Recovering Short Generators of Principal Ideals in Cyclotomic Rings

Ronald Cramer* Léo Ducas† Chris Peikert‡ Oded Regev§

February 25, 2016

Abstract

A handful of recent cryptographic proposals rely on the conjectured hardness of the following problem in the ring of integers of a cyclotomic number field: given a basis of a principal ideal that is guaranteed to have a "rather short" generator, find such a generator. Recently, Bernstein and Campbell-Groves-Shepherd sketched potential attacks against this problem; most notably, the latter authors claimed a *polynomial-time quantum* algorithm. (Alternatively, replacing the quantum component with an algorithm of Biasse and Fieker would yield a *classical subexponential-time* algorithm.) A key claim of Campbell *et al.* is that one step of their algorithm—namely, decoding the *log-unit* lattice of the ring to recover a short generator from an arbitrary one—is classically efficient (whereas the standard approach on general lattices takes exponential time). However, very few convincing details were provided to substantiate this claim.

In this work, we clarify the situation by giving a rigorous proof that the log-unit lattice is indeed efficiently decodable, for any cyclotomic of prime-power index. Combining this with the quantum algorithm from a recent work of Biasse and Song confirms the main claim of Campbell *et al.* Our proof consists of two main technical contributions: the first is a geometrical analysis, using tools from analytic number theory, of the standard generators of the group of cyclotomic units. The second shows that for a wide class of typical distributions of the short generator, a standard lattice-decoding algorithm can recover it, given any generator.

By extending our geometrical analysis, as a second main contribution we obtain an efficient algorithm that, given any generator of a principal ideal (in a prime-power cyclotomic), finds a $2^{\tilde{O}(\sqrt{n})}$-approximate shortest vector in the ideal. Combining this with the result of Biasse and Song yields a quantum polynomial-time algorithm for the $2^{\tilde{O}(\sqrt{n})}$-approximate Shortest Vector Problem on principal ideal lattices.

*Cryptology Group, CWI, Amsterdam, The Netherlands & Mathematical Institute, Leiden University, The Netherlands. Email: cramer@cwi.nl, cramer@math.leidenuniv.nl

†Cryptology Group, CWI, Amsterdam, The Netherlands. Supported by an NWO Free Competition Grant. Email: ducas@cwi.nl

‡Department of Computer Science and Engineering, University of Michigan. Much of this work was done while the author was at the Georgia Institute of Technology. This material is based upon work supported by the National Science Foundation under CAREER Award CCF-1054495, by DARPA under agreement number FA8750-11-C-0096, and by the Alfred P. Sloan Foundation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation, DARPA or the U.S. Government, or the Sloan Foundation. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon.

§Courant Institute of Mathematical Sciences, New York University. Supported by the Simons Collaboration on Algorithms and Geometry and by the National Science Foundation (NSF) under Grant No. CCF-1320188.

1 Introduction

Over the past several years, *lattices* have emerged as an attractive foundation for cryptography. The most efficient (and potentially practical) lattice-based cryptosystems are related to *ideal lattices*, which correspond to ideals in certain families of rings, e.g., $\mathbb{Z}[X]/(X^{2^k}+1)$. Representative works include [HPS98, Mic02, LMPR08, Gen09, LPR10].

More recently, a handful of cryptographic constructions have relied directly on *principal* ideals that have "*relatively short*" generators, which serve as secret keys.[1] These include a simplified variant of Gentry's original fully homomorphic encryption scheme [Gen09] due to Smart and Vercauteren [SV10], the closely related Soliloquy encryption scheme [CGS14], and candidate cryptographic multilinear maps [GGH13, LSS14]. Breaking these systems is no harder than solving the following problem, which we call the *Short Generator of a Principal Ideal Problem* (SG-PIP): given some $\mathbb{Z}$-basis of an ideal that is guaranteed to have a "short" generator g, find a sufficiently short generator (not necessarily g itself).

