sigma}}, V = e^{i\vec{v}\cdot\vec{\sigma}} \in \Gamma$ to be the closest elements to B and C respectively, i.e. $\|B - U\| \leq \varepsilon^2$ and $\|C - V\| \leq \varepsilon^2$. We can now use Fact 1 to find

$$\|B - U\| = \left\|BU^\dagger - I\right\| = \left\|\vec{b} - \vec{u}\right\| + O\left(\left\|\vec{b} - \vec{u}\right\|^3\right) \leq \varepsilon^2.$$

Therefore $\left\|\vec{b}-\vec{u}\right\| = O(\varepsilon^2)$ and similarly $\|\vec{c}-\vec{v}\| = O(\varepsilon^2)$. We then obtain the following relation by the triangle inequality

$$\|A - [\![U,V]\!]\| \leq \left\|A - e^{2i(\vec{b}\times\vec{c})\cdot\sigma}\right\| + \left\|e^{2i(\vec{b}\times\vec{c})\cdot\sigma} - [\![U,V]\!]\right\|.$$

The right term is $O(\varepsilon^3)$ by Fact 4 so we consider the left term.

$$\begin{aligned}
\left\|A - e^{2i(\vec{b}\times\vec{c})\cdot\sigma}\right\| &\leq 2\left\|\vec{b}\times\vec{c} - \vec{u}\times\vec{v}\right\| \\
&= 2\left\|((\vec{b}-\vec{u})+\vec{u})\times((\vec{c}-\vec{v})+\vec{v}) - \vec{u}\times\vec{v}\right\| \\
&= 2\left\|(\vec{b}-\vec{u})\times(\vec{c}-\vec{v}) + \vec{u}\times(\vec{c}-\vec{v}) + (\vec{b}-\vec{u})\times\vec{v}\right\| \\
&= O(\varepsilon^4) + O(\varepsilon^3) + O(\varepsilon^3) \\
&= O(\varepsilon^3)
\end{aligned}$$

Hence we have

$$\|A - [\![U,V]\!]\| = O(\varepsilon^3).$$

□

Lemma 2. Let $\varepsilon > 0$ sufficiently small. There exists $k \in \mathbb{R}$ such that if Γ is an ε^2-net for S_ε then $[\![\Gamma,\Gamma]\!]\Gamma$ is a $k^2\varepsilon^3$-net for $S_{k\varepsilon^{3/2}}$.

Proof. By Lemma 1 There exists $k \in \mathbb{R}$ such that $[\![\Gamma,\Gamma]\!]$ is a $k^2\varepsilon^3$-net for S_{ε^2}. Now pick $A \in S_{k\varepsilon^{3/2}}$. For sufficiently small ε we have $S_{k\varepsilon^{3/2}} \subset S_{\varepsilon^2}$ so we know there exists $W \in \Gamma$ such that $\left\|AW^\dagger - I\right\| \leq \varepsilon^2$. Thus we have $AW^\dagger \in S_{\varepsilon^2}$ and hence we have $U, V \in \Gamma$ such that $\left\|AW^\dagger - [\![U,V]\!]\right\| \leq k^2\varepsilon^3$.

□

Having established Lemma 2 to reduce the size of our net, we apply the idea inductively. Establishing a means to create arbitrarily small nets around the identity. We formalise this in the following corollary.

Corollary 1. If Γ_0 is an ε_0^2-net for S_ε ε_0 sufficiently small. Then $\Gamma_i = [\![\Gamma_{i-1},\Gamma_{i-1}]\!]\Gamma_{i-1}$ is a ε_i^2-net where $\varepsilon_i = (k^2\varepsilon_0)^{(3/2)^i}/k^2$ for some $k \in \mathbb{R}$.

Since each element of Γ_i is composed of gates in Γ_0, we effectively construct a sequence of gates; each taking us closer to the identity.

3 Solovay-Kitaev Theorem

Theorem 1. If Γ is a universal gate set that is closed under inverses. Then we can approximate any $U \in SU(2)$ to any accuracy $\varepsilon > 0$ by a sequence of gates in Γ.

Proof. Pick ε_0 sufficiently small and independent of Γ. We wish to construct Γ_0 - an ε_0^2-net for $SU(2)$. Since $\langle\Gamma\rangle$ is dense, we can take the ε_0^2 neighbourhoods of points in $\langle\Gamma\rangle$ which form a cover for $SU(2)$. We now note that $SU(2)$ is compact and hence has a finite subcover. Taking the centers of the subcover and their inverses we form Γ_0.

