Optimal ancilla-free Clifford+T approximation of z-rotations

Neil J. Ross and Peter Selinger
Department of Mathematics and Statistics
Dalhousie University

Abstract

We consider the problem of approximating arbitrary single-qubit z-rotations by ancilla-free Clifford+T circuits, up to given epsilon. We present a fast new probabilistic algorithm for solving this problem optimally, i.e., for finding the shortest possible circuit whatsoever for the given problem instance. The algorithm requires a factoring oracle (such as a quantum computer). Even in the absence of a factoring oracle, the algorithm is still near-optimal under a mild number-theoretic hypothesis. In this case, the algorithm finds a solution of T-count $m + O(\log(\log(1/\varepsilon)))$, where m is the T-count of the second-to-optimal solution. In the typical case, this yields circuit approximations of T-count $3\log_2(1/\varepsilon) + O(\log(\log(1/\varepsilon)))$. Our algorithm is efficient in practice, and provably efficient under the above-mentioned number-theoretic hypothesis, in the sense that its expected runtime is $O(\text{polylog}(1/\varepsilon))$.

1 Introduction

Practical quantum computing requires the fault-tolerant implementation of a universal gate set. The decomposition of arbitrary unitary operators into gates from this fixed set is then an important problem. Most of the common error correction schemes, including most stabilizer codes and surface codes, permit a relatively inexpensive fault-tolerant implementation of gates from the Clifford group. However, since Clifford gates are not universal for quantum computation, at least one non-Clifford gate must be added to the basic gate set to achieve universality. A common choice for this additional gate is the T-gate or $\pi/8$-gate. The T-gate is not the only possible extension of the Clifford group but it is considered to be the most practical one. This is due to the availability of fault-tolerant implementations of the T-gate. For this reason, the Clifford+T gate set is often considered as the most promising candidate for practical quantum computing.

In this paper, we consider the problem of approximating arbitrary single-qubit z-rotations by Clifford+T circuits up to given ε. Until about two years ago, the state-of-the-art algorithm for this problem was the Solovay-Kitaev algorithm, which yields circuits of size $O(\log^c(1/\varepsilon))$, where $c > 3$. At the other end of the spectrum are algorithms based on exhaustive search. While such algorithms achieve optimal circuit sizes, they have exponential runtimes. For example, the algorithm of [5] is feasible up to $\varepsilon \approx 10^{-4}$. Even the improved algorithm of [11], which combines search-based and other methods and achieves optimal T-counts, still has exponential runtime which makes it feasible only for precisions up to $\varepsilon \approx 10^{-17}$.

Within the last two years, a new generation of efficient number-theoretic algorithms have been proposed for the approximate synthesis problem, achieving circuit sizes of $O(\log(1/\varepsilon))$ with polynomial runtime. Unlike the Solovay-Kitaev algorithm, which is based on a geometric method of successive approximations, these new algorithms are based on solving Diophantine equations. The first such algorithm was due to Kliuchnikov, Maslov, and Mosca [9]. It uses a small number of ancilla qubits to approximate a given single-qubit operator. An improved algorithm was given in [14], which uses no ancillas and achieves T-counts of $K + 4\log_2(1/\varepsilon)$ for approximating arbitrary z-rotations. This compares to the information-theoretic lower bound of $K + 3\log_2(1/\varepsilon)$.

In this paper, we present a fast new probabilistic algorithm for solving the single-qubit approximate synthesis problem. Our algorithm is optimal in an absolute sense, i.e., it finds the shortest possible circuit whatsoever for any given problem instance. To achieve this optimality, the algorithm requires an oracle for integer factoring. Of course, a quantum computer can serve as such an oracle by using Shor's algorithm [15]. But even in the absence of a factoring oracle, our algorithm is still nearly optimal: In this case, under a mild number-theoretic assumption, we can prove that the algorithm finds a solution of T-count $m + O(\log(\log(1/\varepsilon)))$, where m is the T-count of the second-to-optimal solution. In the typical case, m is given by the information-theoretic lower bound $3\log_2(1/\varepsilon)$. Therefore our algorithm, in the absence of a factoring oracle, yields circuit approximations of T-count $3\log_2(1/\varepsilon) + O(\log(\log(1/\varepsilon)))$ in the typical case.

We note that our algorithm is optimal only for the *specific* problem of approximating a given z-rotation by a linear sequence of single-qubit Clifford+T gates. It is already known that even smaller gate counts and/or circuit depths are achievable using additional techniques, such as ancillas, measurements, or state distillation [4, 16, 3, 2]. In particular, the methods of [3] produce so-called repeat-until-success circuits whose expected T-count is below the information theoretic lower bound for deterministic circuits.

It is likely that in the future, the existence of an efficient approximate synthesis algorithm will be considered an essential requirement for any universal gate set proposed for practical quantum computing, of similar importance, say, as the existence of a fault-tolerant implementation. Our algorithm is specialized to the Clifford+T gate set which is arguably the most relevant gate set to practical quantum computing. However, similar number-theoretic methods are also applicable to certain other universal gate sets. For example, Bocharov et al. [1] gave an efficient synthesis algorithm for the Clifford+V gate set that achieves gate counts linear in $\log(1/\varepsilon)$, and Kliuchnikov et al. [8] did the same for the $\langle \mathcal{F}, \mathcal{T} \rangle$ gate set. One may reasonably expect these gate sets to be amenable to the same kind of optimal synthesis that we provide here for the Clifford+T gate set.

2 Overview

Recall that the single-qubit Clifford group is generated by the Hadamard gate H, the phase gate S, and the scalar $\omega = e^{i\pi/4}$. By adding the non-Clifford operator T, one obtains a universal gate set for quantum computing.

$$\omega = e^{i\pi/4}, \quad H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix}.$$

Our goal is to approximate an arbitrary z-rotation

$$R_z(\theta) = e^{-i\theta Z/2} = \begin{pmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{pmatrix}$$

by a Clifford+T operator up to given $\varepsilon > 0$. By a result of Kliuchnikov, Maslov, and Mosca [10], a unitary 2×2-operator can be exactly written as a product of Clifford+T operators if and only if all of its matrix entries belong to the ring $\mathbb{D}[\omega] = \mathbb{Z}[\frac{1}{\sqrt{2}}, i]$. Our strategy is therefore to approximate $R_z(\theta)$ by a unitary operator of the form

$$U = \begin{pmatrix} u & -t^\dagger \\ t & u^\dagger \end{pmatrix},$$

where $u, t \in \mathbb{D}[\omega]$. This problem can be solved in two stages:

(1) find a suitable candidate $u \in \mathbb{D}[\omega]$ that is a good approximation of $e^{-i\theta/2}$;

(2) solve the Diophantine equation $u^\dagger u + t^\dagger t = 1$, to ensure that U is unitary.

Problem (2) can be solved by standard number-theoretic methods. In the interest of self-containedness, we summarize these methods in Section 6 and Appendix C. In general, solving the Diophantine equation in (2) requires the ability to factor large integers. However, if no efficient factoring method is available, then (modulo a mild number-theoretic assumption), the Diophantine equation can still be solved with large enough probability to ensure that at most $O(\log(1/\varepsilon))$ candidates need to be tried.

The main new technical innovation of this paper is a new and optimal solution to problem (1). It turns out that the "suitability" of a candidate u can be expressed as a problem of the form $u \in A$ and $u^\bullet \in B$, where A and B are fixed convex subsets of the complex plane depending only on θ and ε, and $(-)^\bullet$ is the automorphism of the ring $\mathbb{D}[\omega]$ obtained by mapping $\sqrt{2}$ to $-\sqrt{2}$. We call such a problem a *two-dimensional grid problem.* In Sections 4 and 5, we formulate a general algorithm for solving one- and two-dimensional grid problems efficiently. Our algorithm proceeds by transforming the sets A and B using invertible linear operators that preserve solutions to the given problem, which we call *grid operators.* We show that for any two convex sets A and B one can construct a grid operator G such that the action of G on A and B normalizes the grid problem in an appropriate sense. The main technical ingredient that makes our solution efficient is an iterative process for constructing G, which is detailed in Appendix A.

The rest of this paper is organized as follows. In Section 3, we review some basic notions from algebra. We discuss one- and two-dimensional grid problems in Sections 4 and 5, respectively. In Section 6, we show how to solve the relevant Diophantine equation. A detailed description of the main synthesis algorithm is given in Section 7. In Section 8, we analyze the algorithm's correctness, optimality, and complexity. Some experimental results are given in Section 10. For better readability of the main body of the paper, certain technical results are proved in the appendices.

3 Some algebra

We introduce some notation and algebraic prerequisites. The set of natural numbers, including 0, is denoted by $\mathbb{N}$, the ring of integers is denoted by $\mathbb{Z}$, and we let $\omega = e^{i\pi/4} = (1+i)/\sqrt{2}$.

Definition 3.1. *(Extensions of $\mathbb{Z}$)* We are interested in the following rings of algebraic integers:

- $\mathbb{Z}[\sqrt{2}] = \{a + b\sqrt{2} \mid a, b \in \mathbb{Z}\}$, the ring of *quadratic integers with radicand 2*;
- $\mathbb{Z}[\omega] = \{a\omega^3 + b\omega^2 + c\omega + d \mid a, b, c, d \in \mathbb{Z}\}$, the ring of *cyclotomic integers of degree 8*;
- $\mathbb{D} = \mathbb{Z}[\frac{1}{2}] = \{\frac{a}{2^k} \mid a \in \mathbb{Z}, k \in \mathbb{N}\}$, the ring of *dyadic fractions*;
- $\mathbb{D}[\sqrt{2}] = \mathbb{Z}[\frac{1}{\sqrt{2}}] = \{a + b\sqrt{2} \mid a, b \in \mathbb{D}\}$; and
- $\mathbb{D}[\omega] = \mathbb{Z}[\frac{1}{\sqrt{2}}, i] = \{a\omega^3 + b\omega^2 + c\omega + d \mid a, b, c, d \in \mathbb{D}\}$.

We have the inclusions $\mathbb{Z} \subseteq \mathbb{Z}[\sqrt{2}] \subseteq \mathbb{Z}[\omega]$ and $\mathbb{D} \subseteq \mathbb{D}[\sqrt{2}] \subseteq \mathbb{D}[\omega]$. Moreover, $\mathbb{Z} \subseteq \mathbb{D}$, $\mathbb{Z}[\sqrt{2}] \subseteq \mathbb{D}[\sqrt{2}]$, and $\mathbb{Z}[\omega] \subseteq \mathbb{D}[\omega]$. Finally, $\mathbb{Z}[\sqrt{2}]$ and $\mathbb{Z}[\omega]$ are dense in $\mathbb{R}$ and $\mathbb{C}$, respectively.

Definition 3.2. *(Automorphisms)* The following maps are automorphisms of $\mathbb{D}[\omega]$:

- *Complex conjugation*, which we denote $(-)^\dagger$, acts on an arbitrary element of $\mathbb{D}[\omega]$ or $\mathbb{Z}[\omega]$ as follows:
$$(a\omega^3 + b\omega^2 + c\omega + d)^\dagger = -c\omega^3 - b\omega^2 - a\omega + d.$$
- $\sqrt{2}$*-conjugation*, which we denote $(-)^\bullet$, acts on an arbitrary element of $\mathbb{D}[\omega]$ or $\mathbb{Z}[\omega]$ as follows:
$$(a\omega^3 + b\omega^2 + c\omega + d)^\bullet = -a\omega^3 + b\omega^2 - c\omega + d$$

The action of $(-)^\bullet$ on an element of $\mathbb{D}[\sqrt{2}]$ or $\mathbb{Z}[\sqrt{2}]$ is given by $(a + b\sqrt{2})^\bullet = a - b\sqrt{2}$. In particular, this implies that if $t = a + b\sqrt{2}$ is an element of $\mathbb{D}[\sqrt{2}]$ (resp. $\mathbb{Z}[\sqrt{2}]$), then $t^\bullet t = a^2 - 2b^2$ is an element of $\mathbb{D}$ (resp. $\mathbb{Z}$).

Remark 3.3. If α and β are two distinct elements of $\mathbb{Z}[\sqrt{2}]$, then the following inequality holds:

$$|\alpha - \beta| \cdot |\alpha^\bullet - \beta^\bullet| \geqslant 1, \tag{1}$$

This follows from the fact that for $t \in \mathbb{Z}[\sqrt{2}]$, $t^\bullet t$ is an integer and $t^\bullet t = 0$ if and only if $t = 0$.

Definition 3.4. Let $t \in \mathbb{D}[\omega]$ and $k \in \mathbb{N}$. If $\sqrt{2}^k t \in \mathbb{Z}[\omega]$, then we say that k is a *denominator exponent* of t. The smallest such $k \geqslant 0$ is the *least denominator exponent* of t. For $k \in \mathbb{N}$, the elements of $\mathbb{D}[\sqrt{2}]$ (resp. $\mathbb{D}[\omega]$) having k as a denominator exponent form a ring, denoted $\frac{1}{\sqrt{2}^k}\mathbb{Z}[\sqrt{2}]$ (resp. $\frac{1}{\sqrt{2}^k}\mathbb{Z}[\omega]$).

Definition 3.5. We frequently refer to the following elements of $\mathbb{Z}[\sqrt{2}]$ and $\mathbb{Z}[\omega]$:

- $\lambda = 1 + \sqrt{2} \in \mathbb{Z}[\sqrt{2}]$ and
- $\delta = 1 + \omega \in \mathbb{Z}[\omega]$.

Remark 3.6. The element λ is invertible in the ring $\mathbb{Z}[\sqrt{2}]$, with inverse $\lambda^{-1} = -1 + \sqrt{2} = -\lambda^\bullet$. The element δ of the ring $\mathbb{Z}[\omega]$ satisfies $\delta^2 = \lambda\omega\sqrt{2}$ and $\delta^\dagger\delta = \lambda\sqrt{2}$.

4 One-dimensional grid problems

Definition 4.1. Let B be a set of real numbers. The *(real) grid* for B is the set

$$\mathrm{grid}(B) = \{\alpha \in \mathbb{Z}[\sqrt{2}] \mid \alpha^\bullet \in B\}. \tag{2}$$

When B is clear from the context, we refer to the elements of this set as *grid points*.

In the following, we will only be interested in the case where B is a closed interval $[y_0, y_1]$ with $y_0 < y_1$. In this case, the grid is discrete and infinite. It is discrete because the distance between grid points is bounded below by (1). And it is infinite by the density of $\mathbb{Z}[\sqrt{2}]$: there are infinitely many points $\beta \in B \cap \mathbb{Z}[\sqrt{2}]$, and for each of them, $\beta^\bullet$ is a grid point.

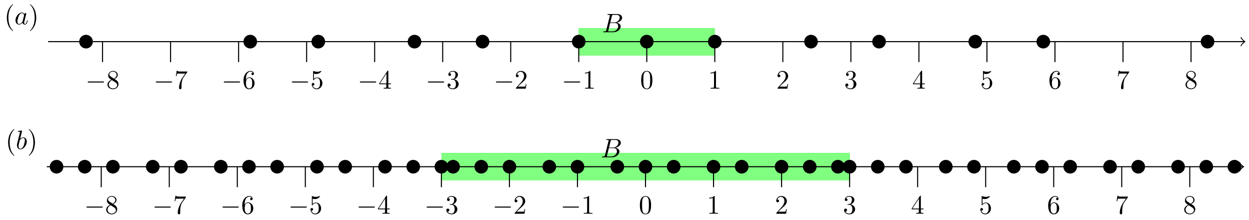


Figure 1: The real grid for two different intervals B. In both cases, the interval B is shown in green, and grid points are shown as black dots.

Example 4.2. Figure 1 illustrates the grids for the intervals $[-1,1]$ and $[-3,3]$, respectively. For example, the first few non-negative points in $\mathrm{grid}([-1,1])$ are 0, 1, $1+\sqrt{2}$, $2+\sqrt{2}$, $2+2\sqrt{2}$, $3+2\sqrt{2}$, and $4+3\sqrt{2}$. As one would expect, the grid for $[-3,3]$ is about three times denser than that for $[-1,1]$. We also note that $B \subseteq B'$ implies $\mathrm{grid}(B) \subseteq \mathrm{grid}(B')$.

Definition 4.3. Let A and B be sets of real numbers. The *one-dimensional grid problem* for A and B is the following:

$$\textbf{One-dimensional grid problem: } \text{Find } \alpha \in \mathbb{Z}[\sqrt{2}] \text{ satisfying } \alpha \in A \text{ and } \alpha^\bullet \in B. \tag{3}$$

Note that (3) can be equivalently written as $\alpha \in A \cap \mathrm{grid}(B)$. In other words, the grid problem is to find points in some given set A that belong to the grid for B. We also refer to the conditions $\alpha \in A$ and $\alpha^\bullet \in B$ as *grid constraints*.

In the case where A and B are finite intervals, the grid problem is guaranteed to have a finite number of solutions. We recall the following facts from [14]:

Lemma 4.4. *Let $A = [x_0, x_1]$ and $B = [y_0, y_1]$ be closed real intervals, such that $x_1 - x_0 = \delta$ and $y_1 - y_0 = \Delta$. If $\delta\Delta < 1$, then the grid problem (3) has at most one solution. If $\delta\Delta \geqslant (1+\sqrt{2})^2$, then the grid problem (3) has at least one solution.*

Proof. Lemmas 16 and 17 of [14]. □

Proposition 4.5. *There is an algorithm for enumerating all solutions of the one-dimensional grid problem for closed intervals $A = [x_0, x_1]$ and $B = [y_0, y_1]$. Moreover, the algorithm is efficient in the sense that it only requires a constant number of arithmetic operations per solution produced.*

Proof. It was already noted in [14, Lemma 17] that there is an efficient algorithm for computing one solution. To see that we can efficiently enumerate all solutions, let $\delta = x_1 - x_0$ and $\Delta = y_1 - y_0$ as before. Recall that $\lambda = 1 + \sqrt{2}$ and that $\lambda^{-1} = -\lambda^\bullet$. The grid problem for the sets A and B is equivalent to the grid problem for $\lambda^{-1}A$ and $-\lambda B$, because $\alpha \in A$ and $\alpha^\bullet \in B$ hold if and only if $\lambda^{-1}\alpha \in \lambda^{-1}A$ and $(\lambda^{-1}\alpha)^\bullet \in -\lambda B$. Using such rescaling, we may without loss of generality assume that $\lambda^{-1} \leqslant \delta < 1$.

Now consider any solution $\alpha = a + b\sqrt{2} \in \mathbb{Z}[\sqrt{2}]$. From $\alpha \in [x_0, x_1]$, we know that $x_0 - b\sqrt{2} \leqslant a \leqslant x_1 - b\sqrt{2}$. But since $x_1 - x_0 < 1$, it follows that for any $b \in \mathbb{Z}$, there is at most one $a \in \mathbb{Z}$ yielding a solution. Moreover, we note that $b = (\alpha - \alpha^\bullet)/\sqrt{2}^3$, so that any solution satisfies $(x_0 - y_1)/\sqrt{2}^3 \leqslant b \leqslant (x_1 - y_0)/\sqrt{2}^3$. The algorithm then proceeds by enumerating all the integers b in the interval $[(x_0 - y_1)/\sqrt{2}^3, (x_1 - y_0)/\sqrt{2}^3]$. For each such b, find the unique integer a (if any) in the interval $[x_0 - b\sqrt{2}, x_1 - b\sqrt{2}]$. Finally, check if $a + b\sqrt{2}$ is a solution. The runtime is governed by the number of $b \in \mathbb{Z}$ that need to be checked, of which there are at most $O(y_1 - y_0) = O(\delta\Delta)$. As a consequence of Lemma 4.4, the total number of solutions is at least $\Omega(\delta\Delta)$, and so the algorithm is efficient. □

Remark 4.6. For the purposes of this paper, by an *arithmetic operation* we mean addition, subtraction, multiplication, division, exponentiation, and logarithm.

Remark 4.7. Since the inputs to the algorithm are real intervals, if we were to give a rigorous complexity-theoretic account, we should also clarify how these intervals are specified (for example, with rational endpoints, endpoints as computable real numbers, etc.). For our purposes, the manner in which an interval $A = [x_0, x_1]$ is specified as an input to the algorithm does not matter very much; it would be sufficient, for example, to assume that we are given rational bounds a, b with $a < x_0 < x_1 < b$, such that $b - a$ exceeds $x_1 - x_0$ by at most a fixed constant factor, as well as a procedure for deciding whether any given point of $\mathbb{D}[\sqrt{2}]$ is in A or not.

5 Two-dimensional grid problems

Recall that $\mathbb{Z}[\omega]$ is a subset of the complex numbers. In what follows, it is often convenient to identify the complex numbers with the Euclidean plane $\mathbb{R}^2$, so we will often interchangeably regard $\mathbb{Z}[\omega]$ as a subset of $\mathbb{C}$ and of $\mathbb{R}^2$.

Definition 5.1. Let B be a subset of $\mathbb{R}^2$. The *(complex) grid* for B is the set

$$\mathrm{Grid}(B) = \{u \in \mathbb{Z}[\omega] \mid u^\bullet \in B\}. \tag{4}$$

We will only be interested in the case where B is a bounded convex set with non-empty interior. In this case, the grid is discrete and infinite, just as in the one-dimensional case.

Example 5.2. Figure 2 illustrates the complex grids for several different convex sets B. Note that the grid has a 90-degree symmetry in (a), a 45-degree symmetry in (b), and a 180-degree symmetry in (c).

Definition 5.3. Let A and B be subsets of $\mathbb{R}^2$. The *two-dimensional grid problem* for A and B is the following:

$$\textbf{Two-dimensional grid problem: } \text{Find } u \in \mathbb{Z}[\omega] \text{ satisfying } u \in A \text{ and } u^\bullet \in B. \tag{5}$$

As in the one-dimensional case, the grid problem can be understood as looking for points in the intersection of the set A and the grid for B. Our goal will be to prove a two-dimensional analogue of Proposition 4.5, namely, that there is an efficient algorithm which, given two bounded convex subsets A and B of $\mathbb{R}^2$ with non-empty interior, enumerates all solutions of the two-dimensional grid problem for A and B.

However, the proof is more complicated than in the one-dimensional case. We will consider several special cases before solving the general problem in Section 5.6.

Remark 5.4. By analogy with Remark 4.7, we should indicate what it means for a bounded convex subset A of $\mathbb{R}^2$ to be "given" as the input to an algorithm. Again, the details of this do not matter much. For our purposes, it will suffice to assume that a convex set is given along with the following information:

- a convex polygon enclosing A, say with rational vertices, and such that the area of the polygon exceeds that of A by at most a fixed constant factor;
- a method to decide, for any given point of $\mathbb{D}[\omega]$, whether it is in A or not; and
- a method to compute the intersection of A with any straight line in $\mathbb{D}[\omega]$. More precisely, given any straight line parameterized as $L(t) = p + tq$, with $p, q \in \mathbb{D}[\omega]$, we can effectively determine the interval $\{t \mid L(t) \in A\}$ in the sense of Remark 4.7.

5.1 Upright rectangles

A special case of the two-dimensional grid problem arises when both A and B are *upright rectangles*, by which we mean sets of the form $[x_0, x_1] \times [y_0, y_1]$. If A and B are upright rectangles, then the two-dimensional grid problem can easily be reduced to the one-dimensional case. We start with a lemma characterizing $\mathbb{Z}[\omega]$.

Lemma 5.5. *A complex number u is in $\mathbb{Z}[\omega]$ if and only it can be written of the form $u = \alpha + \beta i$ or of the form $u = \alpha + \beta i + \omega$, where $\alpha, \beta \in \mathbb{Z}[\sqrt{2}]$.*

Proof. The right-to-left implication is trivial. For the left-to-right implication, let $u = a\omega^3 + b\omega^2 + c\omega + d$, where $a, b, c, d \in \mathbb{Z}$. Noting that $\omega = \frac{1+i}{\sqrt{2}}$, we have

$$u = (d + \frac{c-a}{2}\sqrt{2}) + (b + \frac{c+a}{2}\sqrt{2})i.$$

If $c - a$ (and therefore $c + a$) is even, then u is of the first form, with $\alpha = d + \frac{c-a}{2}\sqrt{2}$ and $\beta = b + \frac{c+a}{2}\sqrt{2}$. If $c - a$ (and therefore $c + a$) is odd, then u is of the second form, with $\alpha = d + \frac{c-a-1}{2}\sqrt{2}$ and $\beta = b + \frac{c+a-1}{2}\sqrt{2}$. $\square$

Lemma 5.6. *There is an algorithm which, given a pair of upright rectangles A and B, enumerates all solutions of the two-dimensional grid problem for A and B. Moreover, the algorithm requires only a constant number of arithmetic operations per solution produced.*

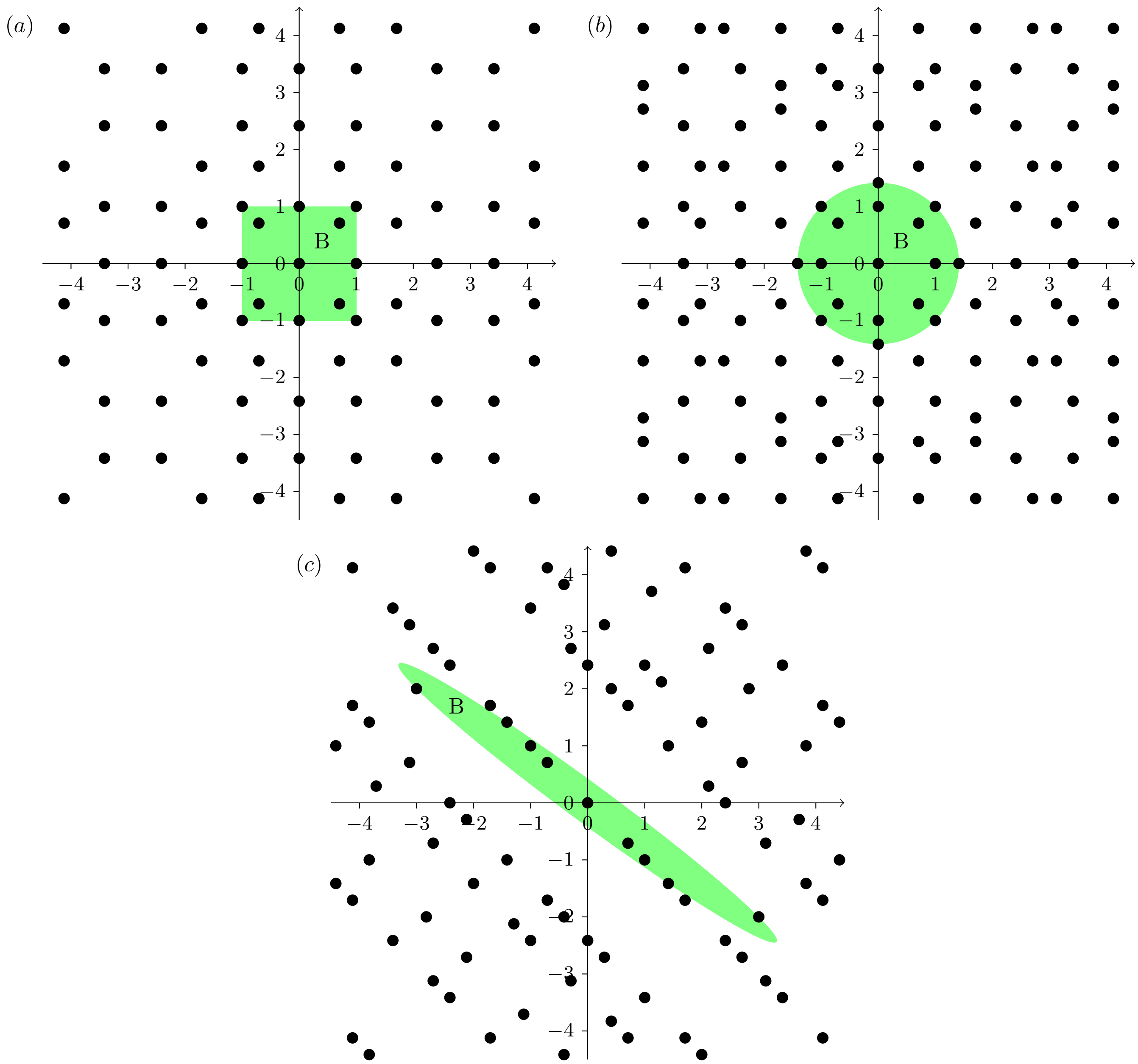


Figure 2: The complex grid for three different convex sets B. In each case, the set B is shown in green, and grid points are shown as black dots. (a) $B = [-1, 1]^2$. (b) $B = \{(x, y) \mid x^2 + y^2 \leqslant 2\}$. (c) $B = \{(x, y) \mid 6x^2 + 16xy + 11y^2 \leqslant 2\}$.