Potential attacks on SG-PIP in certain rings were sketched by Bernstein [Ber14b] and Campbell, Groves, and Shepherd [CGS14]. The basic structure of the attacks, which appears to be folklore in computational number theory, consists of two main parts:

- First, given a $\mathbb{Z}$-basis of the principal ideal, find some arbitrary (not necessarily short) generator of the ideal. For this task, which is known as the *Principal Ideal Problem* (PIP), the state of the art is an algorithm of Biasse and Fieker [BF14, Bia14], whose running time has only a subexponential $2^{n^{2/3+\epsilon}}$ dependence on n, the degree of the ring (over $\mathbb{Z}$). In addition, building on the recent work of Eisenträger *et al.* [EHKS14], polynomial-time *quantum* algorithms for PIP have recently been described in two independent works [CGS14, BS15], the latter of which provides a fully rigorous treatment.

- Second, transform the generator found in the previous phase into a *short* generator, thereby recovering the secret key, or its functional equivalent. The standard approach casts this task as a *closest vector problem* (CVP) on the Dirichlet "log-unit" lattice.

In this work, we focus entirely on the second phase, i.e., on recovering a short generator from any generator. At first, one might suspect that this is a hard problem: in general, the fastest known algorithms for CVP (even allowing quantum) run in exponential $2^{\Omega(n)}$ time [MV10, ADS15], or in less time but with much weaker guarantees on the solution quality (e.g., [LLL82, Bab85, Sch87]). In addition, Bernstein [Ber14b] suggested an algebraic approach that may yield slightly subexponential running times in number fields having many subfields, but it remains to be seen if this proposal can be carried through. Regardless of the method used, it is not obvious *a priori* whether solving CVP on the log-unit lattice yields a sufficiently short generator; much depends on the geometry of the lattice (in the relevant norm) and the quality of the solution.

A promising observation made by several researchers [CGS14, Ber14a] is that the CVP instances arising in the second phase have some implicit structure: the existence of a "rather short" generator (by choice of the secret key) implies that the target point is "somewhat close" to the log-unit lattice; CVP with such a distance guarantee is more commonly known as *bounded-distance decoding* (BDD) and is sometimes easier than the general case of CVP. Indeed, Garg, Gentry and Halevi [GGH13] gave an improved variant of the Gentry-Szydlo algorithm [GS02] which shows that in cyclotomic rings having power-of-two index, BDD on the log-unit lattice is efficiently solvable to within sub-polynomial $n^{-\log\log n}$ distance. However, this threshold is much too small to handle the BDD instances arising in cryptosystems.

Campbell, Groves, and Shepherd [CGS14] were the first to claim an efficient solution to the second phase above. In more detail, they asserted that in cyclotomic rings having power-of-two index, the second phase can

[1] A principal ideal in a commutative ring R is of the form $gR = \{g \cdot r : r \in R\}$ for some $g \in R$, called a *generator* of the ideal.

be accomplished simply by decoding the log-unit lattice using a standard algorithm such as LLL [LLL82]. However, this claim was not accompanied by a proof.[2] Nevertheless, experiments in cryptographically relevant choices of dimension have shown that decoding is indeed practically efficient [She14, Sch15], giving strong evidence that the approach of [CGS14] does indeed work.

Contributions. Our first main contribution is a rigorous proof showing that the second phase above can be solved in polynomial time, in any cyclotomic of prime-power index. Our proof is based on classical ideas and results from analytical number theory, along with some techniques from probability theory, and consists of two main technical contributions. First, in Section 3 we use standard tools from analytical number theory, such as bounds on *Dirichlet L-series*, to elucidate the geometry of a standard set of generators for the group of *cyclotomic units*. (The cyclotomic units correspond either to the log-unit lattice itself, or to a sublattice whose index is conjectured to be quite small.) Using this geometry, in Sections 4 and 5 we show that for a wide class of typical distributions of the secret generator—e.g., Gaussian-like distributions—the naïve "round-off" lattice-decoding algorithm [Len82, Bab85] (using the standard generators of the cyclotomic units) can be used to efficiently recover the secret short generator, given any generator of the ideal.[3] To complement these results, in Appendix B we give concrete numerical data demonstrating that the second phase succeeds for all practical choices of dimension.