Pick $U \in SU(2)$ and find $V_0 \in \Gamma$ such that $\|U - V_0\| = \left\|UV_0^\dagger - I\right\| \leq \varepsilon_0^2$. We then have that $UV_0^\dagger \in S_{\varepsilon_0^2}$. For sufficiently small ε_0 we have $\varepsilon_0^2 < k\varepsilon_0^{3/2} = \varepsilon_1$ for the $k \in \mathbb{R}$ given by Lemma 1. Thus by Corollary 1, we have that Γ_1 is an ε_1^2-net for S_{ε_1} and we can find $V_1 \in \Gamma_1$ such that

$$\left\|UV_0^\dagger V_1^\dagger - I\right\| = \|U - V_1V_0\| \leq \varepsilon_1^2 < k\varepsilon_1^{3/2} = \varepsilon_2.$$

Proceeding inductively, we can find $V_t \in \Gamma_t$ with $\|U - V_t \cdots V_0\| \leq \varepsilon_t^2$. We note that each V_i is composed of 5^i gates. And hence we need $\sum_{i=0}^{t} 5^i = O(5^t)$ gates in Γ_0. Moreover, for an accuracy ε we need $\varepsilon_t^2 = ((k^2\varepsilon_0)^{(3/2)^t}/k^2)^2 \leq \varepsilon$ and solve for t. Seeing as $\frac{3}{2}^{(\log(5)/\log(3/2))} = 5$. We set $c = \frac{\log(5)}{\log(3/2)}$ and we find,

$$\begin{aligned}
((k^2\varepsilon_0)^{(3/2)^t})^2 &\leq k^4\varepsilon \\
\left(\frac{3}{2}\right)^t \ln(k^2\varepsilon_0) &\geq \frac{1}{2}\ln(k^4\varepsilon) \\
\left(\frac{3}{2}\right)^t &\leq \frac{\log(k^4\varepsilon)}{2\log(k^2\varepsilon_0)} \\
\left(\frac{3}{2}\right)^t &\leq \frac{\log(1/k^4\varepsilon)}{2\log(1/k^2\varepsilon_0)} \\
5^t &\leq \frac{\log^c(1/k^4\varepsilon)}{2\log^c(1/k^2\varepsilon_0)}
\end{aligned}$$

Hence we have $O(5^t) = O(\log^c(1/\varepsilon))$.

□

4 Results in Higher Dimensions

As our results hold only for gates in $SU(2)$, and therefore for single qubit systems, we wish to extend the proof for arbitrary qudits. I.e, a proof that holds for $SU(d)$.

Such a result is given in a paper by Dawson and Nielsen [2]. Where an ε-approximate can be made with the paradoxically better result of $O(\log^{2.71}(1/\varepsilon))$ gates because of some redundancy that occurs. If one chooses to ignore this redundancy, a similar proof to the above can be written, giving the familiar bound of $O(\log^{3.97}(1/\varepsilon))$.

4.1 Facts about $SU(d)$

The proof given in this report generalises to $SU(d)$ to the extent that the identities given change slightly but the idea of the proof is the same. For example Fact 1 in $SU(d)$ becomes:

Fact 5. If H is hermitian then

$$\left\|e^{iH} - I\right\| = 2\sin\left(\frac{\|H\|}{2}\right) = \|H\| + O(\|H\|^3).$$

5 Contains an IRREP

In our definition of universal gate set, we insisted that it be closed under inverses. A paper by Bouland and Ozols [3] proves that one can instead insist that our set Γ is instead contains a projective irreducible representation of some finite group G. To approximate any $U \in SU(d)$ to accuracy ε, one only needs $O(\log^{\log_2 |G|}(1/\varepsilon))$ gates in Γ.

5.1 Representations of Groups

An representation of a group G over vector space $\mathbb{C}^n$ is a homomorphism $\sigma : G \to \mathrm{GL}(\mathbb{C}^n)$. Given two representations $\sigma : G \to V$ and $\sigma' \to V'$ we define their direct sum $\sigma \oplus \sigma' : G \to V \oplus V'$ by

$$(\sigma \oplus \sigma')(x) = \begin{pmatrix} \sigma(x) & 0 \\ 0 & \sigma'(x) \end{pmatrix}$$

for all $x \in G$. A representation is then irreducible if it is not the direct sum of two other representations. More specifically, this paper makes use of representations $\sigma : G \to U(d)$ and calls σ projective if there is some function $\theta : G \times G \to \mathbb{R}$ such that for all $g_1, g_2 \in G$ $\sigma(g_1)\sigma(g_2) = e^{i\theta(g_1,g_2)}\sigma(g_1, g_2)$. i.e, σ preserves global phase.

The theorem then becomes

Theorem 2. For any fixed $d \geq 2$, suppose $\Gamma \subset SU(d)$is a finite gate set which densely generates $SU(d)$, and furthermore Γ contains a (projective) irrep of some finite group G. Then there is an algorithm which outputs an ε-approximation to any $U \in SU(d)$ using merely $O(\mathrm{polylog}(1/\varepsilon))$ elements from Γ.

References

[1] Maris Ozols. The Solovay-Kitaev theorem. page 9.

[2] Christopher M. Dawson and Michael A. Nielsen. The Solovay-Kitaev algorithm. *arXiv:quant-ph/0505030*, May 2005. arXiv: quant-ph/0505030.

[3] Adam Bouland and Maris Ozols. Trading inverses for an irrep in the Solovay-Kitaev theorem. *arXiv:1712.09798 [math-ph, physics:quant-ph]*, page 15 pages, 2018. arXiv: 1712.09798.