Proof. By assumption, $A = A_x \times A_y$ and $B = B_x \times B_y$, where A_x, A_y, B_x, and B_y are closed intervals. By Lemma 5.5, any potential solution is of the form $u = \alpha + \beta i$ or $u = \alpha + \beta i + \omega$, where $\alpha, \beta \in \mathbb{Z}[\sqrt{2}]$. When $u = \alpha + \beta i$, then $u^\bullet = \alpha^\bullet + \beta^\bullet i$. Therefore, the two-dimensional grid constraints $u \in A$ and $u^\bullet \in B$ are equivalent to the one-dimensional constraints $\alpha \in A_x$, $\alpha^\bullet \in B_x$ and $\beta \in A_y$, $\beta^\bullet \in B_y$. On the other hand, when $u = \alpha + \beta i + \omega$, let $v = u - \omega = \alpha + \beta i$. Then $v^\bullet = u^\bullet + \omega$, and the constraints $u \in A$ and $u^\bullet \in B$ are equivalent to $v \in A - \omega$ and $v^\bullet \in B + \omega$, which reduces to the one-dimensional constraints $\alpha \in A_x - \frac{1}{\sqrt{2}}$, $\alpha^\bullet \in B_x + \frac{1}{\sqrt{2}}$ and $\beta \in A_y - \frac{1}{\sqrt{2}}$, $\beta^\bullet \in B_y + \frac{1}{\sqrt{2}}$. In both cases, the solutions to the one-dimensional constraints can be efficiently enumerated by Proposition 4.5. □

5.2 Upright sets

We can generalize the method of Section 5.1 to convex sets that are *close* to upright rectangles in a suitable sense.

Definition 5.7. Let A be a bounded convex subset of $\mathbb{R}^2$. The *bounding box* of A, denoted $\mathrm{BBox}(A)$, is the smallest set of the form $[x_0, x_1] \times [y_0, y_1]$ that contains A. The *uprightness* of A, denoted $\mathrm{up}(A)$, is defined to be the ratio of the area of A to the area of its bounding box:

$$\mathrm{up}(A) = \frac{\mathrm{area}(A)}{\mathrm{area}(\mathrm{BBox}(A))}. \tag{6}$$

Therefore, the uprightness is a real number between 0 and 1. We say that A is M*-upright* if $\mathrm{up}(A) \geqslant M$.

Lemma 5.8. *There exists an algorithm which, given a pair A, B of convex M-upright sets, enumerates all solutions of the two-dimensional grid problem for A and B. Moreover, the algorithm requires $O(1/M^2)$ arithmetic operations per solution produced. In particular, when $M > 0$ is fixed, it requires only a constant number of operations per solution.*

Proof. By Lemma 5.6, we can efficiently enumerate the solutions of the grid problem for $\mathrm{BBox}(A)$ and $\mathrm{BBox}(B)$. Moreover, as shown in the proof of Lemma 5.6, the solutions are arranged in rows and columns. For each such candidate solution u, we only need to check whether u is also a solution for A and B. To establish the efficiency of the algorithm, we need to ensure that the total number of solutions is not too small in relation to the total number of candidates produced. To see this, note that, with the exception of trivial cases, when the number of rows or columns is very small, M-uprightness and convexity ensure that the proportion of candidates u that are solutions for A and B is approximately $M^2 : 1$. Therefore, the runtime per solution differs from that of Lemma 5.6 by at most a factor of $O(1/M^2)$. □

Example 5.9. Figure 3 shows three different examples of grid problems. Each example uses the same set $B = [-1, 1] \times [-1, 1]$, shown in green. The grid for B is shown as black dots. The sets A_i are shown in red, for $i = 1, 2, 3$, and their bounding boxes are shown in outline. The typical case of an upright set is A_1. Here, a fixed proportion of grid points from the bounding box of A_1 are elements of A_1. The exceptional case of an upright set is A_2. Its bounding box spans only two columns of the grid. Therefore, although the bounding box contains many grid points, A_2 does not. However, this case is easily dealt with, by solving a one-dimensional grid problem for each of the grid columns separately. Finally, the set A_3 is not upright. In this case, Lemma 5.8 is not helpful, and a priori, it could be a difficult problem to find grid points in A_3.

5.3 Grid operators

The method of Section 5.2 can be further generalized by using certain linear transformations to turn non-upright sets into upright sets. The linear transformations that are useful for this purpose are *special grid operators*:

Definition 5.10. As before, we regard $\mathbb{Z}[\omega]$ as a subset of $\mathbb{R}^2$. A real linear operator $G : \mathbb{R}^2 \to \mathbb{R}^2$ is called a *grid operator* if $G(\mathbb{Z}[\omega]) \subseteq \mathbb{Z}[\omega]$. Moreover, a grid operator G is called *special* if it has determinant ± 1.

Grid operators are characterized by the following lemma.

Lemma 5.11. *Let $G : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear operator, which we can identify with a real 2×2-matrix with real entries. Then G is a grid operator if and only if it is of the form*

$$G = \begin{bmatrix} a + \frac{a'}{\sqrt{2}} & b + \frac{b'}{\sqrt{2}} \\ c + \frac{c'}{\sqrt{2}} & d + \frac{d'}{\sqrt{2}} \end{bmatrix}, \tag{7}$$

where $a, b, c, d, a', b', c', d'$ are integers satisfying $a + b + c + d \equiv 0 \pmod 2$ and $a' \equiv b' \equiv c' \equiv d' \pmod 2$.

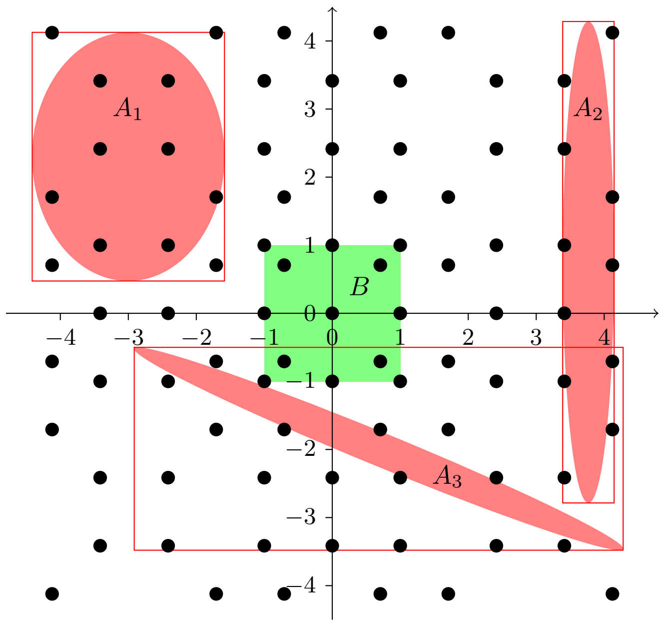


Figure 3: Grid problems for upright and non-upright sets

Proof. By Lemma 5.5, we know that a vector $u \in \mathbb{R}^2$ is in $\mathbb{Z}[\omega]$ if and only if it can be written of the form

$$u = \begin{bmatrix} x_1 + \frac{x_2}{\sqrt{2}} \\ y_1 + \frac{y_2}{\sqrt{2}} \end{bmatrix}, \tag{8}$$

where x_1, x_2, y_1, y_2 are integers and $x_2 \equiv y_2 \pmod 2$. A simple computation then shows that every operator of the form (7) is a grid operator. For the converse, consider an arbitrary grid operator G. We prove the claim by applying G to the three particular points $\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$, $\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$, and $\frac{1}{\sqrt{2}}\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right] \in \mathbb{Z}[\omega]$. From $G\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right] \in \mathbb{Z}[\omega]$ and $G\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right] \in \mathbb{Z}[\omega]$, it follows that the columns of G are of the form (8), so that G is of the form (7), with integers $a, b, c, d, a', b', c', d'$ satisfying $a' \equiv c' \pmod 2$ and $b' \equiv d' \pmod 2$. Moreover, we have

$$G\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} \frac{a'+b'}{2} + \frac{a+b}{\sqrt{2}} \\ \frac{c'+d'}{2} + \frac{c+d}{\sqrt{2}} \end{bmatrix} \in \mathbb{Z}[\omega],$$

which implies $a+b \equiv c+d \pmod 2$ and $a'+b' \equiv c'+d' \equiv 0 \pmod 2$. Together, these conditions imply $a+b+c+d \equiv 0 \pmod 2$ and $a' \equiv b' \equiv c' \equiv d' \pmod 2$, as claimed. □

Remark 5.12. The composition of two (special) grid operators is again a (special) grid operator. If G is a special grid operator, then G is invertible and G^{-1} is a special grid operator. If G is a (special) grid operator, then $G^\bullet$ is a (special) grid operator, defined by applying $(-)^\bullet$ separately to each matrix entry, and satisfying $G^\bullet u^\bullet = (Gu)^\bullet$.

The interest of special grid operators lies in the following fact:

Proposition 5.13. *Let G be a special grid operator, and let A and B be subsets of $\mathbb{R}^2$. Define*

$$G(A) = \{Gu \mid u \in A\},$$
$$G^\bullet(B) = \{G^\bullet u \mid u \in B\}.$$

Then $u \in \mathbb{Z}[\omega]$ is a solution to the two-dimensional grid problem for A and B if and only if Gu is a solution to the two-dimensional grid problem for $G(A)$ and $G^\bullet(B)$. In particular, the two-dimensional grid problem for A and B is computationally equivalent to that for $G(A)$ and $G^\bullet(B)$.

Proof. Let $u \in \mathbb{Z}[\omega]$. Then u is a solution to the grid problem for A and B if and only if $u \in A$ and $u^\bullet \in B$, if and only if $Gu \in G(A)$ and $G^\bullet u^\bullet = (Gu)^\bullet \in G^\bullet(B)$, if and only if Gu is a solution to the grid problem for $G(A)$ and $G^\bullet(B)$. □

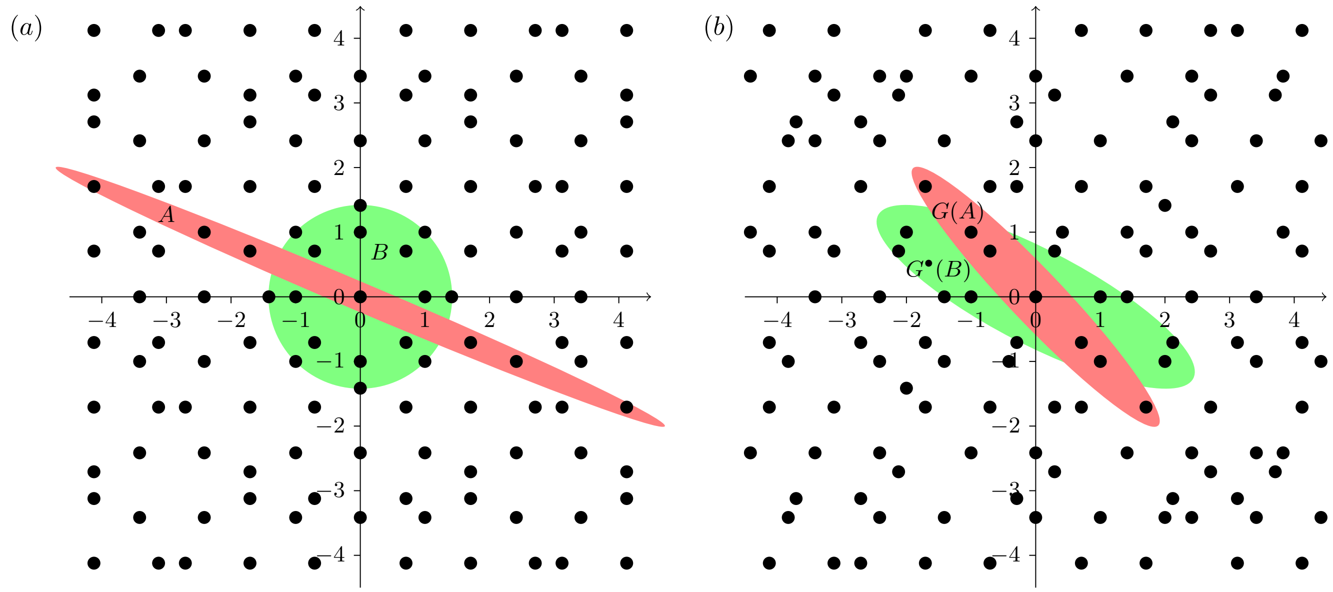


Figure 4: (a) The grid problem for two sets A and B. (b) The grid problem with $G(A)$ and $G^{\bullet}(B)$. Note that the solutions of (a), which are the grid points in the set A, are in one-to-one correspondence with the solutions of (b), which are the grid points in the set $G(A)$.

Example 5.14. Figure 4(a) illustrates the grid problem for a pair of sets A and B. As before, the set B is shown in green, and Grid(B) is shown as black dots. The set A is shown in red, and the solutions to the grid problem are the seven grid points that lie in A. Figure 4(b) shows the grid problem for the sets $G(A)$ and $G^{\bullet}(B)$, where G is the special grid operator

$$G = \begin{bmatrix} 1 & \sqrt{2} \\ 0 & 1 \end{bmatrix}.$$

Note that, as predicted by Proposition 5.13, the solutions of the transformed grid problem are in one-to-one correspondence with those of the original problem; namely, in each case, there are seven solutions.

5.4 Ellipses

Combining the results of Sections 5.2 and 5.3, we know that the grid problem for convex sets A and B can be solved efficiently, provided that we can find a grid operator G such that $G(A)$ and $G^{\bullet}(B)$ are sufficiently upright. Our key technical result is that in case A and B are ellipses, this is always the case.

Definition 5.15. Let D be a positive definite real 2×2-matrix with non-zero determinant, and let $p \in \mathbb{R}^2$ be a point. The *ellipse* defined by D and centered at p is the set

$$E = \{u \in \mathbb{R}^2 \mid (u-p)^{\dagger} D (u-p) \leqslant 1\}.$$

Theorem 5.16. *Suppose $A, B \subseteq \mathbb{R}^2$ are ellipses. Then there exists a grid operator G such that $G(A)$ and $G^{\bullet}(B)$ are $1/6$-upright. Moreover, if A and B are M-upright, then G can be efficiently computed in $O(\log(1/M))$ arithmetic operations.*

Since the proof is long and technical, we give it in Appendix A.

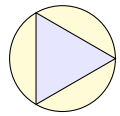
5.5 The enclosing ellipse of a bounded convex set

Our final step in the solution of the two-dimensional grid problem is to generalize Theorem 5.16 from ellipses to arbitrary bounded convex sets with non-empty interior. This can be done because every such set A can be inscribed in an ellipse whose area is not much greater than that of A, as stated in the following proposition.

Proposition 5.17. *Let A be a bounded convex subset of $\mathbb{R}^2$ with non-empty interior. Then there exists an ellipse E such that $A \subseteq E$, and such that*

$$\operatorname{area}(E) \leqslant \frac{4\pi}{3\sqrt{3}} \operatorname{area}(A).$$

The proof is in Appendix B. Note that $\frac{4\pi}{3\sqrt{3}} \approx 2.4184$. We remark that the bound in Proposition 5.17 is sharp; the bound is attained in case A is an equilateral triangle. In this case, the enclosing ellipse is a circle, and the ratio of the areas is exactly $\frac{4\pi}{3\sqrt{3}}$.



5.6 General solution of the two-dimensional grid problem

We are finally in a position to solve the two-dimensional grid problem for arbitrary bounded convex sets of non-empty interior.

Theorem 5.18. *There is an algorithm which, given two bounded convex subset A and B of $\mathbb{R}^2$ with non-empty interior, enumerates all solutions of the two-dimensional grid problem for A and B. Moreover, if A and B are M-upright, then the algorithm requires $O(\log(1/M))$ arithmetic operations overall, plus a constant number of arithmetic operations per solution produced.*

Proof. Given two such sets A and B, we can first find ellipses A' and B' containing A and B, respectively, and whose areas exceed those of A and B by at most a fixed constant factor N. Such ellipses exist by Proposition 5.17; moreover, it is not hard to see that they can be found efficiently if A and B are polygons with rational vertices. Since we assume that each given convex set is equipped with such an enclosing rational polygon (Remark 5.4), A' and B' can be found efficiently for arbitrary given A and B.

Next, by Theorem 5.16, we can use $O(\log(1/M))$ arithmetic operations to find a grid operator G such that $G(A')$ and $G^\bullet(B')$ are 1/6-upright. It follows that $G(A)$ and $G^\bullet(B)$ are $N/6$-upright. By Lemma 5.8, we can efficiently enumerate all solutions u of the grid problem for $G(A)$ and $G^\bullet(B)$. By Proposition 5.13, $G^{-1}u$ then enumerates the solutions to the grid problem for A and B. ∎

Remark 5.19. Note that the complexity of $O(\log(1/M))$ overall operations in Theorem 5.18 is exponentially better than the complexity of $O(1/M^2)$ per candidate we obtained in Lemma 5.8. This improvement is entirely due to the use of grid operators in Theorem 5.16.

5.7 Scaled grid problems

Sometimes we want to find solutions to a grid problem where the points are taken in $\mathbb{D}[\omega]$ instead of $\mathbb{Z}[\omega]$. There are two variants of this problem: we may either enumerate the solutions for a *fixed* denominator exponent, or enumerate all solutions in order of *increasing* least denominator exponent.

Definition 5.20. Let A and B be subsets of $\mathbb{R}^2$. The *two-dimensional scaled grid problem for fixed* $k \geqslant 0$ is to find $u \in \frac{1}{\sqrt{2}^k}\mathbb{Z}[\omega]$ satisfying $u \in A$ and $u^\bullet \in B$. The *two-dimensional scaled grid problem for arbitrary* $k \geqslant 0$ is to find $u \in \mathbb{D}[\omega]$ satisfying $u \in A$ and $u^\bullet \in B$.

Proposition 5.21. *There is an algorithm which, given two bounded convex subsets A and B of $\mathbb{R}^2$ with non-empty interior and an integer $k \geqslant 0$, enumerates all solutions of the two-dimensional scaled grid problem for A, B, and k. Moreover, if A and B are M-upright, then the algorithm requires $O(\log(1/M))$ arithmetic operations overall, plus a constant number of arithmetic operations per solution produced.*

Proof. Note that $u = \frac{1}{\sqrt{2}^k}v$ is a solution to the scaled grid problem for A, B, and k if and only if v is a solution to the (unscaled) grid problem for $\sqrt{2}^k A$ and $(-\sqrt{2})^k B$. The claim then immediately follows from Proposition 5.18. ∎

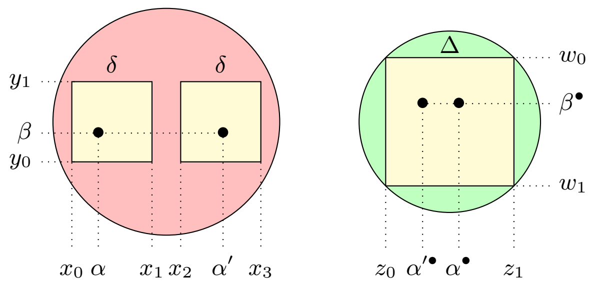
Proposition 5.22. *There is an algorithm which, given two bounded convex subsets A and B of $\mathbb{R}^2$ with non-empty interior, enumerates (the infinite sequence of) all solutions u of the two-dimensional scaled grid problem for A, B, and arbitrary $k \geqslant 0$, in order of increasing k. Moreover, if A and B are M-upright, then the algorithm requires $O(\log(1/M))$ arithmetic operations overall, plus a constant number of arithmetic operations per solution produced.*

Proof. This can be done by applying Lemma 5.21 to each $k = 0, 1, 2, \ldots$, in increasing order. In principle, this method enumerates each solution multiple times, since each solution for k is also a solution for $k+1$. As a slight optimization, such duplicate enumeration can be avoided by noting that for $k > 0$, $u = \frac{1}{\sqrt{2}^k}(a\omega^3 + b\omega^2 + c\omega + d)$ is an element of $\mathbb{Z}[\omega]/\sqrt{2}^k - \mathbb{Z}[\omega]/\sqrt{2}^{k-1}$ if and only if $a - c$ or $b - d$ (or both) are odd. Finally, we note that, because uprightness is invariant under scaling, the grid operator G in the proof of Proposition 5.18 only needs to be computed once, rather than once for every k. □

We end this section with some lower bounds on the number of solutions to two-dimensional scaled grid problems.

Lemma 5.23. *Let A and B be convex subsets of $\mathbb{R}^2$, and let $k \geqslant 0$. Assume A contains a circle of radius r and B contains a circle of radius R, such that $rR \geqslant \frac{1}{2^k}(1+\sqrt{2})^2$. Then the scaled grid problem for k has at least 2 solutions.*

Proof. By scaling the problem by a factor of $\sqrt{2}^k$, we can assume without loss of generality that $k = 0$. Let $\delta = r/\sqrt{2}$ and $\Delta = R\sqrt{2}$, and inscribe two squares of size $\delta \times \delta$ in the first circle, and one square of size $\Delta \times \Delta$ in the second circle, as shown here:



Since $\delta\Delta = rR \geqslant (1+\sqrt{2})^2$, by Lemma 4.4, we can find $\alpha, \alpha', \beta \in \mathbb{Z}[\sqrt{2}]$ such that $\alpha \in [x_0, x_1]$, $\alpha^\bullet \in [z_0, z_1]$, $\alpha' \in [x_2, x_3]$, $\alpha'^\bullet \in [z_0, z_1]$, $\beta \in [y_0, y_1]$, and $\beta^\bullet \in [w_0, w_1]$. Then $u = \alpha + i\beta$ and $v = \alpha' + i\beta$ are two different solutions to the two-dimensional grid problem as claimed. □

Lemma 5.24. *Let A and B be convex subsets of $\mathbb{R}^2$, and assume that the two-dimensional scaled grid problem for k has at least two distinct solutions. Then for all $\ell \geqslant 0$, the scaled grid problem for $k+2\ell$ has at least $2^\ell + 1$ solutions.*

Proof. Let $u \neq v$ be solutions of the scaled grid problem for k. For each $j = 0, 1, \ldots, 2^\ell$, let $\phi = \frac{j}{2^\ell}$, and consider $u_j = \phi u + (1-\phi)v$. Then u_j has denominator exponent $k+2\ell$. Also, u_j is a convex combination of u and v; moreover, since $\phi^\bullet = \phi$, we also know that $u_j^\bullet = \phi u^\bullet + (1-\phi)v^\bullet$ is a convex combination of $u^\bullet$ and $v^\bullet$. Since A and B are convex, it follows that u_j is a solution of the scaled grid problem for $k + 2\ell$, yielding $2^\ell + 1$ distinct solutions. □

Remark 5.25. The bound in Lemma 5.24 is sufficient for our purposes, but it is not tight. In fact, the number of solutions grows as $O(4^k)$.

6 Solving a Diophantine equation

We will be interested in solving equations of the following form: given $\xi \in \mathbb{D}[\sqrt{2}]$, find $t \in \mathbb{D}[\omega]$ such that

$$t^\dagger t = \xi. \tag{9}$$

The following necessary condition is immediate:

Lemma 6.1 (Necessary condition)**.** *The equation (9) has a solution only if $\xi \geqslant 0$ and $\xi^\bullet \geqslant 0$.*

Proof. Assume $t^\dagger t = \xi$. Since t is a complex number, we have $\xi = t^\dagger t \geqslant 0$. Similarly, since $t^\bullet$ is a complex number, we have $\xi^\bullet = (t^\bullet)^\dagger (t^\bullet) \geqslant 0$. □

The following theorem states that the problem of solving the equation (9) can be reduced to the prime factorization problem for integers.