Our second main contribution concerns the questions: in an *arbitrary* principal ideal (of a prime-power cyclotomic), how long can a shortest generator be? And how short of a generator can we find efficiently? In Section 6, we show that for an overwhelming majority of principal ideals, the shortest generator is a $2^{\tilde{\Theta}(\sqrt{n})}$ factor longer than the shortest nonzero vector in the ideal. Moreover, one can efficiently find a generator satisfying this bound, given an arbitrary generator. The first of these facts means that the principal ideals used in the aforementioned cryptographic applications are highly atypical, because their shortest generators are also nearly shortest vectors. The second fact implies that the $2^{\tilde{O}(\sqrt{n})}$-approximate Shortest Vector Problem (SVP) on arbitrary principal ideals reduces to the Principal Ideal Problem.

Implications and discussion. Combining our main contributions with known algorithms for PIP [BF14, Bia14, CGS14, BS15] (which are the computational bottleneck) yields the following two main implications:

- First, there is a quantum polynomial-time, or classical $2^{n^{2/3+\epsilon}}$-time, algorithm for SG-PIP, implying a key-recovery attack for the cryptographic constructions of [SV10, GGH13, LSS14, CGS14].
- Second, there is a quantum polynomial-time algorithm for $2^{\tilde{O}(\sqrt{n})}$-approximate SVP on *principal* ideals in any prime-power cyclotomic. (Note that we do not obtain any improvement over *classical* SVP algorithms, because $2^{n^{2/3}}$ time is sufficient to solve $2^{\tilde{O}(n^{1/3})}$-approximate SVP on arbitrary lattices [Sch87].)

In light of these, an important open problem is to obtain faster classical PIP algorithms, perhaps also using the guarantee that a short generator exists.

A natural question is what effect, if any, these attacks have on other ring-based problems, such as NTRU [HPS98] and ring-LWE [LPR10], which are the heart of many cryptosystems. Specifically, the

[2] The explanation given in [CGS14] is that the secret generator corresponds to a vector that is short relative to the determinant of the log-unit lattice. As far as we can tell, this by itself is not enough to substantiate the claim, as it ignores the geometry of the log-unit lattice and the quality of the output produced by the LLL algorithm.

[3] Strictly speaking, the polynomial running time of this algorithm depends on a number-theoretic conjecture regarding the class numbers $h^+(m)$; see Section 2.4 for details.

theoretical foundation of the ring-LWE problem is the conjectured quantum hardness of approximate-SVP on *arbitrary* ideals, usually in a cyclotomic ring and for (near-)polynomial approximation factors. As far as we can tell, the above-described algorithms do not appear to affect this foundation: the first crucially relies on the existence of an "unusually short" generator, the second is inherently limited to relatively large SVP approximation factors, and both apply only to *principal* ideals. An important question is whether these barriers can be overcome, and if so, whether this leads to attacks on ring-LWE or NTRU themselves.

In a complementary direction, another interesting question is whether the above attacks can be extended to other families of *non-cyclotomic* rings, such as those suggested in [Ber14b]. For this it may suffice to find (by analysis, computation, or both) a suitably good basis of the log-unit lattice, or of a sublattice of not too large index.

Acknowledgments. We thank Dan Bernstein, Jean-François Biasse, Sean Hallgren, Sorina Ionica, Dimitar Jetchev, Paul Kirchner, Shinya Okumara, René Schoof, Alice Silverberg, and Harold M. Stark for comments and many insightful conversations on topics related to this work. We also especially thank Dan Shepherd [She14] for explaining many additional details about the claims made in [CGS14], and for sharing other helpful observations.