Theorem 6.2. *Let $\xi \in \mathbb{D}[\sqrt{2}]$. Note that $\xi^\bullet\xi \in \mathbb{D}$, so we can write $\xi^\bullet\xi = \frac{n}{2^\ell}$ for some $n \in \mathbb{Z}$ and $\ell \in \mathbb{N}$. There exists a probabilistic algorithm which, given ξ and, in case $n \neq 0$, a prime factorization of n, determines whether or not the equation (9) has a solution, and finds a solution if there is one. Moreover, the expected runtime of this algorithm is polynomial in the size of n.*

Theorem 6.2 is a well-known result in computational algebraic number theory. For the benefit readers who are not experts in number theory, we give an elementary and more or less self-contained proof in Appendix C.

7 The approximate synthesis algorithm

7.1 The approximate synthesis problem

Recall that the z-rotation by angle θ is the unitary operator

$$R_z(\theta) = e^{-i\theta Z/2} = \begin{pmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{pmatrix}.$$

Definition 7.1. Given θ and a precision $\varepsilon > 0$, the *approximate synthesis problem for z-rotations* is to find an operator U expressible in the single-qubit Clifford+T gate set, such that

$$\|R_z(\theta) - U\| \leqslant \varepsilon. \tag{10}$$

Moreover, we want the T-count of the operator U to be as small as possible; here, the T-count of a Clifford+T circuit is the number of T-gates appearing in it. The norm in (10) is the operator norm.

It is known from [10] that a single-qubit operator can be exactly represented over the Clifford+T gate set if and only if it can be written of the form

$$U = \begin{pmatrix} u & -t^\dagger\omega^\ell \\ t & u^\dagger\omega^\ell \end{pmatrix}, \tag{11}$$

where $u, t \in \mathbb{D}[\omega]$ and ℓ is an integer. The following lemma shows that, for the purposes of approximate synthesis, we may assume without loss of generality that $\ell = 0$.

Lemma 7.2. *If $\varepsilon < |1 - e^{i\pi/8}|$, then all solutions of the approximate synthesis problem have the form*

$$U = \begin{pmatrix} u & -t^\dagger \\ t & u^\dagger \end{pmatrix}. \tag{12}$$

If $\varepsilon \geqslant |1 - e^{i\pi/8}|$, then there exists a solution of T-count 0 (i.e., a Clifford operator), and it is also of the form (12).

Proof. To prove the first claim, assume $\varepsilon < |1 - e^{i\pi/8}|$. Let U be of the form (11), satisfying (10). Let $e^{i\phi_1}$ and $e^{i\phi_2}$ be the eigenvalues of $UR_z(\theta)^{-1}$, with $\phi_1, \phi_2 \in [-\pi, \pi]$. Using (10), we have $\|I - UR_z(\theta)^{-1}\| \leqslant \varepsilon < |1 - e^{i\pi/8}|$. On the other hand, $\|I - UR_z(\theta)^{-1}\| = \max\{|1 - e^{i\phi_1}|, |1 - e^{i\phi_2}|\}$. It follows that $|1 - e^{i\phi_j}| < |1 - e^{i\pi/8}|$ for $j = 1, 2$, hence $-\pi/8 < \phi_j < \pi/8$, hence $-\pi/4 < \phi_1 + \phi_2 < \pi/4$, so $|1 - e^{i(\phi_1+\phi_2)}| < |1 - e^{i\pi/4}| = |1 - \omega|$. On other hand, we have $e^{i(\phi_1+\phi_2)} = \det(UR_z(\theta)^{-1}) = \omega^\ell$, hence $|1 - \omega^\ell| < |1 - \omega|$, which implies $\omega^\ell = 1$. Therefore, U is of the form (12). To prove the second claim, assume $\varepsilon \geqslant |1 - e^{i\pi/8}|$. Let j be the integer closest to $\frac{2\theta}{\pi}$, so that $|j - \frac{2\theta}{\pi}| \leqslant \frac{1}{2}$, or equivalently, $|j\frac{\pi}{4} - \frac{\theta}{2}| \leqslant \frac{\pi}{8}$. Let U be as in (12), with $u = \omega^{-j}$ and $t = 0$. Then $\|R_z(\theta) - U\| = |e^{i\theta/2} - u^\dagger| = |1 - u^\dagger e^{-i\theta/2}| = |1 - e^{i(j\frac{\pi}{4} - \frac{\theta}{2})}| \leqslant |1 - e^{i\frac{\pi}{8}}| \leqslant \varepsilon$. So (10) holds. But $U = S^j\omega^{-j}$ is a Clifford operator, so has T-count 0. $\square$

Our strategy is therefore to approximate $R_z(\theta)$ by an operator U of the form (12), with $u, t \in \mathbb{D}[\omega]$, and then use the exact synthesis algorithm of [10] to synthesize U into a sequence of Clifford+T gates with minimal T-count. The following lemma relates the T-count of the resulting circuit to the least denominator exponent k of u.

Lemma 7.3. *Let U be a unitary operator as in* (12)*, where $u, t \in \mathbb{D}[\omega]$, and let k be the least denominator exponent of u. Then the T-count of U is either $2k - 2$ or $2k$. Moreover, if $k > 0$ and U has T-count $2k$, then $U' = TUT^\dagger$ has T-count $2k - 2$. We further note that $\|R_z(\theta) - U'\| = \|R_z(\theta) - U\|$, so for the purpose of solving* (10)*, it does not matter whether U or U' is used. Hence, without loss of generality, we may assume that U as in* (12) *always has T-count exactly $2k - 2$ when $k > 0$, and 0 when $k = 0$.*

Proof. Because U is unitary, we have $t^\dagger t + u^\dagger u = 1$. We first claim that t and u have the same least denominator exponent. Indeed, in the ring $\mathbb{Z}[\omega]$, an element s is divisible by $\sqrt{2}$ if and only if $s^\dagger s$ is divisible by 2. The left-to-right implication is obvious, and the right-to-left implication follows, e.g., from Lemma 2 of [6]. Then for any $k \geqslant 0$, we have $\sqrt{2}^k u \in \mathbb{Z}[\omega]$ iff $2^k u^\dagger u \in \mathbb{Z}[\omega]$ iff $2^k(1 - t^\dagger t) \in \mathbb{Z}[\omega]$ iff $2^k t^\dagger t \in \mathbb{Z}[\omega]$ iff $\sqrt{2}^k t \in \mathbb{Z}[\omega]$. This proves that u and t have the same denominator exponents, and in particular, the same least denominator exponent.

The claims about the T-counts of U and U' follow by inspection of Figure 2 of [7]. Using the terminology of Definitions 7.4 and 7.6 of [7], this figure shows every possible k-residue of a Clifford+T operator, modulo a right action of the group $\langle S, X, \omega\rangle$. Because U is of the form (12), only a subset of the k-residues is actually possible, and the figure shows that for this subset, the T-count is $2k$ or $2k - 2$. Moreover, in each of the possible cases where $k > 0$ and U has T-count $2k$, the figure also shows that $U' = TUT^\dagger$ has T-count $2k - 2$.

For the final claim, we have $\|R_z(\theta) - U\| = \|TR_z(\theta)T^\dagger - TUT^\dagger\| = \|R_z(\theta) - U'\|$ because $R_z(\theta)$ and T commute. $\square$

So our task is to find an operator U of the form (12), satisfying (10), and such that the denominator exponent of u is as small as possible. It is useful to first re-express (10) as a property of u. Let $z = e^{-i\theta/2}$. Using $u^\dagger u + t^\dagger t = 1$ and $z^\dagger z = 1$, we have

$$\|R_z(\theta) - U\|^2 = \|u - z\|^2 + \|t\|^2 = (u-z)^\dagger(u-z) + t^\dagger t = u^\dagger u + t^\dagger t - z^\dagger u - u^\dagger z + z^\dagger z = 2 - 2\,\mathrm{Re}(z^\dagger u).$$

So (10) is equivalent to $2 - 2\,\mathrm{Re}(z^\dagger u) \leqslant \varepsilon^2$, or equivalently, $\mathrm{Re}(z^\dagger u) \geqslant 1 - \frac{\varepsilon^2}{2}$. If we identify the complex numbers $z = x + yi$ and $u = a + bi$ with 2-dimensional real vectors $\vec{z} = (x, y)^T$ and $\vec{u} = (a, b)^T$, then $\mathrm{Re}(z^\dagger u)$ is just their inner product $\vec{z} \cdot \vec{u}$, and therefore (10) is equivalent to

$$\vec{z} \cdot \vec{u} \geqslant 1 - \frac{\varepsilon^2}{2}. \tag{13}$$

In summary, the approximate synthesis problem reduces to the following:

Problem 7.4. Given an angle θ and a precision $\varepsilon > 0$, find $u, t \in \mathbb{D}[\omega]$ such that

(a) $t^\dagger t + u^\dagger u = 1$,

(b) $\vec{z} \cdot \vec{u} \geqslant 1 - \varepsilon^2/2$, where $z = e^{-i\theta/2}$, with notation as above,

(c) and such that u has the smallest denominator exponent we can find.

7.2 Reduction to a grid problem and a Diophantine equation

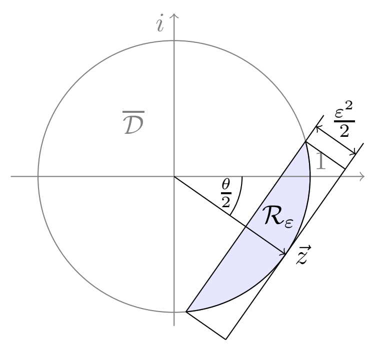
As we will now show, Problem 7.4 can be reduced to a scaled grid problem and a Diophantine equation, and therefore it can be efficiently solved by the methods of Sections 5 and 6. Let $\overline{\mathcal{D}}$ be the closed unit disk, regarded either as a subset of $\mathbb{C}$ or of $\mathbb{R}^2$.

Lemma 7.5. *If $u, t \in \mathbb{D}[\omega]$ and $t^\dagger t + u^\dagger u = 1$, then $u \in \overline{\mathcal{D}}$ and $u^\bullet \in \overline{\mathcal{D}}$.*

Proof. Note that $u^\dagger u = 1 - t^\dagger t \leqslant 1$, so $u \in \overline{\mathcal{D}}$. Similarly, $(u^\bullet)^\dagger(u^\bullet) = 1 - (t^\bullet)^\dagger(t^\bullet) \leqslant 1$, so $u^\bullet \in \overline{\mathcal{D}}$. $\square$

The condition $\vec{z} \cdot \vec{u} \geqslant 1 - \varepsilon^2/2$ from Problem 7.4(b) defines a certain subset of the unit disk, which we call the *ε-region* for θ:

$$\mathcal{R}_\varepsilon = \{\vec{u} \in \overline{\mathcal{D}} \mid \vec{u} \cdot \vec{z} \geqslant 1 - \frac{\varepsilon^2}{2}\}. \tag{14}$$



By Lemma 7.5 and Problem 7.4(b), a *necessary* condition for a solution to Problem 7.4 is that $u \in \mathcal{R}_\varepsilon$ and $u^\bullet \in \overline{\mathcal{D}}$. This is an instance of a scaled two-dimensional grid problem; note that both the ε-region and the unit disk are convex, and are effectively given in the sense of Remark 5.4. Therefore, by Proposition 5.22, there exists an efficient algorithm that enumerates all such u in increasing order of least denominator exponent.

For each $u \in \mathbb{D}[\omega]$ satisfying the grid problem, it remains to check whether the equation $t^\dagger t + u^\dagger u = 1$ has a solution $t \in \mathbb{D}[\omega]$. This is equivalent to solving the Diophantine equation

$$t^\dagger t = 1 - u^\dagger u,$$

which is of the form (9). Let $\xi = 1 - u^\dagger u \in \mathbb{D}[\sqrt{2}]$, and write $\xi^\bullet \xi = \frac{n}{2^\ell}$, where $n \in \mathbb{Z}$ and $\ell \in \mathbb{N}$. By Theorem 6.2, there is an efficient algorithm that can solve this equation (or determine that no solution exists), given a prime factorization of n.

7.3 The main algorithm

Putting together the results of Sections 7.1 and 7.2, we obtain the following algorithm for solving the approximate synthesis problem:

Algorithm 7.6. Given θ and ε, let $A = \mathcal{R}_\varepsilon$ be the ε-region, and let $B = \overline{\mathcal{D}}$ be the unit disk.

1. Use Proposition 5.22 to enumerate the infinite sequence of solutions to the scaled grid problem $u \in A$ and $u^\bullet \in B$, where $u \in \mathbb{D}[\omega]$, in the order of increasing least denominator exponent k.
2. For each such solution u:
   (a) Let $\xi = 1 - u^\dagger u \in \mathbb{D}[\sqrt{2}]$, and write $\xi^\bullet \xi = \frac{n}{2^\ell}$, where $n \in \mathbb{Z}$ and $\ell \geqslant 0$ is minimal.
   (b) Attempt to find a prime factorization of n. If $n \neq 0$ but no prime factorization is found, skip step 2(c) and continue with the next u.
   (c) Use the algorithm of Theorem 6.2 to solve the equation $t^\dagger t = \xi$. If a solution t exists, go to step 3; otherwise, continue with the next u.
3. Define U as in equation (12), let $U' = TUT^\dagger$, and use the exact synthesis algorithm of [10] to find a Clifford+T circuit implementing either U or U', whichever has smaller T-count. Output this circuit and stop.

8 Analysis of the algorithm

8.1 Correctness

Proposition 8.1 (Correctness)**.** *If Algorithm 7.6 terminates, then it yields a valid solution to the approximate synthesis problem, i.e., it yields a Clifford+T circuit approximating $R_z(\theta)$ up to ε.*

Proof. By construction. By steps 2(a) and 2(c) of the algorithm, we have $t^\dagger t + u^\dagger u = 1$, so U is unitary. By step 1 of the algorithm, u belongs to the ε-region, so (13) holds. This implies that U satisfies (10). Moreover, as noted in Lemma 7.3, U' also satisfies (10), so whichever of these operators is returned approximates $R_z(\theta)$ up to ε. □

8.2 Optimality

The optimality of the algorithm hinges on step 2(b), "attempt to find a prime factorization of n". In the presence of a factoring oracle (such as a quantum computer), this can always be done. In this case, Algorithm 7.6 is guaranteed to find an optimal solution to the approximate synthesis problem. In the absence of a factoring oracle, we must potentially discard some candidate solutions u, until we find one for which n can be factored. We analyze these two situations in Propositions 8.2 and 8.8.

Proposition 8.2 (Optimality in the presence of a factoring oracle)**.** *In the presence of an oracle for integer factoring, the circuit returned by Algorithm 7.6 has the smallest T-count of any single-qubit Clifford+T circuit approximating $R_z(\theta)$ up to ε.*

Proof. By construction, step 1 of the algorithm enumerates all solutions u of the scaled grid problem in order of increasing denominator exponent k. Step 2(a) always succeeds, and step 2(b) always succeeds by using the factoring oracle. By Theorem 6.2, step 2(c) succeeds if and only if the equation $t^\dagger t + u^\dagger u = 1$ has a solution. Therefore, when step 2(c) succeeds, the algorithm has found a solution of Problem 7.4 for which u has the lowest possible denominator exponent k. Let m be the T-count of the final solution. By Lemma 7.3, we have $m = 2k - 2$, except when $k = 0$, in which case $m = 0$.

To show that this T-count is minimal, let $\bar{U}$ be any solution of the approximate synthesis problem with T-count $\bar{m}$. By Lemma 7.2, we may assume without loss of generality that

$$\bar{U} = \begin{pmatrix} \bar{u} & -\bar{t}^\dagger \\ \bar{t} & \bar{u}^\dagger \end{pmatrix}.$$

Let $\bar{k}$ be the denominator exponent of $\bar{u}$. By minimality of k, we have $k \leqslant \bar{k}$, hence $m \leqslant \bar{m}$. □

We emphasize that the optimality in Proposition 8.2 is *absolute*, i.e., not merely asymptotic or up to an additive constant. Of all the Clifford+T operators approximating $R_z(\theta)$ to within ε, the algorithm finds one with the lowest T-count.

To analyze the algorithm in the absence of a factoring oracle, we must address the question of how many candidates must be generated before steps 2(b) and 2(c) of the algorithm succeed. In this case, the algorithm may still use any classical algorithm to try to factor the number n in step 2(b), but the amount of effort extended on any particular n must be capped. In our complexity analysis for this case, in Proposition 8.8 below, we conservatively assume that the only n that the algorithm can successfully factor are those n that are already prime. In reality the algorithm might do a little better. In order to complete the analysis, we must rely on a number-theoretic assumption about the distribution of prime numbers.

Hypothesis 8.3. Each number n produced in step 2(a) of Algorithm 7.6 is asymptotically as likely to be prime as a randomly chosen odd number of comparable size. Moreover, the primality of each n can be modelled as an independent random event.

Lemma 8.4. *Each of the numbers n produced in step 2(a) of Algorithm 7.6 satisfies $n \geqslant 0$, and either $n = 0$ or $n \equiv 1 \pmod 8$.*

Proof. See Appendix D. □

Lemma 8.5. *Let u be a candidate produced in step 1 of Algorithm 7.6, let k be its least denominator exponent, and let n be the integer computed in step 2(a). Then $n \leqslant 4^k$.*

Proof. By assumption, we can write $u = v/\sqrt{2}^k$, where $v \in \mathbb{Z}[\omega]$. From step 2(a) of the algorithm, we have $\xi = 1 - u^\dagger u = \frac{1}{2^k}(2^k - v^\dagger v) = \frac{\alpha}{2^k}$, where $\alpha \in \mathbb{Z}[\sqrt{2}]$. Therefore $\xi^\bullet \xi = \frac{\alpha^\bullet \alpha}{2^{2k}} = \frac{n}{2^\ell}$. Since $\alpha^\bullet\alpha$ is an integer and ℓ is minimal, we have $\ell \leqslant 2k$. Also, by assumption, both u and $u^\bullet$ are in the unit disk, so $u^\dagger u \leqslant 1$ and $(u^\bullet)^\dagger(u^\bullet) \leqslant 1$. It follows that $0 \leqslant \xi, \xi^\bullet \leqslant 1$, hence $\xi^\bullet\xi \leqslant 1$. Therefore, $\frac{n}{2^\ell} \leqslant 1$, which implies $n \leqslant 2^\ell \leqslant 4^k$ as claimed. □

Lemma 8.6. *Let $b > 0$ be an arbitrary fixed constant. Then for $a \geqslant 1$,*

$$\sum_{x=1}^{\infty} \left(1 - \frac{1}{a + b\ln x}\right)^x = O(a).$$

Proof. See Appendix E. □

Definition 8.7. Let U', U'' be two solutions of the approximate synthesis problem of the form

$$U' = \begin{pmatrix} u' & -t'^\dagger \\ t' & u'^\dagger \end{pmatrix}, \quad U'' = \begin{pmatrix} u'' & -t''^\dagger \\ t'' & u''^\dagger \end{pmatrix}. \tag{15}$$

We say that U' and U'' are *equivalent solutions* if $u' = u''$.

Proposition 8.8 (Near-optimality in the absence of a factoring oracle)**.** *Let m be the T-count of the solution of the approximate synthesis problem found by Algorithm 7.6 in the absence of an oracle for integer factoring. Then*

(a) The approximate synthesis problem has an expected number of at most $O(\log(1/\varepsilon))$ non-equivalent solutions with T-count less than m.

(b) The expected value of m is $m'' + O(\log(\log(1/\varepsilon)))$, where m' and m'' are the T-counts of the optimal and second-to-optimal solutions of the approximate synthesis problem (up to equivalence), and $m' \leqslant m''$.

Proof. If $\varepsilon \geqslant |1 - e^{i\pi/8}|$, then by Lemma 7.2, there is a solution with T-count 0, and the algorithm easily finds it. In this case, there is nothing to show. So assume without loss of generality that $\varepsilon < |1 - e^{i\pi/8}|$. Then by Lemma 7.2, all solutions are of the form (12).

(a) Consider the list $u_1, u_2, \ldots$ of candidates generated in step 1 of the algorithm. Let $k_1, k_2, \ldots$ be their respective least denominator exponents, and let $n_1, n_2, \ldots$ be the corresponding integers calculated in step 2(a). By Lemma 8.5, we have $n_j \leqslant 4^{k_j}$ for all j. By Hypothesis 8.3, the probability that n_j is prime is asymptotically no smaller than that of a randomly chosen odd integer less than 4^{k_j}, which, by the well-known prime number theorem, is greater than

$$p_j := \frac{2}{\ln(4^{k_j})} = \frac{1}{k_j \ln 2}. \tag{16}$$

Note that u_1 and u_2 are two distinct solutions to the scaled grid problem of step 1 of the algorithm. Since the candidates are enumerated in order of increasing denominator exponent, k_2 is a denominator exponent for both u_1 and u_2. It follows by Lemma 5.24 that there are at least $2^\ell+1$ distinct candidates of denominator exponent $k_2+2\ell$, for all $\ell \geqslant 0$. In other words, for all j, if $j \leqslant 2^\ell+1$, we have $k_j \leqslant k_2+2\ell$. In particular, this holds for $\ell = \lfloor 1+\log_2 j \rfloor$, and therefore,

$$k_j \leqslant k_2 + 2(1+\log_2 j). \tag{17}$$

Combining (17) with (16), we have

$$p_j \geqslant \frac{1}{(k_2+2(1+\log_2 j))\ln 2} = \frac{1}{(k_2+2)\ln 2 + 2\ln j} \tag{18}$$

Let j_0 be the smallest index such that n_{j_0} is prime. By Hypothesis 8.3, we can treat the primality of each n_j as an independent random event. Therefore,

$$\begin{aligned} P(j_0 > j) &= P(n_1,\ldots,n_j \text{ are not prime}) \\ &\leqslant (1-p_1)(1-p_2)\cdots(1-p_j) \\ &\leqslant (1-p_j)^j \\ &\leqslant \left(1 - \frac{1}{(k_2+2)\ln 2 + 2\ln j}\right)^j. \end{aligned}$$

The expected value of j_0 is

$$E(j_0) \;=\; \sum_{j=0}^{\infty} P(j_0 > j) \;\leqslant\; 1 + \sum_{j=1}^{\infty}\left(1 - \frac{1}{(k_2+2)\ln 2 + 2\ln j}\right)^j \;=\; O(k_2), \tag{19}$$

where we have used Lemma 8.6 to estimate the sum.

Next, we will estimate k_2. The width of the ε-region $\mathcal{R}_\varepsilon$, as shown in (14), is $\varepsilon^2/2$ at the widest point, and we can inscribe a disk of radius $r = \varepsilon^2/4$ in it. Also, the closed unit disk $\overline{\mathcal{D}}$ has radius $R = 1$. It follows by Lemma 5.23 that the scaled grid problem for $\mathcal{R}_\varepsilon$ and $\overline{\mathcal{D}}$, as in step 1 of the algorithm, has at least two solutions, provided that

$$rR = \frac{\varepsilon^2}{4} \geqslant \frac{(1+\sqrt{2})^2}{2^k}, \tag{20}$$

or equivalently, provided that

$$k \geqslant 2 + 2\log_2(1+\sqrt{2}) + 2\log_2(1/\varepsilon). \tag{21}$$

We therefore have $k_2 \leqslant k$ for *all* k satisfying (21). It follows that

$$k_2 = O(\log(1/\varepsilon)), \tag{22}$$

and therefore, using (19), also

$$E(j_0) = O(\log(1/\varepsilon)). \tag{23}$$

To finish the proof of part (a), recall that j_0 was defined to be the smallest index such that n_{j_0} is prime. This ensures that step 2(b) of the algorithm succeeds for the candidate u_{j_0}. Furthermore, we have $n \equiv 1 \pmod 8$ by Lemma 8.4, and therefore the equation $t^\dagger t = \xi$ has a solution by Proposition C.26. Hence the remaining steps of the algorithm also succeed for u_{j_0}.

Now let r be the number of non-equivalent solutions of the approximate synthesis problem of T-count strictly less than m. As noted above, any such solution U is of the form (12). Then the least denominator exponent of u is strictly smaller than k_{j_0}, so that $u = u_j$ for some $j < j_0$. In this way, each of the r non-equivalent solutions is mapped to a different index $j < j_0$. It follows that $r < j_0$, and hence $E(r) \leqslant E(j_0) = O(\log(1/\varepsilon))$, as was to be shown.

(b) Let U' be an optimal solution of the approximate synthesis problem, and let U'' be optimal among the solutions that are not equivalent to U'. Let u' and u'' be as in (15), and let k', k'' be the least denominator exponents of u' and u'', respectively, with $k' \leqslant k''$. Let k_2 and j_0 be as in the proof of part (a). Note that, by definition, $k_2 \leqslant k''$. Let k be the least denominator exponent of the solution of the approximate synthesis problem found by the algorithm. Then $k \leqslant k_{j_0}$. Using (17), we have

$$k \leqslant k_{j_0} \leqslant k_2 + 2(1+\log_2 j_0) \leqslant k'' + 2(1+\log_2 j_0).$$

Recall from Lemma 7.3 that $2k-2 \leqslant m \leqslant 2k$, and similarly $2k''-2 \leqslant m'' \leqslant 2k''$. Hence $m \leqslant 2k \leqslant 2k''+4(1+\log_2 j_0) \leqslant m''+6+4\log_2 j_0$. These calculations apply to any one run of the algorithm. Taking expected values over many randomized runs, we therefore have

$$E(m) \leqslant m''+6+4E(\log_2 j_0) \leqslant m''+6+4\log_2 E(j_0). \tag{24}$$

Note that we have used the law $E(\log j_0) \leqslant \log(E(j_0))$, which holds because log is a concave function. Combining (24) with (23), we therefore have the desired result:

$$E(m) = m'' + O(\log(\log(1/\varepsilon))). \tag{25}$$

□

Remark 8.9. In the near-optimal case of Proposition 8.8, our algorithm can additionally be used to compute a firm lower bound for the T-count of any solution of the approximate synthesis problem for the given θ and ε. Namely, the algorithm can consider the first candidate u_j for which the Diophantine equation solver does not fail — i.e., either it solves the equation or it times out. If k_j is the least denominator exponent of u_j, then a lower bound for the T-count is $2k_j - 2$ (or 0 when $k_j = 0$). Note that this is not the usual information-theoretic lower bound that applies in the average case, but an actual lower bound for each particular problem instance.

8.3 Worst-case behavior

In [14], an approximate synthesis algorithms for z-rotations was given that always returns a solution of T-count at most $K + 4\log_2(1/\varepsilon)$, where K is a constant approximately equal to 10. We note that Algorithm 7.6 enumerates all the solutions of the grid problem for the ε-region, whereas the algorithm of [14] only enumerates a subset of the solutions. Also, Algorithm 7.6 can solve the Diophantine equation in all the cases in which the algorithm of [14] can solve it. It follows that in all cases, the solution returned by Algorithm 7.6 is at least as good as that returned by the algorithm of [14]. In particular, Algorithm 7.6 always returns a solution of T-count at most $K + 4\log_2(1/\varepsilon)$. Moreover, it is known from [14, Section 9], that there are certain combinations of θ and ε for which this T-count is actually optimal to within a constant number of gates. Thus our algorithm's performance is $K + 4\log_2(1/\varepsilon)$ in the worst case, but this worst case behavior is only achieved in those cases where it is actually optimal.

One may ask whether a more precise statement can be made about the angles θ for which this worst-case behavior occurs. We believe that this is indeed possible. Let $\mathbb{Q}(\sqrt{2})$ be the field of rational numbers extended with $\sqrt{2}$.

Conjecture 8.10. *Fix an angle θ, and consider the T-count as $\varepsilon \to 0$. Then the T-count behaves as $K + 4\log_2(1/\varepsilon)$ when $\tan\frac{\theta}{2} \in \mathbb{Q}(\sqrt{2})$, and as $K + 3\log_2(1/\varepsilon)$ otherwise.*

Giving a rigorous proof of this conjecture is beyond the scope of this paper and probably difficult, but we will sketch a plausibility argument. We believe that the first part, showing that the T-count behaves no better than $K + 4\log_2(1/\varepsilon)$ when $\tan\frac{\theta}{2} \in \mathbb{Q}(\sqrt{2})$, could be made rigorous without too much difficulty. In addition, we also offer experimental evidence supporting Conjecture 8.10 in Section 10.

The essential point is this. Consider the grid associated to a convex bounded set B as in Figure 5. Note that the slope r of any line passing through two grid points satisfies $r \in \mathbb{Q}(\sqrt{2})$ (unless the line is vertical, in which case the slope is infinite). Conversely, fix some real number r. Let us call a line of slope r a *grid line* if it passes through at least one grid point. We first wish to show that if $r \in \mathbb{Q}(\sqrt{2})$, then the set of grid lines is discrete, i.e., there is a minimum distance between each pair of grid lines. To see why this is so, note that $r \in \frac{1}{n}\mathbb{Z}[\sqrt{2}]$ for some $n \in \mathbb{Z}$. Each grid line is uniquely determined by its y-intercept, i.e., its point of intersection with the y-axis. Consider a grid line with y-intercept b, and passing through the grid point $u \in \mathbb{Z}[\omega]$. Using Lemma 5.5, we can assume without loss of generality that $u = \alpha + \beta i$, where $\alpha, \beta \in \mathbb{Z}[\sqrt{2}]$; the case where $u = \alpha + \beta i + \omega$ is analogous. Then we have $b = \beta - r\alpha$, and therefore $b \in \frac{1}{n}\mathbb{Z}[\sqrt{2}]$. Moreover, the line with y-intercept $b^\bullet$ and slope $r^\bullet$ contains $u^\bullet$, and therefore intersects B. Since B is bounded, it follows that $b^\bullet$ lies in a certain finite interval B' of real numbers determined only by B and $r^\bullet$. The two relations $b \in \frac{1}{n}\mathbb{Z}[\sqrt{2}]$ and $b^\bullet \in B'$ determine a one-dimensional grid problem, and therefore the set of solutions is discrete. This shows that the set of grid lines is discrete when $r \in \mathbb{Q}(\sqrt{2})$.

Recall from (14) that for given θ and ε, the ε-region is bounded by a circular arc and a straight line. The slope of the straight line is $r = -1/\tan\frac{\theta}{2}$, and is independent of ε. If $r \in \mathbb{Q}(\sqrt{2})$, then the ε-region is "parallel" to the grid lines of slope r, as shown in Figure 5(a). In this case, the number of solutions to the grid problem is not dominated by the *area*, but by the *width* of the ε-region. More specifically, for the scaled grid problem of Algorithm 7.6 to have a solution for a given denominator exponent k, it is a necessary condition that the ε-region intersects some grid line. The distance between the grid lines scales as $\frac{1}{2^k}$, and the width of the ε-region is $\frac{\varepsilon^2}{2}$, and therefore, we expect the

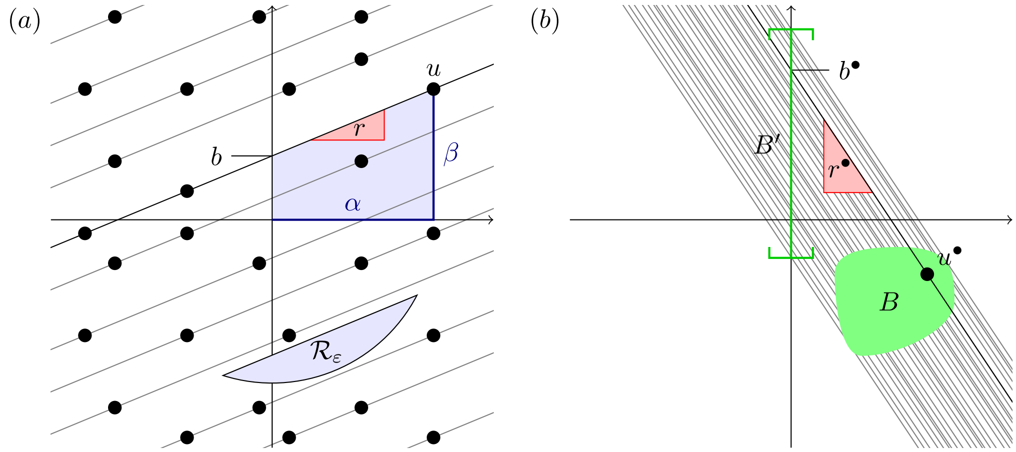


Figure 5: Part (a) shows the grid points and grid lines (of slope $r \in \mathbb{Q}(\sqrt{2})$) for a convex set B. Part (b) shows the set B and the dual grid lines (of slope $r^\bullet$), all of whose y-intercepts lie in an interval B'.

ε-region to intersect at least one grid line when $\frac{1}{2^k} \leqslant c\varepsilon^2$, for some constant c. Equivalently $k \geqslant K + 2\log\frac{1}{\varepsilon}$. By Lemmas 7.3 and 9.7, the T-count scales as $2k$, i.e., $K + 4\log\frac{1}{\varepsilon}$.

On the other hand, if $r \notin \mathbb{Q}(\sqrt{2})$, then the set of grid lines is dense, and in this case, it appears that the solution to the grid problem is dominated by the *area* of the ε-region. In this case, the number of grid points per area scales as 4^k, and the area of the ε-region is proportional to ε^3; hence we expect a grid point to fall in the ε-region when $\frac{1}{4^k} \leqslant c\varepsilon^3$, or equivalently, $2k \geqslant K + 3\log\frac{1}{\varepsilon}$, yielding the typical T-count of $K + 3\log\frac{1}{\varepsilon}$ in that case.

8.4 Time complexity of the algorithm

Proposition 8.11 (Complexity)**.** *Algorithm 7.6 runs in expected time $O(\mathrm{polylog}(1/\varepsilon))$. This is true whether or not a factorization oracle is used.*

Proof. Let M be the uprightness of the ε-region. Let j_0 be the average number of candidates tried in steps 2(a)–(c) of the algorithm, and let k_{j_0} be the least denominator exponent of the final candidate. Let n be the largest integer that appears in step 2(a) of the algorithm.

By Proposition 5.22, step 1 of the algorithm requires $O(\log(1/M))$ arithmetic operations, plus a constant number per candidate. For each of the j_0 candidates, step 2(a) requires $O(1)$ arithmetic operations. Step 2(b) also requires $O(1)$ arithmetic operations, either due to the use of a factoring oracle, or else, because we can put an arbitrary fixed bound on the amount of effort invested in factoring any given integer. At minimum, this will succeed when the integer in question is prime, which is sufficient for the estimates of Proposition 8.8. Step 2(c) requires $O(\mathrm{polylog}(n))$ operations by Theorem 6.2. Finally, step 3 requires $O(k_{j_0})$ arithmetic operations; see, e.g., Theorem 5.1 of [7]. So the total expected number of arithmetic operations is

$$O(\log(1/M)) + j_0 \cdot O(\mathrm{polylog}(n)) + O(k_{j_0}). \tag{26}$$

Recall that the ε-region $\mathcal{R}_\varepsilon$, shown in (14), contains a disk of radius $\varepsilon^2/4$. Hence, $\mathrm{area}(\mathcal{R}_\varepsilon) \geqslant \frac{\pi}{16}\varepsilon^4$. On the other hand, the square $[-1,1]\times[-1,1]$ is a (not very tight) bounding box for $\mathcal{R}_\varepsilon$. It follows that

$$M = \mathrm{up}(\mathcal{R}_\varepsilon) = \frac{\mathrm{area}(\mathcal{R}_\varepsilon)}{\mathrm{area}(\mathrm{BBox}(\mathcal{R}_\varepsilon))} = \Omega(\varepsilon^4),$$

hence $\log(1/M) = O(\log(1/\varepsilon))$. From (23), the expected value of j_0 is $O(\log(1/\varepsilon))$. Combining (17) with (22), we therefore have

$$k_{j_0} \leqslant k_2 + 2(1 + \log_2 j_0) = O(\log(1/\varepsilon)) + O(\log(\log(1/\varepsilon))) = O(\log(1/\varepsilon)).$$

From Lemma 8.5, and the fact that candidates are enumerated in order of increasing denominator exponent, we have $n \leqslant 4^{k_{j_0}}$, hence

$$\mathrm{polylog}(n) = O(\mathrm{poly}(k_{j_0})) = O(\mathrm{polylog}(1/\varepsilon)).$$

Combining all of these estimates with (26), the expected number of arithmetic operations for the algorithm is $O(\mathrm{polylog}(1/\varepsilon))$. Moreover, each individual arithmetic operation can be performed with precision $O(\log(1/\varepsilon))$, taking time $O(\text{polylog}(1/\varepsilon))$. Therefore the total expected time complexity of the algorithm is $O(\text{polylog}(1/\varepsilon))$, as desired. $\square$

9 Approximation up to a phase

So far, we have considered the problem of approximate synthesis "on the nose", i.e., the operator U in Definition 7.1 was literally required to approximate $R_z(\theta)$ in the operator norm. However, it is well-known that global phases have no observable effect in quantum mechanics, so in quantum computing, it is also common to consider the problem of approximate synthesis "up to a phase". This is made precise in the following definition.

Definition 9.1. Given θ and some $\varepsilon > 0$, the *approximate synthesis problem for z-rotations up to a phase* is to find an operator U expressible in the single-qubit Clifford+T gate set, and a unit scalar λ, such that

$$\|R_z(\theta) - \lambda U\| \leqslant \varepsilon. \tag{27}$$

Moreover, it is desirable to find U of smallest possible T-count. As before, the norm in (27) is the operator norm.

In this section, we will give a version of Algorithm 7.6 that optimally solves the approximate synthesis problem up to a phase. The central insight is that it is in fact sufficient to restrict λ to only two possible phases, namely $\lambda = 1$ and $\lambda = \sqrt{\omega} = e^{i\pi/8}$. A similar technique was also used in [11].

First, note that if W is a unitary 2×2-matrix and $\det W = 1$, then $\operatorname{tr} W$ is real. This is obvious, because $\det W = 1$ ensures that the two eigenvalues of W are each other's complex conjugates.

Lemma 9.2. *Let W be a unitary 2×2-matrix, and assume that $\det W = 1$ and $\operatorname{tr} W \geqslant 0$. Then for all unit scalars λ, we have*

$$\|I - W\| \leqslant \|I - \lambda W\|.$$

Proof. We may assume without loss of generality that W is diagonal. Since $\det W = 1$, we can write

$$W = \begin{pmatrix} e^{i\phi} & 0 \\ 0 & e^{-i\phi} \end{pmatrix}$$

for some ϕ. By symmetry, we can assume without loss of generality that $0 \leqslant \phi \leqslant \pi$. Since $\operatorname{tr} W \geqslant 0$, we have $\phi \leqslant \pi/2$. Now consider a unit scalar $\lambda = e^{i\psi}$, where $-\pi \leqslant \psi \leqslant \pi$. Then $\|I - \lambda W\| = \max\{|1 - e^{i(\psi+\phi)}|, |1 - e^{i(\psi-\phi)}|\}$ and $\|I - W\| = |1 - e^{i\phi}|$. If $\psi \geqslant 0$, then $|1 - e^{i\phi}| \leqslant |1 - e^{i(\psi+\phi)}|$. Similarly, if $\psi \leqslant 0$, then $|1 - e^{i\phi}| \leqslant |1 - e^{i(\psi-\phi)}|$. In either case, we have $\|I - W\| \leqslant \|I - \lambda W\|$, as claimed. $\square$

Lemma 9.3. *Fix ε, a unitary operator R with $\det R = 1$, and a Clifford+T operator U. The following are equivalent:*

(1) There exists a unit scalar λ such that

$$\|R - \lambda U\| \leqslant \varepsilon;$$

(2) There exists $n \in \mathbb{Z}$ such that

$$\|R - e^{in\pi/8} U\| \leqslant \varepsilon.$$

Proof. It is obvious that (2) implies (1). For the opposite implication, first note that, because U is a Clifford+T operator, we have $\det U = \omega^k$ for some $k \in \mathbb{Z}$, and therefore $\det(R^{-1}U) = \omega^k$. Let $V = e^{-ik\pi/8} R^{-1} U$, so that $\det V = 1$. If $\operatorname{tr} V \geqslant 0$, let $W = V$; otherwise, let $W = -V$. Either way, we have $W = e^{in\pi/8} R^{-1} U$, where $n \in \mathbb{Z}$, and $\det W = 1$, $\operatorname{tr} W \geqslant 0$. Let $\lambda' = e^{-in\pi/8}\lambda$. By Lemma 9.2, we have

$$\begin{array}{rrcl} & \|I - W\| & \leqslant & \|I - \lambda' W\| \\ \Longrightarrow & \|I - e^{in\pi/8} R^{-1} U\| & \leqslant & \|I - \lambda' e^{in\pi/8} R^{-1} U\| \\ \Longrightarrow & \|R - e^{in\pi/8} U\| & \leqslant & \|R - \lambda' e^{in\pi/8} U\|, \\ \Longrightarrow & \|R - e^{in\pi/8} U\| & \leqslant & \|R - \lambda U\|, \end{array}$$

which implies the desired claim. $\square$

Remark 9.4. A version of Lemma 9.3 also applies to gate sets other than Clifford+T, as long as the gate set has discrete determinants.

Corollary 9.5. *In Definition 9.1, it suffices without loss of generality to consider only the two scalars $\lambda = 1$ and $\lambda = e^{i\pi/8}$.*

Proof. Let U be a Clifford+T operator satisfying (27) for some unit scalar λ. By Lemma 9.3, there exists a λ of the form $e^{in\pi/8}$ also satisfying (27). Then we can write $\lambda = \omega^k\lambda'$, where $k \in \mathbb{Z}$ and $\lambda' \in \{1, e^{i\pi/8}\}$. Letting $U' = \omega^k U$, we have $\lambda' U' = \lambda U$, and therefore

$$\|R_z(\theta) - \lambda' U'\| \leqslant \varepsilon,$$

as claimed. Moreover, since $\omega = e^{i\pi/4}$ is a Clifford operator, U and U' have the same T-count. □

To solve the approximate synthesis problem up to a phase, we therefore need an algorithm for finding optimal solutions of (27) in case $\lambda = 1$ and $\lambda = e^{i\pi/8}$ For $\lambda = 1$, this is of course just Algorithm 7.6. So all that remains to do is to find an algorithm for solving

$$\|R_z(\theta) - e^{i\pi/8} U\| \leqslant \varepsilon. \tag{28}$$

We use a sequence of steps very similar to those of Section 7.1 to reduce this to a grid problem and a Diophantine equation. We first consider the form of U.

Lemma 9.6. *If $\varepsilon < |1 - e^{i\pi/8}|$, then all solutions of (28) have the form*

$$U = \begin{pmatrix} u & -t^\dagger\omega^{-1} \\ t & u^\dagger\omega^{-1} \end{pmatrix}. \tag{29}$$

Proof. This is completely analogous to the proof of Lemma 7.2, using $e^{i\pi/8}U$ in place of U. □

Recall that $\delta = 1 + \omega$, and note that $\frac{\delta}{|\delta|} = e^{i\pi/8}$. Also note that $\delta\omega^{-1} = \delta^\dagger$, and that the element δ is invertible in $\mathbb{D}[\omega]$ with inverse $\delta^{-1} = (\omega - i)/\sqrt{2}$. Suppose that U is of the form (29). Let $u' = \delta u$ and $t' = \delta t$. We have:

$$\begin{aligned} \|R_z(\theta) - e^{i\pi/8}U\| &= \left\| R_z(\theta) - \frac{\delta}{|\delta|}\begin{pmatrix} u & -t^\dagger\omega^{-1} \\ t & u^\dagger\omega^{-1} \end{pmatrix} \right\| \\ &= \left\| R_z(\theta) - \frac{1}{|\delta|}\begin{pmatrix} \delta u & -\delta^\dagger t^\dagger \\ \delta t & \delta^\dagger u^\dagger \end{pmatrix} \right\| \\ &= \left\| R_z(\theta) - \frac{1}{|\delta|}\begin{pmatrix} u' & -t'^\dagger \\ t' & u'^\dagger \end{pmatrix} \right\|. \end{aligned}$$

Recall the definition of the ε-region $\mathcal{R}_\varepsilon$ from (14). Using exactly the same argument as in Section 7, it follows that (28) holds if and only if $\frac{u'}{|\delta|} \in \mathcal{R}_\varepsilon$, i.e., $u' \in |\delta|\mathcal{R}_\varepsilon$.

As before, in order for U to be unitary, of course it must satisfy $u^\dagger u + t^\dagger t = 1$, and a necessary condition for this is $u, u^\bullet \in \overline{\mathcal{D}}$. The latter condition can be equivalently re-expressed in terms of u' by requiring $u' \in |\delta|\,\overline{\mathcal{D}}$ and $u'^\bullet \in |\delta^\bullet|\,\overline{\mathcal{D}}$. Therefore, finding solutions to (28) of the form (29) reduces to the two-dimensional grid problem $u' \in |\delta|\mathcal{R}_\varepsilon$ and $u'^\bullet \in |\delta^\bullet|\,\overline{\mathcal{D}}$, together with the usual Diophantine equation $u^\dagger u + t^\dagger t = 1$. The last remaining piece of the puzzle is to compute the T-count of U, and in particular, to ensure that potential solutions are found in order of increasing T-count.

Lemma 9.7. *Let U be a Clifford+T operator of the form (29), and let k be the least denominator exponent of $u' = \delta u$. Then the T-count of U is either $2k-1$ or $2k+1$. Moreover, if $k > 0$ and U has T-count $2k+1$, then $U' = TUT^\dagger$ has T-count $2k-1$.*

Proof. This can be proved by a tedious but easy induction, analogous to Lemma 7.3. □

We therefore arrive at the following algorithm for solving (28). Here we assume $\varepsilon < |1 - e^{i\pi/8}|$, so that Lemma 9.6 applies.

Algorithm 9.8. Given θ and ε, let $A = |\delta|\mathcal{R}_\varepsilon$, and let $B = |\delta^\bullet|\,\overline{\mathcal{D}}$.

1. Use Proposition 5.22 to enumerate the infinite sequence of solutions to the scaled grid problem $u' \in A$ and $u'^\bullet \in B$, where $u' \in \mathbb{D}[\omega]$, in the order of increasing least denominator exponent k.
2. For each such solution u':
   (a) Let $u = u'/\delta$, let $\xi = 1 - u^\dagger u \in \mathbb{D}[\sqrt{2}]$, and write $\xi^\bullet\xi = \frac{n}{2^\ell}$, where $n \in \mathbb{Z}$ and $\ell \geqslant 0$ is minimal.
   (b) Attempt to find a prime factorization of n. If $n \neq 0$ but no prime factorization is found, skip step 2(c) and continue with the next u'.

(c) Use the algorithm of Theorem 6.2 to solve the equation $t^\dagger t = \xi$. If a solution t exists, go to step 3; otherwise, continue with the next u'.

3. Define U as in equation (29), let $U' = TUT^\dagger$, and use the exact synthesis algorithm of [10] to find a Clifford+T circuit implementing either U or U', whichever has smaller T-count. Output this circuit and stop.

Algorithm 9.8 is optimal in the presence of a factoring oracle, and near-optimal in the absence of a factoring oracle, in the same sense as Algorithm 7.6. Its expected time complexity is $O(\text{polylog}(1/\varepsilon))$. The proofs are completely analogous to those of Section 8. We then arrive at the following composite algorithm for the approximate synthesis problem for z-rotations up to a phase:

Algorithm 9.9 (Approximate synthesis up to a phase)**.** Given θ and ε, run both Algorithms 7.6 and 9.8, and return whichever circuit has the smaller T-count.

Proposition 9.10 (Correctness, time complexity, and optimality)**.** *Algorithm 9.9 yields a valid solution to the approximate synthesis problem up to a phase. Its expected running time is $O(\text{polylog}(1/\varepsilon))$. In the presence of a factoring oracle, the algorithm is optimal, i.e., the returned circuit has the smallest T-count of any single-qubit Clifford+T circuit approximating $R_z(\theta)$ up to ε and up to a phase. Moreover, in the absence of a factoring oracle, the algorithm is near-optimal in the following sense. Let m be the T-count of the solution found. Then:*

(a) The approximate synthesis problem has an expected number of at most $O(\log(1/\varepsilon))$ non-equivalent solutions with T-count less than m.

(b) The expected value of m is $m''' + O(\log(\log(1/\varepsilon)))$, where m''' is the T-count of the third-to-optimal solution (up to equivalence) of the approximate synthesis problem up to a phase.

Proof. The correctness and time complexity of Algorithm 9.9 follows from that of Algorithms 7.6 and 9.8. The optimality results follow from those of Algorithms 7.6 and 9.8, keeping in mind that Algorithm 7.6 finds an optimal (or near-optimal) solution for the phase $\lambda = 1$, Algorithm 9.8 finds an optimal (or near-optimal) solution for the phase $\lambda = e^{i\pi/8}$, and by Corollary 9.5, these are the only two phases that need to be considered.

The only subtlety that must be pointed out is that in part (b) of the near-optimality, we use the T-count of the *third*-to-optimal solution, rather than the second-to-optimal one as in Proposition 8.8. This is because the optimal and second-to-optimal solutions may belong to Algorithms 7.6 and 9.8, respectively, so that it may not be until the third-to-optimal solution that the near-optimality result of either Algorithm 7.6 or Algorithm 9.8 can be invoked. ☐

Remark 9.11. Algorithms 7.6 and 9.8 share the same ε-region up to scaling, and therefore the uprightness computation of Theorem 5.16 only needs to be done once.

Remark 9.12. By Lemmas 7.3 and 9.7, Algorithm 7.6 always produces circuits with even T-count, and Algorithm 9.8 always produces circuits with odd T-count. Instead of running both algorithms to completion, it is possible to interleave the two algorithms, so that all potential solutions are considered in order of increasing T-count. This is a slight optimization which does not, however, affect the asymptotic time complexity.

10 Experimental results

We implemented Algorithm 7.6 in the programming language Haskell; the implementation is available from [13]. Running on a single core of a 3.4GHz Intel i5-3570 CPU, we approximated the operator $R_z(\pi/128)$ up to various ε. The results are summarized in the leftmost four columns of Table 1. For example, here is the approximation of $R_z(\pi/128)$ up to $\varepsilon = 10^{-10}$:

$$\begin{aligned}
u &= \frac{1}{\sqrt{2}^{52}}(-26687414\omega^3 + 10541729\omega^2 + 10614512\omega + 40727366) \\
t &= \frac{1}{\sqrt{2}^{52}}(30805761\omega^3 - 23432014\omega^2 + 2332111\omega + 20133911) \\
U &= \mathtt{HTSHTSHTSHTHTHTHTSHTHTSHTSHTSHTHTHTSHTSHTHTHTSHTHTSHTHTHTHTHTHTHTSHTSHT} \\
&\quad\ \mathtt{SHTHTSHTHTSHTHTHTHTSHTHTHTSHTHTSHTHTHTHTSHTSHTSHTHTHTSHTSHTSHTSHTHTSHTS} \\
&\quad\ \mathtt{HTSHTSHTHTSHTHTSHTSHTHTHTHTHTSHTHTHTHTSHTSHTSHTHTSHTSHTHTHTSHTHTHTHTHTS} \\
&\quad\ \mathtt{HTSHTHTHTHTHTSHTHTHTHTSHTHTHTHTHTHTH}\omega^7
\end{aligned}$$

In addition to ε, the T-count of the computed operator, and the actual error, we have also reported the lower bound on the T-count, which was computed according to Remark 8.9. It can be seen that the actual T-count achieved by

ε	T-count	T-bound	Actual error	Runtime	Candidates	Time/Candidate
10^{-10}	102	$\geqslant 102$	$0.91180 \cdot 10^{-10}$	$0.0190s$	3.0	$0.0064s$
10^{-20}	200	$\geqslant 198$	$0.87670 \cdot 10^{-20}$	$0.0433s$	7.0	$0.0061s$
10^{-30}	298	$\geqslant 298$	$0.99836 \cdot 10^{-30}$	$0.0600s$	7.0	$0.0085s$
10^{-40}	402	$\geqslant 400$	$0.77378 \cdot 10^{-40}$	$0.0976s$	11.7	$0.0084s$
10^{-50}	500	$\geqslant 500$	$0.82008 \cdot 10^{-50}$	$0.1353s$	20.3	$0.0067s$
10^{-60}	602	$\geqslant 596$	$0.61151 \cdot 10^{-60}$	$0.1548s$	16.0	$0.0097s$
10^{-70}	702	$\geqslant 698$	$0.40936 \cdot 10^{-70}$	$0.1931s$	20.9	$0.0093s$
10^{-80}	804	$\geqslant 794$	$0.92372 \cdot 10^{-80}$	$0.2402s$	27.2	$0.0088s$
10^{-90}	898	$\geqslant 898$	$0.96607 \cdot 10^{-90}$	$0.2696s$	22.2	$0.0121s$
10^{-100}	1000	$\geqslant 998$	$0.78879 \cdot 10^{-100}$	$0.3443s$	31.2	$0.0110s$
10^{-200}	1998	$\geqslant 1994$	$0.73266 \cdot 10^{-200}$	$1.1423s$	62.3	$0.0183s$
10^{-500}	4990	$\geqslant 4986$	$0.67156 \cdot 10^{-500}$	$8.6509s$	170.4	$0.0508s$
10^{-1000}	9974	$\geqslant 9966$	$0.80457 \cdot 10^{-1000}$	$47.9300s$	270.4	$0.1773s$
10^{-2000}	19942	$\geqslant 19934$	$0.88272 \cdot 10^{-2000}$	$383.1024s$	556.7	$0.6881s$

Table 1: Experimental results. The first four columns report the T-count, computed lower bound on the T-count, and the actual error for approximating the operator $R_z(\pi/128)$ up to various ε. The remaining columns report the runtime for each ε, averaged over 50 independent runs of Algorithm 7.6 with random angles θ. The runtime is further broken down into average number of candidates tried per run of the algorithm, and time spent per candidate.