2 Preliminaries

We denote column vectors by lower-case bold letters (e.g., $\mathbf{x}$) and matrices by upper-case bold letters (e.g., $\mathbf{X}$). We often adopt the nonstandard, but very useful, convention of indexing rows and columns by particular finite sets (not necessarily $\{1, \ldots, n\}$), and identify a matrix with its indexed set of column vectors. The canonical scalar product over $\mathbb{R}^n$ and over $\mathbb{C}^n$ is denoted $\langle \cdot, \cdot \rangle$, and $\|\cdot\|$ denotes the Euclidean norm. For a complex number $z \in \mathbb{C}$, $\overline{z}$ denotes its complex conjugate, and $|z| = \sqrt{z \cdot \overline{z}}$ denotes its magnitude.

2.1 Lattices and BDD

A *lattice* $\mathcal{L}$ is a discrete additive subgroup of $\mathbb{R}^n$ for some positive integer n. The *minimum distance* of $\mathcal{L}$ is $\lambda_1(\mathcal{L}) := \min_{\mathbf{v} \in \mathcal{L} \setminus \{\mathbf{0}\}} \|\mathbf{v}\|$, the length of a shortest nonzero lattice vector. Every lattice is generated as the integer linear combinations of some (non-unique) $\mathbb{R}$-linearly independent *basis* vectors $\mathbf{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$, as $\mathcal{L} = \mathcal{L}(\mathbf{B}) := \{\sum_{j=1}^{k} \mathbb{Z} \cdot \mathbf{b}_j\}$, where $k \leq n$ is called the *rank* of the lattice.

Letting span denote the $\mathbb{R}$-linear span of a set, the *dual* basis $\mathbf{B}^\vee = \{\mathbf{b}_1^\vee, \ldots, \mathbf{b}_k^\vee\} \subset \mathrm{span}(\mathbf{B})$ and dual lattice $\mathcal{L} = \mathcal{L}(\mathbf{B}^\vee)$ are defined to satisfy $\langle \mathbf{b}_j^\vee, \mathbf{b}_{j'} \rangle = \delta_{j,j'}$ for all j, j', where the Kronecker delta $\delta_{j,j'} = 1$ if $j = j'$, and is 0 otherwise. In other words, $\mathbf{B}^t \cdot \mathbf{B}^\vee = (\mathbf{B}^\vee)^t \cdot \mathbf{B}$ is the identity matrix.

In this work we deal with a computational problem on lattices called *bounded-distance decoding* (BDD): given a lattice basis $\mathbf{B} \subset \mathbb{R}^n$ of $\mathcal{L} = \mathcal{L}(\mathbf{B})$ and a target point $\mathbf{t} \in \mathrm{span}(\mathcal{L})$ with the guarantee that $\min_{\mathbf{v} \in \mathcal{L}} \|\mathbf{v} - \mathbf{t}\| \leq r$ for some known $r < \lambda_1(\mathcal{L})/2$, find the unique $\mathbf{v} \in \mathcal{L}$ closest to $\mathbf{t}$ (i.e., such that $\|\mathbf{v} - \mathbf{t}\| \leq r$). In fact, in our context $\mathbf{B}$ and r will be fixed in advance, and $\mathbf{t}$ is the only input that may vary.

A standard approach to solve BDD (and related problems) is the "round-off" algorithm of [Bab85], which simply returns $\mathbf{B} \cdot \lfloor (\mathbf{B}^\vee)^t \cdot \mathbf{t} \rceil$, where the rounding function $\lfloor c \rceil := \lfloor c + \frac{1}{2} \rfloor \in \mathbb{Z}$ is applied to each coordinate independently. (Notice that $(\mathbf{B}^\vee)^t \cdot \mathbf{t}$ is the coefficient vector of $\mathbf{t}$ with respect to basis $\mathbf{B}$.) We recall the following standard fact about this algorithm, and include a brief proof for completeness.