	$\varepsilon = 10^{-100}$	$\varepsilon = 10^{-200}$	$\varepsilon = 10^{-300}$	$\varepsilon = 10^{-400}$	$\varepsilon = 10^{-500}$
$\theta = 2\tan^{-1}(2+3\sqrt{2})$	1320	2646	3976	5306	6636
$\theta = 2\tan^{-1}(5/3)$	1314	2646	3966	5306	6636
$\theta = 2\tan^{-1}((2+7\sqrt{2})/5)$	1308	2642	3968	5294	6630
$\theta = 2\tan^{-1}(2+3\sqrt{3})$	1002	1996	2994	3992	4990
$\theta = 2\tan^{-1}(\sqrt{5})$	998	1998	2996	3990	4990
$\theta = 2\tan^{-1}(1+11\sqrt{7})$	1000	1998	2996	3992	4990
$4\log_2(1/\varepsilon)$	1329	2658	3986	5315	6644
$3\log_2(1/\varepsilon)$	997	1993	2990	3986	4983

Table 2: Worst-case and typical T-counts. The top part of the table shows the T-counts for various combinations of θ and ε. The bottom part of the table shows the actual values of $4\log_2(1/\varepsilon)$ and $3\log_2(1/\varepsilon)$ for each ε, rounded to the closest integer. As predicted by Conjecture 8.10, the T-counts are close to $4\log_2(1/\varepsilon)$ when $\theta = 2\tan^{-1}(r)$, for $r \in \mathbb{Q}(\sqrt{2})$, and close to $3\log_2(1/\varepsilon)$ otherwise.

the algorithm exceeds this lower bound by at most a very small amount, which is consistent with $O(\log(\log(1/\varepsilon)))$ as predicted by Proposition 8.8(b). This excess could be further reduced by increasing the amount of effort spent on factoring, and will become zero in the presence of a factoring oracle.

Table 1 also shows the runtime as a function of ε. Since the enumeration of solutions to the grid problem in the algorithm is deterministic, the algorithm tends to find the same one or two solutions each time it is run with the same parameters. For this reason, we have averaged the runtimes in Table 1 over 50 runs of the algorithm with random angles θ, for each ε. As shown in Table 1, we were able to achieve approximations up to $\varepsilon = 10^{-1000}$ with a T-count of under 10000 in less than 50 seconds on average. This compares to a T-count of 13300 and an average runtime of 504.8 seconds reported in [14] on the same hardware. We also note that the experimental runtimes in Table 1 are consistent with the polynomial runtime predicted by Proposition 8.11, and appear to be $O(\log^3(1/\varepsilon))$.

Table 2 shows T-counts as a function of θ and ε. As predicted by Conjecture 8.10, the T-counts are close to $4\log_2(1/\varepsilon)$ when $\theta = 2\tan^{-1}(r)$, for $r \in \mathbb{Q}(\sqrt{2})$, and close to $3\log_2(1/\varepsilon)$ otherwise.

11 Conclusion

We have presented a fast new probabilistic algorithm for approximating arbitrary single-qubit z-rotations by Clifford+T circuits. Our algorithm is optimal in the presence of a factoring oracle, i.e., it finds the shortest possible circuit whatsoever for the given problem instance. In the absence of a factoring oracle, our algorithm is still nearly optimal. In particular, no quantum computer is required to run our algorithm. The main technical innovation of this paper is a new efficient algorithm for solving two-dimensional grid problems, such as the ones that arise in candidate selection for approximate synthesis. We solved this problem by an iterative method that successively increases the uprightness of a pair of convex sets until the problem is in a form where it can be solved directly.

It is an interesting question whether a similar algorithm can be found for giving optimal or near-optimal approximations of arbitrary single-qubit operators. In its current form, our method only applies to z-rotations. Since any arbitrary single-qubit operator can be decomposed into three z-rotations using Euler angles, our algorithm can currently achieve a T-count of $9\log_2(1/\varepsilon)+O(\log(\log(1/\varepsilon)))$. (Interestingly, the average case gate complexity can always be achieved in this situation, because if the operator to be approximated happens to exhibit worst-case behavior, we can always change it by multiplying it by a small number of random Clifford+T gates). However, the information-theoretic lower bound in this situation remains $K+3\log_2(1/\varepsilon)$, so there is still potential for improvement.

12 Acknowledgements

This research was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC). This research was supported by the Intelligence Advanced Research Projects Activity (IARPA) via Department of Interior National Business Center contract number D12PC00527. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. Disclaimer: The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, DoI/NBC, or the U.S. Government.

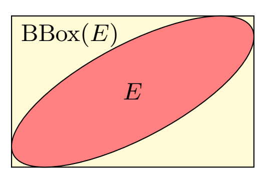
A Proof of Theorem 5.16

This appendix contains a proof of Theorem 5.16. We start by reformulating the problem in more convenient terms. Recall that the notion of uprightness introduced in Section 5.2 was defined for an arbitrary bounded convex subset of $\mathbb{R}^2$. If the set in question is an ellipse, we can expand the definition of uprightness into an explicit expression. Recall from Definition 5.15 that an ellipse centered at p can be written as $E=\{u\in\mathbb{R}^2 \mid (u-p)^\dagger D(u-p)\leqslant 1\}$, where D is a positive definite matrix whose entries are, e.g., as follows:

$$D=\begin{bmatrix} a & b \\ b & d \end{bmatrix}.$$

We can compute the area of E and the area of its bounding box using D. Indeed, we have $\mathrm{area}(E)=\pi/\sqrt{\det(D)}$ and $\mathrm{area}(\mathrm{BBox}(E))=4\sqrt{ad}/\det(D)$. Substituting these in Definition 5.7 yields the desired expression for uprightness:

$$\mathrm{up}(E)=\frac{\mathrm{area}(E)}{\mathrm{area}(\mathrm{BBox}(E))}=\frac{\pi}{4}\sqrt{\frac{\det(D)}{ad}}. \tag{30}$$



It follows that the uprightness of E is invariant under translation and scalar multiplication.

Recall that $\lambda=\sqrt{2}+1$. The matrix D corresponding to an ellipse E has determinant 1 if and only if it can be written in the form

$$D=\begin{bmatrix} e\lambda^{-z} & b \\ b & e\lambda^{z} \end{bmatrix} \tag{31}$$

for some $b,e,z\in\mathbb{R}$ with $e>0$ and $e^2=b^2+1$. In this case, the definition of uprightness (30) simplifies to

$$\mathrm{up}(E)=\frac{\pi}{4e^2}=\frac{\pi}{4\sqrt{b^2+1}}. \tag{32}$$

Equivalently, if $\mathrm{up}(E) = M$, then

$$b^2 = \frac{\pi^2}{16M^2} - 1. \tag{33}$$

Since Theorem 5.16 deals with pairs of ellipses, it is convenient to introduce the following terminology for discussing pairs of matrices.

Definition A.1. A *state* is a pair of real symmetric positive definite matrices of determinant 1. Given a state (D, Δ) with

$$D = \begin{bmatrix} e\lambda^{-z} & b \\ b & e\lambda^{z} \end{bmatrix} \qquad \Delta = \begin{bmatrix} \varepsilon\lambda^{-\zeta} & \beta \\ \beta & \varepsilon\lambda^{\zeta} \end{bmatrix} \tag{34}$$

we define its *skew* as $\texttt{Skew}(D, \Delta) = b^2 + \beta^2$ and its *bias* as $\texttt{Bias}(D, \Delta) = \zeta - z$.

Note that the skew of a state is small if and only if both b^2 and β^2 are small, which happens, by (32), if and only if the ellipses corresponding to D and Δ both have large uprightness. So our strategy for increasing the uprightness will be to reduce the skew. In what follows, we use (D, Δ) to denote an arbitrary state and always assume that the entries of D and Δ are given as in (34). For future reference, we record here another useful property of states.

Remark A.2. If (D, Δ) is a state with $b \geqslant 0$, then $-be \leqslant -b^2$. Indeed:

$$e^2 = b^2 + 1 \implies e^2 \geqslant b^2 \implies e \geqslant b \implies -be \leqslant -b^2.$$

Similarly, if $b \leqslant 0$, then $be \leqslant -b^2$. Analogous inequalities also hold for β and ε.

The action of a grid operator on an ellipse can be adapted to states in a natural way, provided that the operator is special.

Definition A.3. The action of special grid operators on states is defined as follows. Here, $G^\dagger$ denotes the transpose of G, and $G^\bullet$ is defined by applying $(-)^\bullet$ separately to each matrix entry, as in Remark 5.12.

$$(D, \Delta) \cdot G = (G^\dagger D G, G^{\bullet\dagger} \Delta G^\bullet).$$

Lemma A.4. *Let (D, Δ) be a state, and let A and B be the ellipses centered at the origin that are defined by D and Δ, respectively. Then the ellipses $G(A)$ and $G^\bullet(B)$ are defined by the matrices D' and Δ', where*

$$(D', \Delta') = (D, \Delta) \cdot G^{-1}$$

Proof. We have

$$\begin{aligned} G(A) &= \{G(u) \in \mathbb{R}^2 \mid u^\dagger D u \leqslant 1\} \\ &= \{v \in \mathbb{R}^2 \mid (G^{-1}v)^\dagger D (G^{-1}v) \leqslant 1\} \\ &= \{v \in \mathbb{R}^2 \mid v^\dagger (G^{-1})^\dagger D G^{-1} v \leqslant 1\}, \end{aligned}$$

so the ellipse $G(A)$ is defined by the positive operator $D' = (G^{-1})^\dagger D G^{-1}$. The proof for $G^\bullet(B)$ is similar. □

The main ingredient in the proof of Theorem 5.16 is the following Step Lemma.

Lemma A.5 (Step Lemma)**.** *For any state (D, Δ), if $\texttt{Skew}(D, \Delta) \geqslant 15$, then there exists a special grid operator G such that $\texttt{Skew}((D, \Delta) \cdot G) \leqslant 0.9\ \texttt{Skew}(D, \Delta)$. Moreover, G can be computed using a constant number of arithmetic operations.*

Before proving the Step Lemma, we show how it can be used to derive Theorem 5.16, whose statement we reproduce here.

Theorem. *Suppose $A, B \subseteq \mathbb{R}^2$ are ellipses. Then there exists a grid operator G such that $G(A)$ and $G^\bullet(B)$ are $1/6$-upright. Moreover, if A and B are M-upright, then G can be efficiently computed in $O(\log(1/M))$ arithmetic operations.*

Proof. Let D and Δ be the matrices defining A and B respectively, in the sense of Definition 5.15. Since uprightness is invariant under translations and scaling, we may without loss of generality assume that both ellipses are centered at the origin, and that $\det D = \det \Delta = 1$.

The pair (D, Δ) is a state. By applying Lemma A.5 repeatedly, we get grid operators $G_1, \ldots, G_n$ such that:

$$\texttt{Skew}((D, \Delta) \cdot G_1 \ldots G_n) \leqslant 15. \tag{35}$$

$$R = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \quad A = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 1 & \sqrt{2} \\ 0 & 1 \end{bmatrix}$$

$$K = \frac{1}{\sqrt{2}}\begin{bmatrix} -\lambda^{-1} & -1 \\ \lambda & 1 \end{bmatrix} \quad X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Figure 6: List of useful grid operators.

Now let $(D', \Delta') = (D, \Delta) \cdot G_1 \dots G_n$ and set $G = (G_1 \cdots G_n)^{-1}$. By Lemma A.4, the ellipses $G(A)$ and $G^\bullet(B)$ are defined by the matrices D' and Δ', respectively. Let b and β be the anti-diagonal entries of the matrices D' and Δ', respectively. We have:

$$b^2 + \beta^2 = \texttt{Skew}(D', \Delta') = \texttt{Skew}((D, \Delta) \cdot G^{-1}) = \texttt{Skew}((D, \Delta) \cdot G_1 \dots G_n) \leqslant 15,$$

hence $b^2 \leqslant 15$ and $\beta^2 \leqslant 15$. Using (32), we get

$$\mathrm{up}(G(A)) = \frac{\pi}{4\sqrt{b^2+1}} \geqslant \frac{\pi}{4\sqrt{16}} \geqslant 1/6 \quad \text{and} \quad \mathrm{up}(G^\bullet(B)) = \frac{\pi}{4\sqrt{\beta^2+1}} \geqslant \frac{\pi}{4\sqrt{16}} \geqslant 1/6,$$

as desired.

To bound the number of operations, note that each application of G_j reduces the skew by at least 10 percent. Therefore, the number n in (35) satisfies $n \leqslant \log_{0.9}(15/\texttt{Skew}(D, \Delta)) = O(\log(\texttt{Skew}(D, \Delta)))$. Using (33), we have

$$\log(\texttt{Skew}(D, \Delta)) = \log(b^2 + \beta^2) \leqslant \log((\frac{\pi^2}{16M^2} - 1) + (\frac{\pi^2}{16M^2} - 1)) = O(\log(1/M)).$$

It follows that the computation of G requires $O(\log(1/M))$ applications of the Step Lemma, each of which requires a constant number of arithmetic operations, proving the final claim of the theorem. □

The remainder of this appendix is devoted to proving the Step Lemma. To each state, we associate the pair (z, ζ). The proof of the Step Lemma is essentially a case distinction on the location of the pair (z, ζ) in the plane. We find coverings of the plane with the property that if the point (z, ζ) belongs to some region $\mathcal{O}$ of our covering, then it is easy to compute a special grid operator G such that $\texttt{Skew}((D, \Delta) \cdot G) \leqslant 0.9\ \texttt{Skew}(D, \Delta)$. The relevant grid operators are given in Figure 6. Each one of the next 5 subsections is dedicated to a particular region of the plane. We prove the Step Lemma in Section A.6.

A.1 The Shift Lemma

In this section, we consider states (D, Δ) such that $|\texttt{Bias}(D, \Delta)| > 1$. Any such state can be "shifted" to a state (D', Δ') of equal skew but with $|\texttt{Bias}(D', \Delta')| \leqslant 1$.

Definition A.6. The *shift operators* σ and τ are defined by:

$$\sigma = \sqrt{\lambda^{-1}}\begin{bmatrix} \lambda & 0 \\ 0 & 1 \end{bmatrix}, \tau = \sqrt{\lambda^{-1}}\begin{bmatrix} 1 & 0 \\ 0 & -\lambda \end{bmatrix}$$

Even though σ and τ are not grid operators, we can use them to define an operation on states called a *shift by* k. By abuse of notation, we write this operation as an action.

Definition A.7. Given a state (D, Δ) and $k \in \mathbb{Z}$, the *k-shift of* (D, Δ) is defined as:

$$(D, \Delta) \cdot \mathrm{Shift}^k = (\sigma^k D \sigma^k, \tau^k \Delta \tau^k).$$

The notation $(D, \Delta) \cdot \mathrm{Shift}^k$ is justified by the following lemma.

Lemma A.8. *The shift of a state is a state and moreover:*

$$\texttt{Skew}((D, \Delta) \cdot \mathit{Shift}^k) = \texttt{Skew}(D, \Delta) \quad \textit{and} \quad \texttt{Bias}((D, \Delta) \cdot \mathit{Shift}^k) = \texttt{Bias}(D, \Delta) + 2k$$

Proof. Compute $(D,\Delta)\cdot \text{Shift}^k$:

$$\begin{aligned}
(D,\Delta)\cdot \text{Shift}^k &= (\sigma^k D\sigma^k, \tau^k\Delta\tau^k)\\
&= (\sigma^k \begin{bmatrix} e\lambda^{-z} & b \\ b & e\lambda^{z}\end{bmatrix}\sigma^k, \tau^k \begin{bmatrix} \varepsilon\lambda^{-\zeta} & \beta \\ \beta & \varepsilon\lambda^{\zeta}\end{bmatrix}\tau^k)\\
&= (\begin{bmatrix} e\lambda^{-z+k} & b \\ b & e\lambda^{z-k}\end{bmatrix}, \begin{bmatrix} \varepsilon\lambda^{-\zeta-k} & (-1)^k\beta \\ (-1)^k\beta & \varepsilon\lambda^{\zeta+k}\end{bmatrix})
\end{aligned}$$

The resulting matrices are clearly symmetric and positive definite. Moreover, since σ^k and τ^k have determinant ± 1, both $\sigma^k D\sigma^k$ and $\tau^k\Delta\tau^k$ have determinant 1. Finally:

- $\texttt{Skew}((D,\Delta)\cdot\text{Shift}^k) = b^2 + ((-1)^k\beta)^2 = b^2+\beta^2 = \texttt{Skew}(D,\Delta)$ and
- $\texttt{Bias}((D,\Delta)\cdot\text{Shift}^k) = (\zeta+k)-(z-k) = \texttt{Bias}(D,\Delta)+2k$,

which completes the proof. □

For every special grid operator G, there is a special grid operator G' whose action on a state corresponds to shifting the state by k, applying G and then shifting the state by $-k$.

Lemma A.9. *If G is a special grid operator and $k\in\mathbb{Z}$, then $G' = \sigma^k G\sigma^k$ is a special grid operator and moreover $G'^\bullet = (-\tau)^k G^\bullet \tau^k$.*

Proof. It suffices to show this for $k=1$. Suppose $G = \begin{bmatrix} w & x \\ y & z\end{bmatrix}$ is a special grid operator and note that:

$$G' = \sigma G\sigma = \begin{bmatrix} \lambda w & x \\ y & \lambda^{-1}z\end{bmatrix} = \begin{bmatrix} \lambda^{-1} & 0 \\ 0 & \lambda^{-1}\end{bmatrix}\begin{bmatrix} \lambda & 0 \\ 0 & 1\end{bmatrix} G \begin{bmatrix} \lambda & 0 \\ 0 & 1\end{bmatrix}.$$

Since all the factors in the above product are grid operators, the result is also a grid operator. Moreover $\det(\sigma G\sigma) = \det(G) = 1$ so that $\sigma G\sigma$ is special. Finally:

$$G'^\bullet = (\sigma G\sigma)^\bullet = \begin{bmatrix} \lambda^\bullet w^\bullet & x^\bullet \\ y^\bullet & (\lambda^{-1})^\bullet z^\bullet\end{bmatrix} = \begin{bmatrix} -\lambda^{-1} w^\bullet & x^\bullet \\ y^\bullet & -\lambda z^\bullet\end{bmatrix} = -\tau G^\bullet\tau.$$

□

Lemma A.10. *If G is a grid operator, then:*

$$(((D,\Delta)\cdot \mathit{Shift}^k)\cdot G)\cdot \mathit{Shift}^k = (D,\Delta)\cdot(\sigma^k G\sigma^k).$$

Proof. Write $G' = \sigma^k G\sigma^k$. Simple computation then yields the result:

$$\begin{aligned}
(((D,\Delta)\cdot\text{Shift}^k)\cdot G)\cdot\text{Shift}^k &= ((\sigma^k D\sigma^k,\, \tau^k\Delta\tau^k)\cdot G)\cdot\text{Shift}^k\\
&= (G^\dagger\sigma^k D\sigma^k G,\, G^{\bullet\dagger}\tau^k\Delta\tau^k G^\bullet)\cdot\text{Shift}^k\\
&= (\sigma^k G^\dagger\sigma^k D\sigma^k G\sigma^k,\, \tau^k G^{\bullet\dagger}\tau^k\Delta\tau^k G^\bullet\tau^k)\\
&= (\sigma^k G^\dagger\sigma^k D\sigma^k G\sigma^k,\, ((-\tau)^k G^{\bullet\dagger}\tau^k)\Delta((-\tau)^k G^\bullet\tau^k))\\
&= (G'^\dagger D G',\, G'^{\bullet\dagger}\Delta G'^\bullet)\\
&= (D,\Delta)\cdot G'\\
&= (D,\Delta)\cdot(\sigma^k G\sigma^k).
\end{aligned}$$

□

Shifts allow us to consider only states (D,Δ) with $\texttt{Bias}(D,\Delta)\in[-1,1]$ in the proof of the Step Lemma.

Lemma A.11. *If the Step Lemma holds for all states (D,Δ) with $\texttt{Bias}(D,\Delta)\in[-1,1]$, then it holds for all states.*

Proof. Let (D,Δ) be some state with $\texttt{Skew}(D,\Delta)\geqslant 15$. Let $x = \texttt{Bias}(D,\Delta)$ and set $k = \lfloor \frac{1-x}{2}\rfloor$. Then by Lemma A.8, we have $\texttt{Skew}((D,\Delta)\cdot\text{Shift}^k) = \texttt{Skew}(D,\Delta)$ and $\texttt{Bias}((D,\Delta)\cdot\text{Shift}^k)\in[-1,1]$. Then by assumption, there exists a special grid operator G such that $\texttt{Skew}(((D,\Delta)\cdot\text{Shift}^k)\cdot G)\leqslant 0.9\,\texttt{Skew}((D,\Delta)\cdot\text{Shift}^k)$. Now by Lemma A.9 we know that $G' = \sigma^k\, G\,\sigma^k$ is a special grid operator. Moreover, by Lemma A.10 and A.8, we have:

$$\begin{aligned}
\texttt{Skew}((D,\Delta)\cdot G') &= \texttt{Skew}((((D,\Delta)\cdot\text{Shift}^k)\cdot G)\cdot\text{Shift}^k)\\
&= \texttt{Skew}(((D,\Delta)\cdot\text{Shift}^k)\cdot G)\\
&\leqslant 0.9\,\texttt{Skew}((D,\Delta)\cdot\text{Shift}^k)\\
&= 0.9\,\texttt{Skew}(D,\Delta),
\end{aligned}$$

which completes the proof. □

A.2 The R Lemma

Definition A.12. The *hyperbolic sine in base* λ is defined as:

$$\sinh_\lambda(x) = \frac{\lambda^x - \lambda^{-x}}{2}.$$

Lemma A.13. *Recall the operator R from Figure 6. If (D,Δ) is such that $\texttt{Skew}(D,\Delta) \geqslant 15$, and such that $-0.8 \leqslant z \leqslant 0.8$ and $-0.8 \leqslant \zeta \leqslant 0.8$, then:*

$$\texttt{Skew}((D,\Delta)\cdot R) \leqslant 0.9\ \texttt{Skew}(D,\Delta).$$

Proof. Compute the action of R on (D,Δ):

$$\begin{aligned}
R^\dagger D R &= \frac{1}{2}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} e\lambda^{-z} & b \\ b & e\lambda^{z} \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \\
&= \begin{bmatrix} \cdots & \frac{e(\lambda^z-\lambda^{-z})}{2} \\ \frac{e(\lambda^z-\lambda^{-z})}{2} & \cdots \end{bmatrix} = \begin{bmatrix} \cdots & e\sinh_\lambda(z) \\ e\sinh_\lambda(z) & \cdots \end{bmatrix}, \\
R^{\bullet\dagger}\Delta R^\bullet &= \frac{1}{2}\begin{bmatrix} -1 & -1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} \varepsilon\lambda^{-\zeta} & \beta \\ \beta & \varepsilon\lambda^{\zeta} \end{bmatrix}\begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix} \\
&= \begin{bmatrix} \cdots & \frac{\varepsilon(\lambda^\zeta-\lambda^{-\zeta})}{2} \\ \frac{\varepsilon(\lambda^\zeta-\lambda^{-\zeta})}{2} & \cdots \end{bmatrix} = \begin{bmatrix} \cdots & \varepsilon\sinh_\lambda(\zeta) \\ \varepsilon\sinh_\lambda(\zeta) & \cdots \end{bmatrix}.
\end{aligned}$$

Therefore $\texttt{Skew}((D,\Delta)\cdot R) = e^2\sinh^2_\lambda(z) + \varepsilon^2\sinh^2_\lambda(\zeta)$. But recall that $e^2 = b^2+1$ and $\varepsilon^2 = \beta^2+1$, so that in fact:

$$\texttt{Skew}((D,\Delta)\cdot R) = (b^2+1)\sinh^2_\lambda(z) + (\beta^2+1)\sinh^2_\lambda(\zeta).$$

We assumed $-0.8 \leqslant z,\zeta \leqslant 0.8$ and this implies that $\sinh^2_\lambda(\zeta), \sinh^2_\lambda(z) \leqslant \sinh^2_\lambda(0.8)$. Writing $y = \sinh^2_\lambda(0.8)$ for brevity, and using the assumption that $\texttt{Skew}(D,\Delta) \geqslant 15$, we get:

$$\begin{aligned}
\texttt{Skew}((D,\Delta)\cdot R) &= (b^2+1)\sinh^2_\lambda(z) + (\beta^2+1)\sinh^2_\lambda(\zeta) \\
&\leqslant (b^2+1)y + (\beta^2+1)y \\
&= (b^2+\beta^2+2)y \\
&\leqslant \texttt{Skew}(D,\Delta)(1+\tfrac{2}{15})y.
\end{aligned}$$

This completes the proof, since $(1+\frac{2}{15})y = (1+\frac{2}{15})\sinh^2_\lambda(0.8) \approx 0.663 \leqslant 0.9$. ◻

A.3 The K Lemma

Definition A.14. The *hyperbolic cosine in base* λ is defined as:

$$\cosh_\lambda(x) = \frac{\lambda^x + \lambda^{-x}}{2}.$$

Lemma A.15. *Recall the operator K from Figure 6. If (D,Δ) is such that $\texttt{Bias}(D,\Delta) \in [-1,1]$, $\texttt{Skew}(D,\Delta) \geqslant 15$, and such that $b,\beta \geqslant 0$, $z \leqslant 0.3$, and $0.8 \leqslant \zeta$, then:*

$$\texttt{Skew}((D,\Delta)\cdot K) \leqslant 0.9\ \texttt{Skew}(D,\Delta).$$

Proof. Compute the action of K on (D,Δ):

$$\begin{aligned}
&K^\dagger D K \\
&= \frac{1}{2}\begin{bmatrix} -\lambda^{-1} & \lambda \\ -1 & 1 \end{bmatrix}\begin{bmatrix} e\lambda^{-z} & b \\ b & e\lambda^{z} \end{bmatrix}\begin{bmatrix} -\lambda^{-1} & -1 \\ \lambda & 1 \end{bmatrix} \\
&= \frac{1}{2}\begin{bmatrix} \cdots & e(\lambda^{z+1}+\lambda^{-z-1}) - 2\sqrt{2}b \\ e(\lambda^{z+1}+\lambda^{-z-1}) - 2\sqrt{2}b & \cdots \end{bmatrix} \\
&= \begin{bmatrix} \cdots & e\cosh_\lambda(z+1) - \sqrt{2}b \\ e\cosh_\lambda(z+1) - \sqrt{2}b & \cdots \end{bmatrix},
\end{aligned}$$

$$
\begin{aligned}
&K^{\bullet\dagger}\Delta K^{\bullet} \\
&\quad= \frac{1}{2}\begin{bmatrix} \lambda & -\lambda^{-1} \\ -1 & 1 \end{bmatrix}\begin{bmatrix} \varepsilon\lambda^{-\zeta} & \beta \\ \beta & \varepsilon\lambda^{\zeta} \end{bmatrix}\begin{bmatrix} \lambda & -1 \\ -\lambda^{-1} & 1 \end{bmatrix} \\
&\quad= \frac{1}{2}\begin{bmatrix} \dots & -\varepsilon(\lambda^{\zeta-1}+\lambda^{-\zeta+1})+2\sqrt{2}\beta \\ -\varepsilon(\lambda^{\zeta-1}+\lambda^{-\zeta+1})+2\sqrt{2}\beta & \dots \end{bmatrix} \\
&\quad= \begin{bmatrix} \dots & \sqrt{2}\beta-\varepsilon\cosh_\lambda(\zeta-1) \\ \sqrt{2}\beta-\varepsilon\cosh_\lambda(\zeta-1) & \dots \end{bmatrix}.
\end{aligned}
$$

Therefore:

$$
\texttt{Skew}((D,\Delta)\cdot K) = (\sqrt{2}b - e\cosh_\lambda(z+1))^2 + (\sqrt{2}\beta - \varepsilon\cosh_\lambda(\zeta-1))^2. \tag{36}
$$

But recall that $e^2 = b^2+1$, and from Remark A.2 that $b \geqslant 0$ implies $-be \leqslant -b^2$, so:

$$
\begin{aligned}
&(\sqrt{2}b - e\cosh_\lambda(z+1))^2 \\
&\quad= 2b^2 - 2\sqrt{2}\,be\cosh_\lambda(z+1) + e^2\cosh^2_\lambda(z+1) \\
&\quad\leqslant 2b^2 - 2\sqrt{2}\,b^2\cosh_\lambda(z+1) + (b^2+1)\cosh^2_\lambda(z+1) \\
&\quad= b^2(2-2\sqrt{2}\cosh_\lambda(z+1)+\cosh^2_\lambda(z+1)) + \cosh^2_\lambda(z+1) \\
&\quad= b^2(\sqrt{2}-\cosh_\lambda(z+1))^2 + \cosh^2_\lambda(z+1).
\end{aligned} \tag{37}
$$

Reasoning analogously, we also have

$$
(\sqrt{2}\beta - \varepsilon\cosh_\lambda(\zeta-1))^2 \quad\leqslant\quad \beta^2(\sqrt{2}-\cosh_\lambda(\zeta-1))^2 + \cosh^2_\lambda(\zeta-1). \tag{38}
$$

By assumption, $\texttt{Bias}(D,\Delta) \in [-1,1]$, thus $\zeta \leqslant z+1$. This, together with the assumptions $0.8 \leqslant \zeta$ and $z \leqslant 0.3$, implies that both $z+1$ and $\zeta-1$ are in the interval $[-0.2, 1.3]$. On this interval, the function $\cosh^2_\lambda(x)$ assumes its maximum at $x = 1.3$, and the function $f(x) = (\sqrt{2}-\cosh_\lambda(x))^2$ assumes its maximum at $x = 0$. Therefore,

$$
b^2(\sqrt{2}-\cosh_\lambda(z+1))^2 + \cosh^2_\lambda(z+1) \quad\leqslant\quad b^2(\sqrt{2}-\cosh_\lambda(0))^2 + \cosh^2_\lambda(1.3). \tag{39}
$$

and

$$
\beta^2(\sqrt{2}-\cosh_\lambda(\zeta-1))^2 + \cosh^2_\lambda(\zeta-1) \quad\leqslant\quad \beta^2(\sqrt{2}-\cosh_\lambda(0))^2 + \cosh^2_\lambda(1.3). \tag{40}
$$

Combining (36)–(40), together with the assumption that $\texttt{Skew}(D,\Delta) \geqslant 15$, yields:

$$
\begin{aligned}
\texttt{Skew}((D,\Delta)\cdot K) &= (\sqrt{2}b - e\cosh_\lambda(z+1))^2 + (\sqrt{2}\beta - \varepsilon\cosh_\lambda(\zeta-1))^2 \\
&\leqslant (b^2+\beta^2)(\sqrt{2}-\cosh_\lambda(0))^2 + 2\cosh^2_\lambda(1.3) \\
&= \texttt{Skew}(D,\Delta)(\sqrt{2}-\cosh_\lambda(0))^2 + 2\cosh^2_\lambda(1.3) \\
&\leqslant \texttt{Skew}(D,\Delta)((\sqrt{2}-\cosh_\lambda(0))^2 + \frac{2}{15}\cosh^2_\lambda(1.3))
\end{aligned}
$$

This completes the proof since $(\sqrt{2}-\cosh_\lambda(0))^2 + \frac{2}{15}\cosh^2_\lambda(1.3) \approx 0.571 \leqslant 0.9$. ☐

A.4 The A Lemma

Definition A.16. Let $g(x) = (1-2x)^2$.

Lemma A.17. *Recall the operator A from Figure 6. If (D,Δ) is such that $\texttt{Bias}(D,\Delta) \in [-1,1]$, $\texttt{Skew}(D,\Delta) \geqslant 15$, and such that $b, \beta \geqslant 0$ and $0.3 \leqslant z, \zeta$, then there exists $n \in \mathbb{Z}$ such that:*

$$
\texttt{Skew}((D,\Delta)\cdot A^n) \leqslant 0.9\ \texttt{Skew}(D,\Delta).
$$

Proof. Let $c = \min\{z, \zeta\}$ and $n = \max\{1, \lfloor\frac{\lambda^c}{2}\rfloor\}$. Compute the action of A^n on (D,Δ):

$$
\begin{aligned}
A^{n\dagger}DA^n &= \begin{bmatrix} 1 & 0 \\ -2n & 1 \end{bmatrix}\begin{bmatrix} e\lambda^{-z} & b \\ b & e\lambda^{z} \end{bmatrix}\begin{bmatrix} 1 & -2n \\ 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} \dots & b-2ne\lambda^{-z} \\ b-2ne\lambda^{-z} & \dots \end{bmatrix},
\end{aligned}
$$

$$
\begin{aligned}
A^{n\bullet\dagger}\Delta A^{n\bullet} &= A^{n\dagger}\Delta A^n \\
&= \begin{bmatrix} \dots & \beta-2n\varepsilon\lambda^{-\zeta} \\ \beta-2n\varepsilon\lambda^{-\zeta} & \dots \end{bmatrix}.
\end{aligned}
$$

Therefore:

$$\texttt{Skew}((D,\Delta)\cdot A^n) = (b - 2ne\lambda^{-z})^2 + (\beta - 2n\varepsilon\lambda^{-\zeta})^2$$

But recall that $e^2 = b^2 + 1$ and $\varepsilon^2 = \beta^2 + 1$, and from Remark A.2 that $b, \beta \geqslant 0$ implies $-be \leqslant -b^2$ and $-\varepsilon\beta \leqslant -\beta^2$. Using these facts, we can expand the above formula as follows:

$$\begin{aligned}
&\texttt{Skew}((D,\Delta)\cdot A^n)\\
&= (b - 2ne\lambda^{-z})^2 + (\beta - 2n\varepsilon\lambda^{-\zeta})^2\\
&= b^2 - 4nbe\lambda^{-z} + 4n^2e^2\lambda^{-2z} + \beta^2 - 4n\beta\varepsilon\lambda^{-\zeta} + 4n^2\varepsilon^2\lambda^{-2\zeta}\\
&\leqslant b^2 - 4nb^2\lambda^{-z} + 4n^2(b^2+1)\lambda^{-2z} + \beta^2 - 4n\beta^2\lambda^{-\zeta} + 4n^2(\beta^2+1)\lambda^{-2\zeta}\\
&= b^2(1 - 4n\lambda^{-z} + 4n^2\lambda^{-2z}) + \beta^2(1 - 4n\lambda^{-\zeta} + 4n^2\lambda^{-2\zeta}) + 4n^2(\lambda^{-2z} + \lambda^{-2\zeta})\\
&= b^2(1 - 2n\lambda^{-z})^2 + \beta^2(1 - 2n\lambda^{-\zeta})^2 + 4n^2(\lambda^{-2z} + \lambda^{-2\zeta})\\
&= b^2 g(n\lambda^{-z}) + \beta^2 g(n\lambda^{-\zeta}) + 4n^2(\lambda^{-2z} + \lambda^{-2\zeta}).
\end{aligned}$$

Writing $y = \max\{g(n\lambda^{-z}), g(n\lambda^{-\zeta})\}$ for brevity, and using the assumption that $\texttt{Skew}(D,\Delta) \geqslant 15$ together with the fact that $c \leqslant z, \zeta$, we get:

$$\begin{aligned}
\texttt{Skew}((D,\Delta)\cdot A^n) &\leqslant b^2y + \beta^2y + 8n^2\lambda^{-2c}\\
&= \texttt{Skew}(D,\Delta)y + 8n^2\lambda^{-2c}\\
&\leqslant \texttt{Skew}(D,\Delta)(y + \frac{8}{15}n^2\lambda^{-2c}).
\end{aligned}$$

To finish the proof, it remains to show that $y + \frac{8}{15}n^2\lambda^{-2c} \leqslant 0.9$. There are two cases:

- If $\lfloor\frac{\lambda^c}{2}\rfloor \geqslant 1$, then $\frac{\lambda^c}{4} \leqslant n \leqslant \frac{\lambda^c}{2}$. From $n \leqslant \frac{\lambda^c}{2}$, we have $2n\lambda^{-c} \leqslant 1$, and so $\frac{8}{15}n^2\lambda^{-2c} \leqslant \frac{2}{15}$. Moreover, because $\texttt{Bias}(D,\Delta) \in [-1,1]$, we have $c \leqslant z, \zeta \leqslant c+1$. Hence $\frac{1}{4\lambda} = \frac{\lambda^c}{4}\lambda^{-c-1} \leqslant n\lambda^{-c-1} \leqslant n\lambda^{-z}, n\lambda^{-\zeta} \leqslant n\lambda^{-c} \leqslant \frac{1}{2}$. On the interval $[\frac{1}{4\lambda}, \frac{1}{2}]$, the function $g(x)$ assumes its maximum at $x = \frac{1}{4\lambda}$. This implies that $y \leqslant g(\frac{1}{4\lambda})$. This completes the present case since we get:

$$y + \frac{8}{15}n^2\lambda^{-2c} \leqslant g(\frac{1}{4\lambda}) + \frac{2}{15} \approx 0.762 \leqslant 0.9.$$

- If $\lfloor\frac{\lambda^c}{2}\rfloor < 1$, then $n = 1$ and $\lambda^c < 2$. From $0.3 \leqslant c$, we have $\frac{8}{15}n^2\lambda^{-2c} \leqslant \frac{8}{15}\lambda^{-0.6}$. Moreover, because $\texttt{Bias}(D,\Delta) \in [-1,1]$, we have $0.3 \leqslant c \leqslant z, \zeta \leqslant c+1$. With $\lambda^c \leqslant 2$, this implies that $\frac{1}{2\lambda} \leqslant \lambda^{-c-1} \leqslant \lambda^{-z}, \lambda^{-\zeta} \leqslant \lambda^{-0.3}$. Therefore both λ^{-z} and $\lambda^{-\zeta}$ are in the interval $[\frac{1}{2\lambda}, \lambda^{-0.3}]$. On this interval, the function $g(x)$ assumes its maximum at $x = \frac{1}{2\lambda}$, and therefore $y \leqslant g(\frac{1}{2\lambda})$. This completes the proof since:

$$y + \frac{8}{15}n^2\lambda^{-2c} \leqslant g(\frac{1}{2\lambda}) + \frac{8}{15}\lambda^{-0.6} \approx 0.657 \leqslant 0.9.$$

□

A.5 The B Lemma

Definition A.18. Let $h(x) = (1 - \sqrt{2}x)^2$.

Lemma A.19. *Recall the operator B from Figure 6. If (D,Δ) is such that $\texttt{Bias}(D,\Delta) \in [-1,1]$, $\texttt{Skew}(D,\Delta) \geqslant 15$, and such that $b \leqslant 0 \leqslant \beta$ and $-0.2 \leqslant z, \zeta$, then there exists $n \in \mathbb{Z}$ such that:*

$$\texttt{Skew}((D,\Delta)\cdot B^n) \leqslant 0.9\ \texttt{Skew}(D,\Delta).$$

Proof. Let $c = \min\{z, \zeta\}$, $n = \max\{1, \lfloor\frac{\lambda^c}{\sqrt{2}}\rfloor\}$ and compute the action of B^n on (D,Δ):

$$\begin{aligned}
B^{n\dagger}DB^n &= \begin{bmatrix} 1 & 0 \\ \sqrt{2}n & 1 \end{bmatrix}\begin{bmatrix} e\lambda^{-z} & b \\ b & e\lambda^{z} \end{bmatrix}\begin{bmatrix} 1 & \sqrt{2}n \\ 0 & 1 \end{bmatrix}\\
&= \begin{bmatrix} \dots & b + \sqrt{2}\,ne\lambda^{-z} \\ b + \sqrt{2}\,ne\lambda^{-z} & \dots \end{bmatrix},
\end{aligned}$$

$$\begin{aligned}
B^{n\bullet\dagger}DB^{n\bullet} &= \begin{bmatrix} 1 & 0 \\ -\sqrt{2}n & 1 \end{bmatrix}\begin{bmatrix} \varepsilon\lambda^{-\zeta} & \beta \\ \beta & \varepsilon\lambda^{\zeta} \end{bmatrix}\begin{bmatrix} 1 & -\sqrt{2}n \\ 0 & 1 \end{bmatrix}\\
&= \begin{bmatrix} \dots & \beta - \sqrt{2}\,n\varepsilon\lambda^{-\zeta} \\ \beta - \sqrt{2}\,n\varepsilon\lambda^{-\zeta} & \dots \end{bmatrix}.
\end{aligned}$$

Therefore:

$$\texttt{Skew}((D,\Delta)\cdot B^n) = (b+\sqrt{2}\,ne\lambda^{-z})^2 + (\beta - \sqrt{2}\,n\varepsilon\lambda^{-\zeta})^2.$$

But recall that $e^2 = b^2+1$, that $\varepsilon^2 = \beta^2+1$, and from Remark A.2 that $b \leqslant 0 \leqslant \beta$ implies $be \leqslant -b^2$ and $-\beta\varepsilon \leqslant -\beta^2$. Using these facts, we can expand the above formula as follows:

$$\begin{array}{ll} \multicolumn{2}{l}{\texttt{Skew}((D,\Delta)\cdot B^n)} \\ = & (b+\sqrt{2}\,ne\lambda^{-z})^2 + (\beta - \sqrt{2}\,n\varepsilon\lambda^{-\zeta})^2 \\ = & b^2 + 2\sqrt{2}\,nbe\lambda^{-z} + 2n^2e^2\lambda^{-2z} + \beta^2 - 2\sqrt{2}\,n\beta\varepsilon\lambda^{-\zeta} + 2n^2\varepsilon^2\lambda^{-2\zeta} \\ \leqslant & b^2 - 2\sqrt{2}\,nb^2\lambda^{-z} + 2n^2(b^2+1)\lambda^{-2z} + \beta^2 - 2\sqrt{2}\,n\beta^2\lambda^{-\zeta} + 2n^2(\beta^2+1)\lambda^{-2\zeta} \\ = & b^2(1-2\sqrt{2}\,n\lambda^{-z} + 2n^2\lambda^{-2z}) + \beta^2(1-2\sqrt{2}\,n\lambda^{-\zeta} + 2n^2\lambda^{-2\zeta}) + 2n^2(\lambda^{-2z}+\lambda^{-2\zeta}) \\ = & b^2(1-\sqrt{2}\,n\lambda^{-z})^2 + \beta^2(1-\sqrt{2}\,n\lambda^{-\zeta})^2 + 2n^2(\lambda^{-2z}+\lambda^{-2\zeta}). \\ = & b^2h(n\lambda^{-z}) + \beta^2h(n\lambda^{-\zeta}) + 2n^2(\lambda^{-2z}+\lambda^{-2\zeta}). \end{array}$$

Writing $y = \max\{h(n\lambda^{-z}), h(n\lambda^{-\zeta})\}$ for brevity, and using the assumption that $\texttt{Skew}(D,\Delta) \geqslant 15$, together with the fact that $c \leqslant z,\zeta$, we get:

$$\begin{array}{rcl} \texttt{Skew}((D,\Delta)\cdot B^n) & \leqslant & b^2y + \beta^2y + 4n^2\lambda^{-2c} \\ & = & \texttt{Skew}(D,\Delta)y + 4n^2\lambda^{-2c} \\ & \leqslant & \texttt{Skew}(D,\Delta)(y + \dfrac{4}{15}n^2\lambda^{-2c}). \end{array}$$

To finish the proof, it remains to show that $y + \frac{4}{15}n^2\lambda^{-2c} \leqslant 0.9$. There are two cases:

- If $\lfloor \frac{\lambda^c}{\sqrt{2}} \rfloor \geqslant 1$, then $\frac{\lambda^c}{2\sqrt{2}} \leqslant n \leqslant \frac{\lambda^c}{\sqrt{2}}$. From $n \leqslant \frac{\lambda^c}{\sqrt{2}}$, we have $2n^2\lambda^{-2c} \leqslant 1$, and so $\frac{4n^2\lambda^{-2c}}{15} \leqslant \frac{2}{15}$. Moreover, because $\texttt{Bias}(D,\Delta) \in [-1,1]$, we have $c \leqslant z,\zeta \leqslant c+1$. Hence $\frac{1}{2\sqrt{2}\,\lambda} = \frac{\lambda^c}{2\sqrt{2}}\lambda^{-c-1} \leqslant n\lambda^{-c-1} \leqslant n\lambda^{-z}, n\lambda^{-\zeta} \leqslant n\lambda^{-c} \leqslant \frac{1}{\sqrt{2}}$. On the interval $[\frac{1}{2\sqrt{2}\,\lambda}, \frac{1}{\sqrt{2}}]$, the function $h(x)$ assumes its maximum at $x = \frac{1}{2\sqrt{2}\,\lambda}$. This implies that $y \leqslant h(\frac{1}{2\sqrt{2}\,\lambda})$. This completes the present case since we get:

  $$y + \frac{4}{15}n^2\lambda^{-2c} \leqslant h(\frac{1}{2\sqrt{2}\,\lambda}) + \frac{2}{15} \approx 0.762 \leqslant 0.9.$$

- If $\lfloor \frac{\lambda^c}{\sqrt{2}} \rfloor < 1$, then $n = 1$ and $\lambda^c < \sqrt{2}$. From $-0.2 \leqslant c$, we have $\frac{4}{15}n^2\lambda^{-2c} \leqslant \frac{4}{15}\lambda^{0.4}$. Moreover, because $\texttt{Bias}(D,\Delta) \in [-1,1]$, we have $-0.2 \leqslant c \leqslant z,\zeta \leqslant c+1$. With $\lambda^c \leqslant \sqrt{2}$, this implies that $\frac{1}{\sqrt{2}\,\lambda} \leqslant \lambda^{-c-1} \leqslant \lambda^{-z}, \lambda^{-\zeta} \leqslant \lambda^{0.2}$. Therefore both λ^{-z} and $\lambda^{-\zeta}$ are in the interval $[\frac{1}{\sqrt{2}\,\lambda}, \lambda^{0.2}]$. On this interval, the function $h(x)$ assumes its maximum at $x = \lambda^{0.2}$, and therefore $y \leqslant h(\lambda^{0.2})$. This completes the proof since:

  $$y + \frac{4}{15}n^2\lambda^{-2c} \leqslant h(\lambda^{0.2}) + \frac{4}{15}\lambda^{0.4} \approx 0.851 \leqslant 0.9.$$

□

A.6 Proof of the Step Lemma

The proof of the Step Lemma is now basically a case distinction, using the cases enumerated in Sections A.1–A.5, as well as some additional symmetric cases. In particular, the following remark will allow us to use the grid operators X and Z to reduce the number of cases to consider.

Remark A.20. The grid operator Z negates the anti-diagonal entries of a state while the operator X swaps the diagonal entries of a state. This follows by simple computation:

$$(D,\Delta)\cdot Z = \left(\begin{bmatrix} e\lambda^{-z} & -b \\ -b & e\lambda^{z} \end{bmatrix}, \begin{bmatrix} \varepsilon\lambda^{-\zeta} & -\beta \\ -\beta & \varepsilon\lambda^{\zeta} \end{bmatrix} \right), \qquad (D,\Delta)\cdot X = \left(\begin{bmatrix} e\lambda^{z} & b \\ b & e\lambda^{-z} \end{bmatrix}, \begin{bmatrix} \varepsilon\lambda^{\zeta} & \beta \\ \beta & \varepsilon\lambda^{-\zeta} \end{bmatrix} \right).$$

Moreover, $\texttt{Bias}((D,\Delta)\cdot Z) = \texttt{Bias}(D,\Delta)$ and $\texttt{Bias}((D,\Delta)\cdot X) = -\texttt{Bias}(D,\Delta)$.

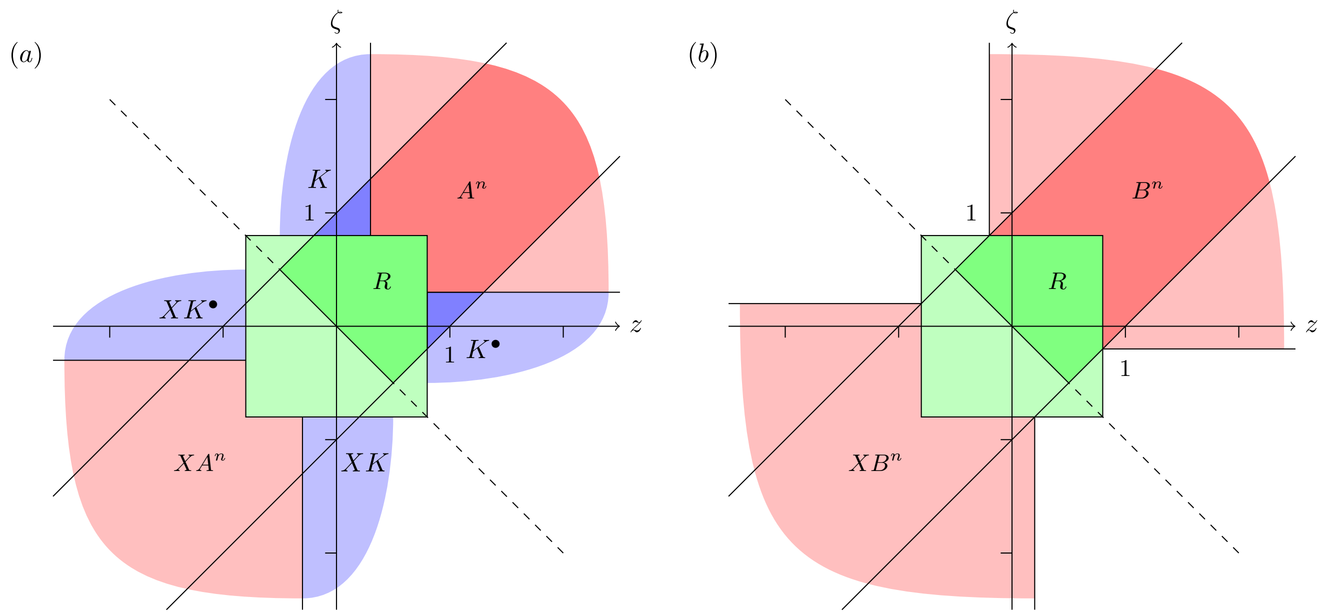


Figure 7: (a) A covering of the region $z - \zeta \in [-1, 1]$ and $z + \zeta \geqslant 0$ for the case $b \geqslant 0$. (b) A covering of the region $z - \zeta \in [-1, 1]$ and $z + \zeta \geqslant 0$ for the case $b \leqslant 0$.

Lemma (Step Lemma)**.** *For any state* (D, Δ)*, if* $\texttt{Skew}(D, \Delta) \geqslant 15$*, then there exists a special grid operator* G *such that* $\texttt{Skew}((D, \Delta) \cdot G) \leqslant 0.9\ \texttt{Skew}(D, \Delta)$*. Moreover,* G *can be computed using a constant number of arithmetic operations.*

Proof. Let (D, Δ) be a state such that $\texttt{Skew}(D, \Delta) \geqslant 15$. By Lemma A.11 we can assume w.l.o.g. that $\texttt{Bias}(D, \Delta) \in [-1, 1]$. Moreover, by Remark A.20, we can also assume that $\beta \geqslant 0$ and $z + \zeta \geqslant 0$. Note that the application of the grid operators X and/or Z in Remark A.20 preserves the fact that $\texttt{Bias}(D, \Delta) \in [-1, 1]$. We now treat in turn the cases $b \geqslant 0$ and $b \leqslant 0$.