Claim 2.1. *Let $\mathcal{L} \subset \mathbb{R}^n$ be a lattice with basis $\mathbf{B}$, and let $\mathbf{t} = \mathbf{v} + \mathbf{e} \in \mathbb{R}^n$ for some $\mathbf{v} \in \mathcal{L}$, $\mathbf{e} \in \mathbb{R}^n$. If $\langle \mathbf{b}_j^\vee, \mathbf{e} \rangle \in [-\frac{1}{2}, \frac{1}{2})$ for all j, then on input $\mathbf{t}$ and basis $\mathbf{B}$, the round-off algorithm outputs $\mathbf{v}$.*

Proof. Because $\mathbf{v} = \mathbf{B}\mathbf{z}$ for some integer vector $\mathbf{z}$, we have $(\mathbf{B}^\vee)^t \cdot \mathbf{t} = \mathbf{z} + (\mathbf{B}^\vee)^t \cdot \mathbf{e}$, so by hypothesis on the $\langle \mathbf{b}_j, \mathbf{e} \rangle$, we have $\lfloor (\mathbf{B}^\vee)^t \cdot \mathbf{t} \rceil = \mathbf{z}$. The claim follows. □

2.2 Circulant Matrices

We recall some standard facts about *circulant* matrices for a finite abelian group $(G, \cdot)$, and their relationship with the *characters* of the group. See, e.g., see [Lan02] for further details and proofs.

Definition 2.2 (Circulant matrix). For a vector $\mathbf{a} = (a_g)_{g \in G}$ indexed by G, the *G-circulant* matrix associated with $\mathbf{a}$ is the G-by-G matrix whose (i, j)th entry is $a_{ij^{-1}}$.

Note that the transpose of any G-circulant matrix (associated with $(a_g)_{g \in G}$) is also a G-circulant matrix (associated with $(a_{g^{-1}})_{g \in G}$).

Definition 2.3 (Character group). A *character* is a group morphism $\chi \colon G \to \{u \in \mathbb{C} : |u| = 1\}$, i.e., $\chi(g \cdot h) = \chi(g) \cdot \chi(h)$ for all $g, h \in G$. The *character group* $(\hat{G}, \cdot)$ is the set of characters of G, with the group operation being the usual multiplication of functions, i.e., $(\chi \cdot \psi)(g) = \chi(g) \cdot \psi(g)$.

A basic fact is that $|\hat{G}| = |G|$. Notice that for a character $\chi \in \hat{G}$, we have $\overline{\chi(g)} = \chi(g)^{-1} = \chi(g^{-1})$. We identify χ with the vector $(\chi(g))_{g \in G}$. Then all characters χ have Euclidean norm $\|\chi\| = \sqrt{|G|}$, because

$$\langle \chi, \chi \rangle = \sum_{g \in G} \chi(g) \cdot \overline{\chi(g)} = \sum_{g \in G} 1 = |G|.$$

Moreover, distinct characters χ, ψ are orthogonal:

$$\langle \chi, \psi \rangle = \sum_{g \in G} \chi(g) \cdot \overline{\psi(g)} = \sum_{g \in G} (\chi \cdot \psi^{-1})(g) = 0.$$

Therefore, the complex G-by-$\hat{G}$ matrix

$$\mathbf{P}_G := |G|^{-1/2} \cdot \big(\chi(g)\big)_{g \in G, \chi \in \hat{G}}$$

is unitary, i.e., $\mathbf{P}_G^{-1} = \mathbf{P}_G^*$, the conjugate transpose of $\mathbf{P}_G$.

Lemma 2.4. *A complex matrix $\mathbf{A}$ is G-circulant if and only if the $\hat{G}$-by-$\hat{G}$ matrix $\mathbf{P}_G^{-1} \cdot \mathbf{A} \cdot \mathbf{P}_G$ is diagonal; equivalently, the columns of $\mathbf{P}_G$ are the eigenvectors of $\mathbf{A}$. If $\mathbf{A}$ is the G-circulant matrix associated with $\mathbf{a} = (a_g)_{g \in G}$, its eigenvalue corresponding to $\chi \in \hat{G}$ is $\lambda_\chi = \langle \mathbf{a}, \chi \rangle = \sum_{g \in G} a_g \cdot \overline{\chi(g)}$.*

It follows that every row and column of $\mathbf{A}$ has squared Euclidean norm

$$\|\mathbf{a}\|^2 = \|\mathbf{P}_G^* \cdot \mathbf{a}\|^2 = |G|^{-1} \cdot \sum_{\chi \in \hat{G}} |\lambda_\chi|^2.$$

It also follows that $\mathbf{A}^{-1}$ (when defined) is G-circulant, with eigenvalue λ_χ^{-1} for eigenvector χ.