Case 1 $b \geqslant 0$. A covering of the strip defined by $z - \zeta \in [-1, 1]$ and $z + \zeta \geqslant 0$ is depicted in Figure 7(a). The R region (in green) and the A region (in red) are defined as the intersection of this space with $\{(z, \zeta) \mid -0.8 \leqslant z, \zeta \leqslant 0.8\}$ and $\{(z, \zeta) \mid z \leqslant 0.3 \text{ and } 0.8 \leqslant \zeta\}$ respectively. The K and $K^\bullet$ regions (both in blue) fill the remaining space. We now consider in turn the possible locations of the pair (z, ζ) in this covering.

1. If $-0.8 \leqslant z, \zeta \leqslant 0.8$, then $\texttt{Skew}((D, \Delta) \cdot R) \leqslant 0.9\ \texttt{Skew}(D, \Delta)$ by Lemma A.13.
2. If $z \leqslant 0.3$ and $0.8 \leqslant \zeta$, then $\texttt{Skew}((D, \Delta) \cdot K) \leqslant 0.9\ \texttt{Skew}(D, \Delta)$ by Lemma A.15.
3. If $0.3 \leqslant z, \zeta$, then there exists $n \in \mathbb{Z}$ such that $\texttt{Skew}((D, \Delta) \cdot A^n) \leqslant 0.9\ \texttt{Skew}(D, \Delta)$ by Lemma A.17.
4. If $0.8 \leqslant z$ and $\zeta \leqslant 0.3$, then note that $(D, \Delta) \cdot K^\bullet = (\Delta, D) \cdot K$, and therefore by Lemma A.15:

$$\texttt{Skew}((D, \Delta) \cdot K^\bullet) = \texttt{Skew}((\Delta, D) \cdot K) \leqslant 0.9\ \texttt{Skew}(\Delta, D) = 0.9\ \texttt{Skew}(D, \Delta).$$

Case 2 $b \leqslant 0$. As above, we use a covering of the strip defined by $z - \zeta \in [-1, 1]$ and $z + \zeta \geqslant 0$ and consider the possible locations of (z, ζ) in this space. The relevant covering is depicted in Figure 7(b), where the R region (in green) is defined as above and the B region (in red) is defined as the intersection of the strip with $\{(z, \zeta) \mid z, \zeta \geqslant -0.2\}$.

1. If $-0.8 \leqslant z, \zeta \leqslant 0.8$, then $\texttt{Skew}((D, \Delta) \cdot R) \leqslant 0.9\ \texttt{Skew}(D, \Delta)$ by Lemma A.13.
2. If $z, \zeta \geqslant -0.2$ then there exists $n \in \mathbb{Z}$ such that $\texttt{Skew}((D, \Delta) \cdot B^n) \leqslant 0.9\ \texttt{Skew}(D, \Delta)$ by Lemma A.19.

Finally, note that only a constant number of calculations are required to decide which of the above cases applies. Moreover, each case only requires a constant number of operations. Specifically, the computation of k and σ^k in Lemma A.11, of n and A^n in Lemma A.17, and of n and B^n in Lemma A.19 each require just a fixed number of operations, and each of the remaining cases produces a fixed grid operator. $\square$

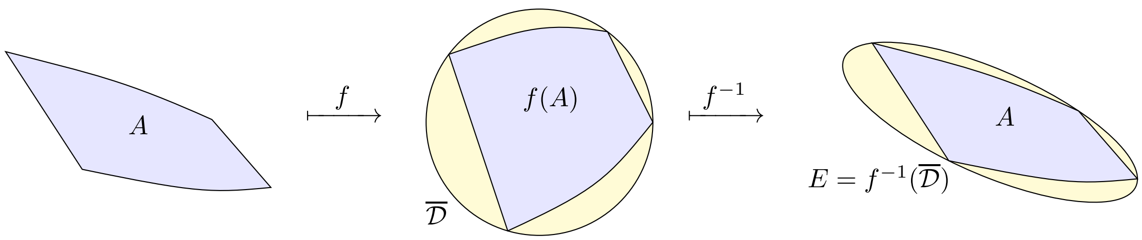
B Proof of Proposition 5.17

We prove Proposition 5.17, whose statement we reproduce here.

Proposition. *Let A be a bounded convex subset of $\mathbb{R}^2$ with non-empty interior. Then there exists an ellipse E such that $A \subseteq E$, and such that*

$$\text{area}(E) \leqslant \frac{4\pi}{3\sqrt{3}}\,\text{area}(A). \tag{41}$$

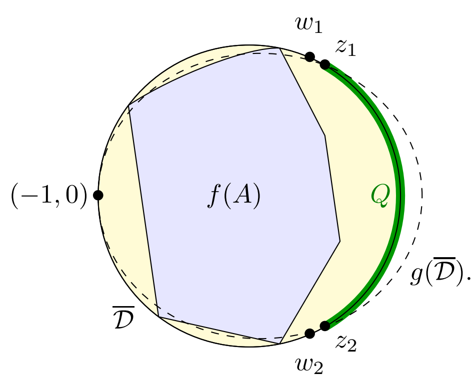
Proof. We may assume without loss of generality that A is compact, for it it is not, we can replace A by its closure, which has the same area as A because A is convex. Let $\overline{\mathcal{D}}$ be the closed unit disk, and consider the collection $\text{Aff}(A, \overline{\mathcal{D}})$ of all affine transformations $f : \mathbb{R}^2 \to \mathbb{R}^2$ satisfying $f(A) \subseteq \overline{\mathcal{D}}$. Then $\text{Aff}(A, \overline{\mathcal{D}})$, with the natural topology, is a compact set. Therefore, there exists some $f \in \text{Aff}(A, \overline{\mathcal{D}})$ maximizing the area of $f(A)$. We claim that f is invertible. Indeed, since A is bounded, there exists some $\lambda > 0$ such that $\lambda A \subseteq \overline{\mathcal{D}}$. Since λA has non-zero area, and multiplication by λ is an affine map, it follows that $f(A)$ has non-zero area as well, and so f is invertible. Let $E = f^{-1}(\overline{\mathcal{D}})$. We claim that E is the desired ellipse.



We have $A \subseteq E$ by construction. We must show (41). Because affine transformations preserve ratios of areas, we may equivalently show

$$\text{area}(\overline{\mathcal{D}}) \leqslant \frac{4\pi}{3\sqrt{3}}\,\text{area}(f(A)).$$

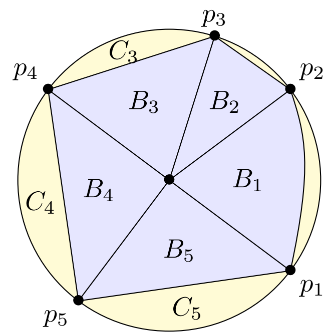
Let $\partial\overline{\mathcal{D}}$ be the boundary of $\overline{\mathcal{D}}$, and consider points p where $f(A)$ "touches" the boundary, i.e., points $p \in f(A) \cap \partial\overline{\mathcal{D}}$. We claim that any arc segment of $\partial\overline{\mathcal{D}}$ of length $2\pi/3$ radians (120 degrees) contains at least one such point. To prove this, assume, for the sake of contradiction, that there is such an arc segment Q not containing any point of $f(A)$. By rotational symmetry, we may without loss of generality assume that Q is the arc from $-\pi/3$ to $\pi/3$ radians on the unit circle. Let z_1 and z_2 be the endpoints of Q, as shown here:



Since both Q and $f(A)$ are compact, there exists some $d > 0$ such that the distance between any point of A and any point of Q is at least d. Let w_1 and w_2 be the two points on the unit circle whose distance from Q is $d/2$. Now consider the affine transformation g that fixes $(-1, 0)$ and maps w_1 to z_1 and w_2 to z_2. Then $g(\overline{\mathcal{D}})$ is an ellipse whose boundary passes through the points $(-1, 0)$, z_1, and z_2, shown as a dashed line in the above illustration. It is therefore bisected by Q. Since the map g moves no point of the unit disk by more than distance d, the set $g(f(A))$ does not intersect Q. It follows that $g(f(A))$ is contained in $\overline{\mathcal{D}}$. On the other hand, a calculation shows that the area of $g(f(A))$ is slightly greater than that of $f(A)$, contradicting the assumption that the area of $f(A)$ was maximal.

We have proved that every arc segment of length $2\pi/3$ radians on the boundary of $\overline{\mathcal{D}}$ contains a point of $f(A)$. It follows that there is some finite cyclic sequence of points $p_1, \ldots, p_n \in f(A) \cap \partial\overline{\mathcal{D}}$ such that consecutive points are no more than $2\pi/3$ radians apart. By connecting each p_i to the center, we partition each of the sets $f(A)$ and $\overline{\mathcal{D}}$ into n

pieces $B_1, \ldots, B_n$ and $C_1, \ldots, C_n$, respectively.



The fact that the inner angles are less than $2\pi/3$ immediately implies that $\text{area}(C_i) \leqslant \frac{4\pi}{3\sqrt{3}}\,\text{area}(B_i)$ for all i; hence also $\text{area}(\overline{\mathcal{D}}) \leqslant \frac{4\pi}{3\sqrt{3}}\,\text{area}(f(A))$. This finishes the proof of the proposition. $\square$

C Proof of Theorem 6.2

Consider the rings $\mathbb{Z}$ and $\mathbb{D}$, together with their respective extensions $\mathbb{Z}[\sqrt{2}]$, $\mathbb{Z}[\omega]$ and $\mathbb{D}[\sqrt{2}]$, $\mathbb{D}[\omega]$, as introduced in Section 3. We wish to give an efficient method for solving equations of the form $t^\dagger t = \xi$, for given $\xi \in \mathbb{D}[\sqrt{2}]$ and unknown $t \in \mathbb{D}[\omega]$. To do this, we first classify the primes in $\mathbb{Z}[\sqrt{2}]$ and $\mathbb{Z}[\omega]$, then show how to solve the equation $t^\dagger t = \xi$ in the case where $\xi \in \mathbb{Z}[\sqrt{2}]$ and $t \in \mathbb{Z}[\omega]$, which we finally extend to the general case. None of the results in this section are original; they are well-known from the theory of cyclotomic fields, which goes back to the 19th century work of Gauss and Kummer. Nevertheless, we hope that the following detailed treatment may be useful to readers who are not experts in algebraic number theory. Moreover, we hope to give sufficient details to enable an interested reader, in principle, to implement the algorithm. This will also aid in our complexity analysis.

A fundamental property of the rings $\mathbb{Z}$, $\mathbb{Z}[\sqrt{2}]$, and $\mathbb{Z}[\omega]$ is that they are *Euclidean domains*; this implies that the notions of divisibility, greatest common divisor, and unique prime factorization all make sense in these rings. Recall that in a ring, a *unit* is an invertible element. In a Euclidean domain, we write $x \mid y$ if x is a divisor of y, and $x \sim y$ if $x \mid y$ and $y \mid x$; equivalently, $x \sim y$ iff there exists a unit u that $xu = y$. An element x is *prime* if x is not a unit, and $x = ab$ implies that either a or b is a unit. Note that if x is prime and $x \mid ab$, then $x \mid a$ or $x \mid b$; this follows from Euclid's algorithm.

C.1 Units in $\mathbb{Z}[\sqrt{2}]$

Definition C.1. We say that $\xi \in \mathbb{Z}[\sqrt{2}]$ is *doubly positive* if $\xi \geqslant 0$ and $\xi^\bullet \geqslant 0$.

Lemma C.2. *The units of $\mathbb{Z}[\sqrt{2}]$ are of the form $u = (-1)^n \lambda^m$, where $\lambda = 1 + \sqrt{2}$. Moreover, a unit u is doubly positive if and only if u is a square in $\mathbb{Z}[\sqrt{2}]$.*

Proof. Lemma 10 of [14]. $\square$

Lemma C.3. *Let $\xi \in \mathbb{Z}[\sqrt{2}]$, and consider $n = \xi^\bullet \xi$. If n is a unit in $\mathbb{Z}$, then ξ is a unit in $\mathbb{Z}[\sqrt{2}]$.*

Proof. If n is a unit, then there exists $m \in \mathbb{Z}$ such that $nm = 1$, hence $\xi\xi^\bullet m = 1$, hence ξ is invertible in $\mathbb{Z}[\sqrt{2}]$. $\square$

C.2 Primes in $\mathbb{Z}[\sqrt{2}]$

Lemma C.4. *Let $\xi \in \mathbb{Z}[\sqrt{2}]$, and consider $n = \xi^\bullet \xi$. If n is prime in $\mathbb{Z}$, then ξ is prime in $\mathbb{Z}[\sqrt{2}]$.*

Proof. Suppose $\xi = \alpha\beta$ in $\mathbb{Z}[\sqrt{2}]$, and consider $n = \xi^\bullet\xi = \alpha^\bullet\alpha\beta^\bullet\beta$. Since $\alpha^\bullet\alpha$ and $\beta^\bullet\beta$ are integers and n is prime, we must have that either $\alpha^\bullet\alpha$ or $\beta^\bullet\beta$ is a unit in $\mathbb{Z}$; hence α or β is a unit in $\mathbb{Z}[\sqrt{2}]$. So ξ is prime. $\square$

Lemma C.5. *For every prime ξ of $\mathbb{Z}[\sqrt{2}]$, there exists a unique (up to a unit) prime p of $\mathbb{Z}$ such that $\xi \mid p$.*

Proof. To show existence, consider $n = \xi^\bullet\xi$. Note that $\xi \neq 0$, hence $n \neq 0$. Let $n = p_1 p_2 \cdots p_k$ be a prime factorization of n. Since $\xi \mid p_1 p_2 \cdots p_k$ and ξ is prime in $\mathbb{Z}[\sqrt{2}]$, there exists some i such that $\xi \mid p_i$.

To show uniqueness, assume $\xi \mid p$ and $\xi \mid q$, where $p \not\sim q$. Then $\gcd(p, q) = 1$, hence by Euclid's algorithm, we can write $1 = np + mq$, for integers n and m. Then $\xi \mid np + mq = 1$, which is absurd since ξ is prime. $\square$

Lemma C.6. *Let ξ be a prime of $\mathbb{Z}[\sqrt{2}]$, and let p be the unique (up to a unit) prime of $\mathbb{Z}$ such that $\xi \mid p$. Then exactly one of the following holds: $\xi \sim p$ or $\xi^\bullet \xi \sim p$.*

Proof. First note that at most one of these properties can hold, because otherwise $\xi \sim \xi^\bullet \xi$, which implies that $\xi^\bullet$ is a unit, which is absurd since it is prime.

We now show that at least one of the properties holds. Since $\xi \mid p$ and p is an integer, we have $\xi^\bullet \mid p$, hence $\xi^\bullet \xi \mid p^2$. Also, $\xi^\bullet \xi$ is an integer, so either $\xi^\bullet \xi \sim 1$, $\xi^\bullet \xi \sim p$, or $\xi^\bullet \xi \sim p^2$. The first of these cases is absurd, since ξ would then be a unit. In the second case, we have $\xi^\bullet \xi \sim p$, which is to be shown. In the third case, since $\xi \mid p$, there is $\alpha \in \mathbb{Z}[\sqrt{2}]$ such that $\xi\alpha = p$. Then we have $\xi^\bullet \xi \alpha^\bullet \alpha = p^2 \sim \xi^\bullet \xi$, which implies that α is a unit, hence $\xi \sim p$. □

Lemma C.7. *Let p be a prime of $\mathbb{Z}$. Then the prime factorization of p in $\mathbb{Z}[\sqrt{2}]$ consists of one or two factors.*

Proof. Let ξ be some prime factor of p in $\mathbb{Z}[\sqrt{2}]$. By Lemma C.6, either $\xi \sim p$ or $\xi^\bullet \xi \sim p$. Either way, this gives a prime factorization of p in $\mathbb{Z}[\sqrt{2}]$. □

We now determine the prime factorization in $\mathbb{Z}[\sqrt{2}]$ of every prime p of $\mathbb{Z}$. It turns out that there are 5 cases, depending whether p is even, or $p \equiv 1, 3, 5, 7 \pmod 8$.

Lemma C.8. *The prime factorization of $p = 2$ in $\mathbb{Z}[\sqrt{2}]$ is $p = \sqrt{2} \cdot \sqrt{2}$. The prime factorization of $p = -2$ in $\mathbb{Z}[\sqrt{2}]$ is $p = -\sqrt{2} \cdot \sqrt{2}$.*

Proof. We only need to show that $\sqrt{2}$ is prime in $\mathbb{Z}[\sqrt{2}]$, but this follows from Lemma C.4. □

Lemma C.9. *Let p be a prime of $\mathbb{Z}$ such that $p \equiv 3 \pmod 8$ or $p \equiv 5 \pmod 8$. Then p is prime in $\mathbb{Z}[\sqrt{2}]$.*

Proof. Let ξ be a prime factor of p in $\mathbb{Z}[\sqrt{2}]$. By Lemma C.6, we know that $\xi \sim p$ or $\xi^\bullet \xi \sim p$. In the former case, p is prime and we are done. In the latter case, writing $\xi = a + b\sqrt{2}$, we have $\xi^\bullet \xi = a^2 - 2b^2 = \pm p$, hence $a^2 - 2b^2 \equiv \pm 3 \pmod 8$. By easy case distinction, we see that a^2 can only be 0, 1, or 4 modulo 8, and $2b^2$ can only be 0 or 2 modulo 8, so $a^2 - 2b^2 \equiv \pm 3 \pmod 8$ is plainly impossible. □

Lemma C.10. *Let $\xi \in \mathbb{Z}[\sqrt{2}]$ such that $\xi \sim \xi^\bullet$. Then either $\xi \sim n$, or $\xi \sim n\sqrt{2}$, for some $n \in \mathbb{Z}$.*

Proof. By assumption, $\xi \sim \xi^\bullet$, so there exists a unit $u \in \mathbb{Z}[\sqrt{2}]$ such that $\xi^\bullet = u\xi$. By Lemma C.2, the units of $\mathbb{Z}[\sqrt{2}]$ are exactly of the form $u = (-1)^n \lambda^m$, where $\lambda = 1 + \sqrt{2}$. Applying the automorphism to $\xi^\bullet = u\xi$, we get $\xi = u^\bullet \xi^\bullet$, hence $\xi^\bullet \xi = u^\bullet u \xi^\bullet \xi$, hence $u^\bullet u = 1$, hence $(\lambda^\bullet \lambda)^m = 1$. But $\lambda^\bullet \lambda = -1$, so m is even; say $m = 2k$. Let $\xi' = \lambda^k \xi$. Note that $\xi' \sim \xi$. Then we have

$$\xi'^\bullet = (\lambda^\bullet)^k \xi^\bullet = (-\lambda)^{-k} u\xi = (-1)^{-k} \lambda^{-k} (-1)^n \lambda^m \xi = \pm \lambda^k \xi = \pm \xi'. \tag{42}$$

We can write $\xi' = a + b\sqrt{2}$ for some $a, b \in \mathbb{Z}$. From (42), we have either $a - b\sqrt{2} = a + b\sqrt{2}$, or $a - b\sqrt{2} = -(a + b\sqrt{2})$. In the first case, $b = 0$, and therefore $\xi \sim \xi' = a$. In the second case, $a = 0$, and therefore $\xi \sim \xi' = b\sqrt{2}$. □

Lemma C.11. *Let p be a prime of $\mathbb{Z}$ such that $p \equiv 1 \pmod 8$ or $p \equiv 7 \pmod 8$. Then p has a prime factorization of the form $p \sim \xi^\bullet \xi$ in $\mathbb{Z}[\sqrt{2}]$; moreover, $\xi \not\sim \xi^\bullet$, so that the two prime factors are distinct.*

Proof. It is well-known (by quadratic reciprocity) that if p is a prime with $p \equiv \pm 1 \pmod 8$, then 2 is a quadratic residue modulo p. Therefore, the equation $x^2 \equiv 2 \pmod p$ has an integer solution; let x be such a solution. Let $\alpha = x + \sqrt{2}$. Then $\alpha^\bullet \alpha = x^2 - 2 \equiv 0 \pmod p$, and hence $p \mid \alpha^\bullet \alpha$. On the other hand, clearly $p \nmid \alpha$ (since $\alpha/p \notin \mathbb{Z}[\sqrt{2}]$), and similarly $p \nmid \alpha^\bullet$, which shows that p is not a prime in $\mathbb{Z}[\sqrt{2}]$. By Lemma C.6, $p \sim \xi^\bullet \xi$ for some prime ξ of $\mathbb{Z}[\sqrt{2}]$.

The final claim follows from Lemma C.10, because if $\xi \sim \xi^\bullet$, then either $\xi \sim n$ or $\xi \sim n\sqrt{2}$. In the first case, $n^2 \mid \xi^\bullet \xi \sim p$, but p is prime, so that $n = \pm 1$, contradicting that ξ is prime. In the second case, $2n^2 \mid \xi^\bullet \xi \sim p$, contradicting the assumption that p is odd. □

Lemma C.12. *Let p be a prime of $\mathbb{Z}$. A prime factorization of p in $\mathbb{Z}[\sqrt{2}]$ can be computed in probabilistic polynomial time.*

Proof. Assume without loss of generality that $p > 0$. If $p = 2$, then $p = \sqrt{2}^2$ is the desired prime factorization by Lemma C.8. If $p \equiv 3, 5 \pmod 8$, then p is already prime in $\mathbb{Z}[\sqrt{2}]$ by Lemma C.9. The remaining case is when $p \equiv 1, 7 \pmod 8$. In this case the prime factorization is of the form $p \sim \xi^\bullet \xi$ by Lemma C.11. Moreover, the proof of Lemma C.11 indicates how such ξ can be computed, namely as $\xi = \gcd(p, x + \sqrt{2})$, where x is a solution of $x^2 \equiv 2 \pmod p$. The equation $x^2 \equiv 2 \pmod p$ can be solved in probabilistic polynomial time by a well-known algorithm, see [12]. □

C.3 Primes in $\mathbb{Z}[\omega]$

Lemma C.13. *Let ξ be a prime in $\mathbb{Z}[\sqrt{2}]$. Then either ξ is prime in $\mathbb{Z}[\omega]$, or else $\xi \sim t^\dagger t$ where t is some prime of $\mathbb{Z}[\omega]$. In particular, the prime factorization of ξ in $\mathbb{Z}[\omega]$ consists of either one or two factors.*

Proof. This is similar to the proof of Lemma C.7. Let t be some prime factor of ξ in $\mathbb{Z}[\omega]$. Then $t \,|\, \xi$, therefore $t^\dagger \,|\, \xi^\dagger = \xi$, therefore $t^\dagger t \,|\, \xi^2$ in $\mathbb{Z}[\omega]$. But both $t^\dagger t$ and ξ are elements of $\mathbb{Z}[\sqrt{2}]$, so $t^\dagger t \,|\, \xi^2$ in $\mathbb{Z}[\sqrt{2}]$. Since ξ is prime in $\mathbb{Z}[\sqrt{2}]$, there are only three possibilities: $t^\dagger t \sim 1$, $t^\dagger t \sim \xi$, and $t^\dagger t \sim \xi^2$. In the first case, t is a unit, contradicting the fact that it is prime in $\mathbb{Z}[\omega]$. In the second case, we are done. In the third case, since $t \,|\, \xi$, there is $a \in \mathbb{Z}[\omega]$ such that $at = \xi$. Then $a^\dagger a t^\dagger t = \xi^\dagger \xi = \xi^2 = t^\dagger t$, which implies $a^\dagger a = 1$. Therefore a is a unit, and $t \sim \xi$. Therefore ξ is prime in $\mathbb{Z}[\omega]$ and we are done. □

C.4 The Diophantine equation $t^\dagger t = \xi$

We are interested in solving equations of the form

$$t^\dagger t = \xi, \tag{43}$$

where $\xi \in \mathbb{D}[\sqrt{2}]$ is given, and $t \in \mathbb{D}[\omega]$ is unknown.

Definition C.14. Recall that for elements $\xi, \xi' \in \mathbb{Z}[\sqrt{2}]$, the notation $\xi \sim \xi'$ means that ξ, ξ' differ by a unit, i.e., there exists a unit $u \in \mathbb{Z}[\sqrt{2}]$ such that $\xi = u\xi'$. We extend this notation also to the ring $\mathbb{D}[\sqrt{2}]$: for $\xi, \xi' \in \mathbb{D}[\sqrt{2}]$, we write $\xi \sim \xi'$ iff there exists a unit $u \in \mathbb{Z}[\sqrt{2}]$ such that $\xi = u\xi'$. Note that we have taken u to be a unit of the ring $\mathbb{Z}[\sqrt{2}]$, not of $\mathbb{D}[\sqrt{2}]$.

It will often be convenient to replace (43) by the following weaker condition.

Definition C.15. We say that $\xi \in \mathbb{D}[\sqrt{2}]$ is $\dagger$-*decomposable* if the equation

$$t^\dagger t \sim \xi \tag{44}$$

has a solution $t \in \mathbb{D}[\omega]$.

Solutions to (43) can be recovered from solutions to (44) by using the following observation:

Lemma C.16. *Let $\xi \in \mathbb{D}[\sqrt{2}]$. Then equation* (43) *has a solution if and only if ξ is doubly positive and $\dagger$-decomposable.*

Proof. If ξ is a solution to (43), then it is obviously $\dagger$-decomposable. It is also doubly positive by Lemma 6.1. Conversely, assume $t^\dagger t \sim \xi$. Then there exists a unit u of $\mathbb{Z}[\sqrt{2}]$ such that $\xi = ut^\dagger t$. Since both ξ and $t^\dagger t$ are doubly positive, it follows that u is doubly positive, and hence u is a square of the ring $\mathbb{Z}[\sqrt{2}]$ by Lemma C.2; say $u = v^2$. Since $v \in \mathbb{Z}[\sqrt{2}]$, we have $v = v^\dagger$. Setting $t' = vt$, we have $\xi = v^\dagger v t^\dagger t = t'^\dagger t'$, which finishes the proof. □

C.5 The case $\xi \in \mathbb{Z}[\sqrt{2}]$

Lemma C.17. *Suppose $t^\dagger t \sim \xi$ for $t \in \mathbb{D}[\omega]$ and $\xi \in \mathbb{Z}[\sqrt{2}]$. Then $t \in \mathbb{Z}[\omega]$.*

Proof. Note that, in $\mathbb{Z}[\omega]$, we have $\sqrt{2}^k \,|\, t$ if and only if $2^k \,|\, t^\dagger t$. Choose $t' \in \mathbb{Z}[\omega]$ and $k \geqslant 0$ such that $t = t'/\sqrt{2}^k$. Then $t'^\dagger t' = 2^k \xi$, hence $2^k \,|\, t'^\dagger t'$, hence $\sqrt{2}^k \,|\, t'$, hence $t \in \mathbb{Z}[\omega]$. □

Lemma C.18. *If x, y, z are three elements of a Euclidean domain, then*

$$\gcd(xy, z) \,|\, \gcd(x, z) \cdot \gcd(y, z).$$

Proof. By considering the prime factorization of z. □

Lemma C.19. *Suppose $\xi = \alpha\beta$, where $\alpha, \beta \in \mathbb{Z}[\sqrt{2}]$ and $\gcd(\alpha, \beta) = 1$. Then ξ is $\dagger$-decomposable iff α and β are $\dagger$-decomposable.*

Proof. For the right-to-left implication, assume $\alpha \sim s^\dagger s$ and $\beta \sim r^\dagger r$. Clearly $\xi = \alpha\beta \sim (sr)^\dagger sr$. For the left-to-right implication, assume $\xi \sim t^\dagger t$. Note that $t \in \mathbb{Z}[\omega]$ by Lemma C.17. To show that α is $\dagger$-decomposable, let $s = \gcd(t, \alpha)$. We claim that $s^\dagger s \sim \alpha$. Clearly, since $s \,|\, \alpha$ and $s^\dagger \,|\, \alpha$, we have $s^\dagger s \,|\, \alpha^2$. One the other hand, we know that $s^\dagger s \,|\, t^\dagger t \sim \xi = \alpha\beta$. Since α and β are relatively prime, it follows that $s^\dagger s \,|\, \alpha$. Conversely, note that by Lemma C.18, $\alpha = \gcd(t^\dagger t, \alpha) \,|\, \gcd(t^\dagger, \alpha) \cdot \gcd(t, \alpha) = s^\dagger s$. Therefore $s^\dagger s \sim \alpha$ and α is $\dagger$-decomposable. The argument for β is similar. □

Lemma C.20. *Suppose that $\xi \in \mathbb{Z}[\sqrt{2}]$ is prime. Let $p > 0$ be the unique positive prime in $\mathbb{Z}$ such that $\xi \mid p$. Then ξ is $\dagger$-decomposable if and only if $p = 2$ or $p \equiv 1, 3, 5 \pmod 8$.*

Proof. We consider each case in turn. If $p = 2$, then $\xi \sim \sqrt{2}$. Let $\delta = 1 + \omega$; a simple calculation shows that $\delta^\dagger\delta = \lambda\sqrt{2} \sim \xi$. So ξ is $\dagger$-decomposable.

If $p \equiv 1 \pmod 4$, then by quadratic reciprocity, there exists some integer u such that $u^2 \equiv -1 \pmod p$. Therefore $\xi \mid p \mid u^2 + 1 = (u+i)(u-i)$. Let $t = \gcd(\xi, u+i)$. We claim that $\xi \sim t^\dagger t$. Note that $t \mid \xi$, hence $t^\dagger \mid \xi^\dagger = \xi$, hence $t^\dagger t \mid \xi^2$. Since ξ is prime in $\mathbb{Z}[\sqrt{2}]$, there are 3 possibilities: $t^\dagger t \sim 1$, $t^\dagger t \sim \xi$, or $t^\dagger t \sim \xi^2$. The first case is not possible, because in this case, t would be a unit, so that ξ is relatively prime to $u+i$, hence to $u-i$, hence to u^2+1, contradicting $\xi \mid u^2+1$. In the second case, we have $t^\dagger t \sim \xi$, which was to be shown. In the third case, we have $t^\dagger t \sim \xi^2$. Since $t \mid \xi$, there exists some $s \in \mathbb{Z}[\omega]$ such that $ts = \xi$. But then $t^\dagger t s^\dagger s = \xi^\dagger\xi = \xi^2 \sim t^\dagger t$, so that s is a unit. In this case, we have $t \sim \xi$, therefore $\xi \mid u+i$, therefore also $\xi = \xi^\dagger \mid u-i$, hence $\xi \mid (u+i) - (u-i) \sim 2$, contradicting the fact that p is the only prime integer divisible by ξ.

The case $p \equiv 3 \pmod 8$ is very similar. In this case, by quadratic reciprocity, -2 is a square modulo p, so that there exists some integer u such that $u^2 \equiv -2 \pmod p$. Therefore $\xi \mid p \mid u^2+2 = (u+i\sqrt{2})(u-i\sqrt{2})$. Let $t = \gcd(\xi, u+i\sqrt{2})$. We claim that $\xi \sim t^\dagger t$. Note that $t \mid \xi$, hence $t^\dagger \mid \xi^\dagger = \xi$, so $t^\dagger t \mid \xi^2$. So again we have the three possibilities $t^\dagger t \sim 1$, $t^\dagger t \sim \xi$, or $t^\dagger t \sim \xi^2$. The first case is not possible, because in this case, t would be a unit, so that ξ is relatively prime to $u+i\sqrt{2}$, hence to $u-i\sqrt{2}$, hence to u^2+2, contradicting $\xi \mid u^2+2$. In the second case, we have $t^\dagger t \sim \xi$, which was to be shown. In the third case, we have $t^\dagger t \sim \xi^2$. Since $t \mid \xi$, there exists some $s \in \mathbb{Z}[\omega]$ such that $ts = \xi$. But then $t^\dagger t s^\dagger s = \xi^\dagger \xi = \xi^2 \sim t^\dagger t$, so that s is a unit. In this case, we have $t \sim \xi$, therefore $\xi \mid u+i\sqrt{2}$, therefore also $\xi = \xi^\dagger \mid u-i\sqrt{2}$, hence $\xi \mid i(u-i\sqrt{2}) - i(u+i\sqrt{2}) = 2\sqrt{2}$. On the other hand, $\xi \mid p$, therefore $\xi \mid \gcd(p, 2\sqrt{2}) = 1$, contradicting the fact that ξ is not a unit.

Finally, if $p \equiv 7 \pmod 8$, then $p \sim \xi^\bullet\xi$ by Lemma C.11. Assume that ξ is $\dagger$-decomposable as $\xi \sim t^\dagger t$. Then

$$(t^\dagger t)^\bullet (t^\dagger t) \sim \xi^\bullet \xi \sim p.$$

Note that $t^\bullet t \in \mathbb{Z}[i]$, so we can write $t^\bullet t = a + bi$ for some $a, b \in \mathbb{Z}$. But then we have

$$p \sim (t^\bullet t)(t^\bullet t)^\dagger = (a+bi)(a-bi) = a^2 + b^2.$$

But a^2 and b^2 can only be congruent to 0, 1, or 4 modulo 8, contradicting $p \equiv 7 \pmod 8$. Therefore ξ is not $\dagger$-decomposable, which is what was claimed. □

Lemma C.21. *Suppose $\xi \in \mathbb{Z}[\sqrt{2}]$ is prime. Let $p > 0$ be the unique positive prime in $\mathbb{Z}$ such that $\xi \mid p$. Let m be a positive integer. Then ξ^m is $\dagger$-decomposable if and only if m is even or $p \equiv 1, 2, 3, 5 \pmod 8$.*

Proof. If m is even, then note that $\xi \in \mathbb{Z}[\sqrt{2}]$, so $\xi = \xi^\dagger$; therefore, $t = \xi^{m/2}$ will be a solution. If $p \equiv 1, 2, 3, 5 \pmod 8$, then by Lemma C.20, there exists $s \in \mathbb{Z}[\omega]$ with $s^\dagger s \sim \xi$; therefore $t = s^m$ is a solution of $t^\dagger t \sim \xi^m$. The only remaining case is then m is odd and $p \equiv 7 \pmod 8$. In this case, there can be no solution. For assume on the contrary that $t^\dagger t = \xi^m$. Then as in the proof of Lemma C.20, we can write $t^\bullet t = a + bi$ for some $a, b \in \mathbb{Z}$, and we get

$$p^m \sim (t^\bullet t)(t^\bullet t)^\dagger = (a+bi)(a-bi) = a^2 + b^2,$$

contradicting $p^m \equiv 7 \pmod 8$. □

Remark C.22. The proofs of Lemmas C.20 and C.21 are constructive, and immediately yield efficient algorithms for determining t, provided that we have an efficient method of solving $u^2 \equiv -1 \pmod p$ when p is a prime congruent to $1 \pmod 4$ and of solving $u^2 \equiv -2 \pmod p$ when p is a prime congruent to $3 \pmod 8$. These last problems can be solved in probabilistic polynomial time by a well-known algorithm, see [12].

Lemma C.23. *Given $\xi \in \mathbb{Z}[\sqrt{2}]$, together with its prime factorization in $\mathbb{Z}[\sqrt{2}]$, there exists an algorithm that determines, in probabilistic polynomial time, whether the equation $t^\dagger t \sim \xi$ has a solution or not, and finds a solution if there is one.*

Proof. Let $\xi \sim \xi_1^{m_1}\xi_2^{m_2}\cdots\xi_k^{m_k}$ be a prime factorization of ξ in $\mathbb{Z}[\sqrt{2}]$, where $\xi_1, \ldots, \xi_k$ are distinct (i.e., pairwise non-associate) primes. By Lemma C.19, $t^\dagger t \sim \xi$ has a solution if and only if $t^\dagger t \sim \xi_i^{m_i}$ has a solution for all i. By Lemma C.21, $t^\dagger t \sim \xi_i^{m_i}$ has a solution if and only if m_i is even or $p \equiv 1, 2, 3, 5 \pmod 8$. Since these conditions are easy to check, we can therefore determine the existence of a solution efficiently, i.e., in polynomial time.

Moreover, once a solution has been shown to exist, the actual solution can be efficiently computed. Namely, for each i, find t_i such that $t_i^\dagger t_i \sim \xi_i^{m_i}$ as in Lemma C.21. Then $t \sim t_1 t_2 \cdots t_k$ satisfies $t^\dagger t \sim \xi_1^{m_1}\xi_2^{m_2}\cdots\xi_k^{m_k} \sim \xi$. By Remark C.22, all of this can be computed in probabilistic polynomial time. □

Proposition C.24. *Let $\xi \in \mathbb{Z}[\sqrt{2}]$, and let $n = \xi^\bullet \xi$. Note that n is an integer. Given the prime factorization of n, there exists an algorithm that determines, in probabilistic polynomial time, whether the equation $t^\dagger t \sim \xi$ has a solution or not, and finds a solution if there is one.*

Proof. Suppose that $n = \pm p_1^{m_1} \cdots p_k^{m_k}$ is the prime factorization of n, where $p_1, \ldots, p_k$ are distinct positive primes. Each p_i can be efficiently factored into primes of $\mathbb{Z}[\sqrt{2}]$ by Lemma C.12; this yields the prime factorization of $n = \xi^\bullet \xi$ in $\mathbb{Z}[\sqrt{2}]$. From this, it is easy to obtain a prime factorization of ξ. The rest of the claim then follows from Lemma C.23. □

C.6 The case $\xi \in \mathbb{D}[\sqrt{2}]$

The following lemma can be used to reduce the problem of †-decomposability in $\mathbb{D}[\sqrt{2}]$ to †-decomposability in $\mathbb{Z}[\sqrt{2}]$.

Lemma C.25. *Consider $\xi \in \mathbb{D}[\sqrt{2}]$. Then ξ is †-decomposable if and only if $\sqrt{2}\,\xi$ is †-decomposable.*

Proof. Recall that $\delta = 1 + \omega$ satisfies $\delta^\dagger \delta = \lambda\sqrt{2} \sim \sqrt{2}$. Also note that δ is invertible in $\mathbb{D}[\omega]$; specifically, $\delta^{-1} = \delta\lambda^{-1}\omega^{-1}/\sqrt{2} \in \mathbb{D}[\omega]$. Assume that ξ is †-decomposable, so that $t^\dagger t \sim \xi$. Letting $t' = \delta t$, we have $t'^\dagger t' = \delta^\dagger \delta t^\dagger t \sim \sqrt{2} t^\dagger t \sim \sqrt{2}\,\xi$, so that $\sqrt{2}\,\xi$ is †-decomposable. The converse is proved similarly, using δ^{-1} instead of δ. □

We can now prove Theorem 6.2, whose statement we reproduce here.

Theorem. *Let $\xi \in \mathbb{D}[\sqrt{2}]$. Note that $\xi^\bullet \xi \in \mathbb{D}$, so we can write $\xi^\bullet \xi = \frac{n}{2^\ell}$ for some $n \in \mathbb{Z}$ and $\ell \in \mathbb{N}$. There exists a probabilistic algorithm which, given ξ and, in case $n \neq 0$, a prime factorization of n, determines whether or not the equation $t^\dagger t = \xi$ has a solution, and finds a solution if there is one. Moreover, the expected runtime of this algorithm is polynomial in the size of n.*

Proof. Let ξ be an element of $\mathbb{D}[\sqrt{2}]$ with $\xi^\bullet \xi = \frac{n}{2^\ell}$ for some $n \in \mathbb{Z}$. First, the algorithm can easily check whether ξ is doubly positive, and if it is not, there is no solution. Next, if $n = 0$, then $\xi = 0$, and $t = 0$ is obviously a solution, so there is nothing to do. Otherwise, let $\xi' = \sqrt{2}^\ell \xi$. Since $\xi' \in \mathbb{Z}[\sqrt{2}]$ and $\xi'^\bullet \xi' = n$, and a factorization of n is given, we can use Proposition C.24 to efficiently determine whether $s^\dagger s \sim \xi'$ has a solution. By Lemma C.25 this is the case if and only if the equation $t^\dagger t \sim \xi$ also has a solution, in which case it is given by $t = \delta^{-\ell} s$. Finally, since ξ is doubly positive, we can solve $t'^\dagger t' = \xi$ by Lemma C.16. □

We conclude this appendix with a useful fact: the algorithm of Theorem 6.2 always succeeds in case n is a prime that is congruent to 1 modulo 8.

Proposition C.26. *Let $\xi = \mathbb{D}[\sqrt{2}]$ be doubly positive, with $\xi^\bullet \xi = \frac{n}{2^\ell}$ for some $n \in \mathbb{Z}$. If n is prime and $n \equiv 1 \pmod 8$, then the equation $t^\dagger t = \xi$ has a solution.*

Proof. Let $\xi' = \sqrt{2}^\ell \xi$. Then $\xi' \in \mathbb{Z}[\sqrt{2}]$ and $\xi'^\bullet \xi' = n$. Then ξ' is prime by Lemma C.4 and †-decomposable by Lemma C.20. By Lemma C.25, ξ is †-decomposable, so that the equation $t^\dagger t = \xi$ can be solved by Lemma C.16. □

D Proof of Lemma 8.4

Definition D.1. Recall that $\delta = 1 + \omega$. Every element $u \in \mathbb{D}[\omega]$ can be written in the form

$$u = \frac{1}{\delta^k}(a\omega^3 + b\omega^2 + c\omega + d), \tag{45}$$

where $a, b, c, d \in \mathbb{Z}$ and $k \geqslant 0$. The smallest k such that u can be written in this form is called the *least δ-exponent* of u.

Remark D.2. A calculation shows that $\frac{1}{\delta}(a\omega^3 + b\omega^2 + c\omega + d)$ is equal to

$$\frac{1}{2}\left[\,(a - b + c - d)\omega^3 + (a + b - c + d)\omega^2 + (-a + b + c - d)\omega + (a - b + c + d)\,\right].$$

It follows that an element $a\omega^3 + b\omega^2 + c\omega + d \in \mathbb{Z}[\omega]$ is divisible by δ if and only if $a + b + c + d$ is even.

We can now prove Lemma 8.4, whose statement we reproduce here.

Lemma. *Each of the numbers n produced in step 2(a) of Algorithm 7.6 satisfies $n \geqslant 0$, and either $n = 0$ or $n \equiv 1 \pmod 8$.*

Proof. Recall that in step 2(a) of Algorithm 7.6, we are given $u \in \mathbb{D}[\omega]$ such that $u \in \overline{\mathcal{D}}$ and $u^\bullet \in \overline{\mathcal{D}}$. We let $\xi = 1 - u^\dagger u$ and write $\xi^\bullet \xi = \frac{n}{2^\ell}$, where $\ell \geqslant 0$ is minimal. We must show that $n \geqslant 0$, and that either $n = 0$ or $n \equiv 1 \pmod 8$.

The first claim is trivial, since by assumption $u^\dagger u \leqslant 1$ and $(u^\dagger u)^\bullet \leqslant 1$, and therefore $\xi, \xi^\bullet \geqslant 0$, which implies $n \geqslant 0$. For the second claim, write u in the form (45), with least δ-exponent k.

- Case 1: $k \leqslant 1$. In this case, one can show by a direct calculation (for example with the help of the algorithm from Proposition 5.21) that the scaled two-dimensional grid problem only has the 9 solutions $u = 0$ and $u = \omega^j$, where $j = 0, \ldots, 7$. In the first case, $n = 1$, and in the remaining 8 cases, $n = 0$, so the claim follows.

- Case 2: $k \geqslant 2$. We calculate
$$u^\dagger u = \frac{1}{(\delta^\dagger \delta)^k}(A + B\sqrt{2}),$$
where $A = a^2+b^2+c^2+d^2$ and $B = cd+bc+ab-da$. Note that, since k is the least δ-exponent of u, Remark D.2 implies that $a+b+c+d$ is odd. It easily follows that A is odd; moreover, since $B \equiv (a+c)(b+d) \pmod 2$, it also follows that B is even. Also note that $(\delta^\dagger\delta)^k = (\lambda\sqrt{2})^k$ is an element of $\mathbb{Z}[\sqrt{2}]$, and is divisible by 2 since $k \geqslant 2$. Therefore, we have $(\delta^\dagger\delta)^k = C + D\sqrt{2}$ for some $C, D \in \mathbb{Z}$ where C, D are both even. We further calculate
$$\xi = 1 - u^\dagger u = \frac{1}{(\delta^\dagger\delta)^k}(C + D\sqrt{2} - A - B\sqrt{2}) = \frac{1}{(\delta^\dagger\delta)^k}(x + y\sqrt{2}),$$
where $x = C - A$ is odd and $y = D - B$ is even. Noting that $(\delta^\dagger\delta)^\bullet(\delta^\dagger\delta) = 2$ and that $x^2 - 2y^2 \equiv 1 \pmod 8$, we therefore have
$$\xi^\bullet \xi = \frac{1}{((\delta^\dagger\delta)^\bullet(\delta^\dagger\delta))^k}(x^2 - 2y^2) = \frac{1}{2^k}(x^2 - 2y^2).$$
It follows that $\ell = k$ and $n = x^2 - 2y^2$, and therefore $n \equiv 1 \pmod 8$, which was to be shown. $\square$

E Proof of Lemma 8.6

Lemma E.1. *(a)* $\sqrt{x} \geqslant \ln x$ *for all* $x > 0$.

(b) $3\sqrt{x} \geqslant (\ln x)^2$ *for all* $x \geqslant 1$.

(c) $(1 - \frac{a}{x})^x \leqslant e^{-a}$, *for all* $x \geqslant a > 0$.

Proof. By elementary calculus. For (a) and (b), note that the functions $f(x) = \ln x/\sqrt{x}$ and $g(x) = (\ln x)^2/3\sqrt{x}$, on the given domains, take their maxima at $x = e^2$ and $x = e^4$, respectively, and in both cases, the maximum is less than 1. For (c), first note that for all $z \in \mathbb{R}$, $z + 1 \leqslant e^z$; the claim follows by letting $z = -\frac{a}{x}$. $\square$

We can now prove Lemma 8.6, whose statement we reproduce here.

Lemma. *Let $b > 0$ be an arbitrary fixed constant. Then for $a \geqslant 1$,*
$$\sum_{x=1}^{\infty}\left(1 - \frac{1}{a + b\ln x}\right)^x = O(a).$$

Proof. Let
$$x_0 = 16a^2 + 144b^2. \tag{46}$$
We claim that for all $x \geqslant x_0$,
$$\left(1 - \frac{1}{a + b\ln x}\right)^x \leqslant \frac{1}{x^2}. \tag{47}$$
Indeed, from $x \geqslant 16a^2$, we get
$$\frac{\sqrt{x}}{2} \geqslant 2a. \tag{48}$$
From $x \geqslant 144b^2$, we get
$$\frac{\sqrt{x}}{6} \geqslant 2b. \tag{49}$$

Combining (48) and (49) with Lemma E.1(a) and (b), we have

$$x = \frac{\sqrt{x}}{2} \cdot \sqrt{x} + \frac{\sqrt{x}}{6} \cdot 3\sqrt{x} \geqslant 2a \ln x + 2b(\ln x)^2 = 2 \ln x(a + b \ln x),$$

hence

$$\frac{1}{a + b \ln x} \geqslant \frac{2 \ln x}{x},$$

hence

$$\left(1 - \frac{1}{a + b \ln x}\right)^x \leqslant \left(1 - \frac{2 \ln x}{x}\right)^x \leqslant e^{-2 \ln x} = \frac{1}{x^2}$$

where the final inequality uses Lemma E.1(c). This finishes the proof of (47). The lemma now immediately follows, because we have

$$\begin{aligned}
\sum_{x=1}^{\infty} \left(1 - \frac{1}{a + b \ln x}\right)^x &= \sum_{x=1}^{\lfloor x_0 \rfloor} \left(1 - \frac{1}{a + b \ln x}\right)^x + \sum_{x=\lfloor x_0 \rfloor + 1}^{\infty} \left(1 - \frac{1}{a + b \ln x}\right)^x \\
&\leqslant \sum_{x=1}^{\lfloor x_0 \rfloor} \left(1 - \frac{1}{a + b \ln x_0}\right)^x + \sum_{x=\lfloor x_0 \rfloor + 1}^{\infty} \frac{1}{x^2} \\
&\leqslant \sum_{x=0}^{\infty} \left(1 - \frac{1}{a + b \ln x_0}\right)^x + \sum_{x=1}^{\infty} \frac{1}{x^2} \\
&= a + b \ln x_0 + \frac{\pi^2}{6} = a + b \ln(16a^2 + 144b^2) + \frac{\pi^2}{6} = O(a).
\end{aligned}$$

□

References

[1] A. Bocharov, Y. Gurevich, and K. M. Svore. Efficient decomposition of single-qubit gates into V basis circuits. *Phys. Rev. A*, 88:012313 (13 pages), 2013. Also available from `arXiv:1303.1411`.

[2] A. Bocharov, M. Roetteler, and K. M. Svore. Efficient synthesis of probabilistic quantum circuits with fallback. *Physical Review A*, 91:052317, May 2015. Also available from `arXiv:1409.3552`.

[3] A. Bocharov, M. Roetteler, and K. M. Svore. Efficient synthesis of universal repeat-until-success quantum circuits. *Phys. Rev. Lett.*, 114:080502, Feb. 2015. Also available from `arXiv:1404.5320`.

[4] G. Duclos-Cianci and K. M. Svore. Distillation of nonstabilizer states for universal quantum computation. *Phys. Rev. A*, 88:042325 (7 pages), 2013. Also available from `arXiv:1210.1980`.

[5] A. G. Fowler. Constructing arbitrary Steane code single logical qubit fault-tolerant gates. *Quantum Information and Computation*, 11(9–10):867–873, 2011. Also available from `arXiv:quant-ph/0411206`.

[6] B. Giles and P. Selinger. Exact synthesis of multiqubit Clifford+T circuits. *Physical Review A*, 87:032332, 2013. Also available from `arXiv:1212.0506`.

[7] B. Giles and P. Selinger. Remarks on Matsumoto and Amano's normal form for single-qubit Clifford+T operators. `arXiv:1312.6584`, Dec. 2013.

[8] V. Kliuchnikov, A. Bocharov, and K. M. Svore. Asymptotically optimal topological quantum compiling. *Phys. Rev. Lett.*, 112:140504, Apr 2014. Also available from `arXiv:1310.4150`.

[9] V. Kliuchnikov, D. Maslov, and M. Mosca. Asymptotically optimal approximation of single qubit unitaries by Clifford and T circuits using a constant number of ancillary qubits. *Phys. Rev. Lett.*, 110:190502 (5 pages), 2013. Also available from `arXiv:1212.0822v2`.

[10] V. Kliuchnikov, D. Maslov, and M. Mosca. Fast and efficient exact synthesis of single qubit unitaries generated by Clifford and T gates. *Quantum Information and Computation*, 13(7–8):607–630, 2013. Also available from `arXiv:1206.5236v4`.

[11] V. Kliuchnikov, D. Maslov, and M. Mosca. Practical approximation of single-qubit unitaries by single-qubit quantum Clifford and T circuits. *IEEE Transactions on Computers*, 65(1):161–172, Jan 2016. Also available from `arXiv:1212.6964v2`.

[12] M. O. Rabin. Probabilistic algorithms in finite fields. *SIAM Journal on Computing*, 9(2):273–280, 1980.

[13] N. J. Ross and P. Selinger. Exact and approximate synthesis of quantum circuits, version 0.3.0.1. Software implementation, `http://www.mathstat.dal.ca/~selinger/newsynth/`, 2015.

[14] P. Selinger. Efficient Clifford+T approximation of single-qubit operators. *Quantum Information and Computation*, 15(1-2):159–180, Jan. 2015. Also available from `arXiv:1212.6253`.

[15] P. W. Shor. Algorithms for quantum computation: discrete logarithms and factoring. In *Proceedings of the 35th Annual Symposium on Foundations of Computer Science*, pages 124–134, 1994. Also available from `arXiv:quant-ph/9508027`.

[16] N. Wiebe and V. Kliuchnikov. Floating point representations in quantum circuit synthesis. *New Journal of Physics*, 15:093041 (24 pages), 2013. Also available from `arXiv:1305.5528`.