Proof. Suppose that $\mathbf{A}$ is G-circulant, and let $\chi \in \hat{G}$ be a character of G. Then

$$(\mathbf{A} \cdot \chi)_g = \sum_{h \in G} a_{gh^{-1}} \cdot \chi(h) = \Big(\sum_{k \in G} a_k \cdot \overline{\chi(k)} \Big) \cdot \chi(g),$$

where in the final equality we have substituted $k = gh^{-1}$ and used $\chi(h) = \overline{\chi(k)} \cdot \chi(g)$. So $\mathbf{A} \cdot \chi = \lambda_\chi \cdot \chi$.

For the other direction, it suffices by linearity to show that $\mathbf{A}_\chi = \mathbf{P}_G \cdot \mathbf{D}_\chi \cdot \mathbf{P}_G^{-1}$ is G-circulant for every $\chi \in \hat{G}$, where $\mathbf{D}_\chi$ is the diagonal $\hat{G}$-by-$\hat{G}$ matrix with 1 in its (χ, χ)th entry and zeros elsewhere. Indeed, by definition of $\mathbf{P}_G$ and because $\mathbf{P}_G^{-1} = \mathbf{P}_G^*$, the (i, j)th entry of $\mathbf{A}_\chi$ is simply $|G|^{-1} \cdot \chi(i) \cdot \overline{\chi(j)} = |G|^{-1} \cdot \chi(ij^{-1})$, which depends only on ij^{-1} as required. □

2.3 Dirichlet Characters and L-Series

A *Dirichlet character* χ is a character of $\mathbb{Z}_k^*$ for some positive integer k. Note that if $k|\ell$ then χ induces a character of $\mathbb{Z}_\ell^*$ via the natural morphism $\mathbb{Z}_\ell^* \to \mathbb{Z}_k^*$, so we can equivalently view χ as being defined modulo either k or ℓ. The *conductor* f_χ of χ is the smallest positive f such that χ is induced by a Dirichlet character modulo f. The character is said to be *even* if $\chi(-1) = 1$; note that the even Dirichlet characters correspond with the characters of $\mathbb{Z}_k^*/\{\pm 1\}$. The character is said to be *quadratic* if all its values are real (i.e., ± 1), and it is not the constant 1 character (which is known as the principal character). Following the convention used in [Was97], we often implicitly extend χ to a completely multiplicative function from $\mathbb{Z}$ to $\mathbb{C}$, by considering it as modulo its conductor k (i.e., as a primitive character) and letting $\chi(a) = 0$ if $\gcd(a, k) > 1$.

Definition 2.5 (Dirichlet L-Series). For a Dirichlet character χ, the Dirichlet *L-function* $L(\cdot, \chi)$ is defined as the formal series

$$L(s, \chi) = \sum_{k \geq 1} \frac{\chi(k)}{k^s}.$$

For any Dirichlet character χ, the series $L(s, \chi)$ is absolutely convergent for all $s \in \mathbb{C}$ with $\Re(s) > 1$. It is also known that $L(1, \chi)$ converges and is nonzero for any non-principal Dirichlet character (i.e., $\chi \neq 1$). We have the following asymptotic bounds on its value; we will only use the lower bounds.

Theorem 2.6. *There exists a $C > 0$ such that, for any non-quadratic character χ of conductor $f > 1$,*

$$\frac{1}{\ell(f)} \leq |L(1, \chi)| \leq \ell(f) \quad \textit{where } \ell(f) = C \ln f. \tag{1}$$

Moreover, for any quadratic character χ,

$$|L(1, \chi)| \geq \frac{1}{C\sqrt{f}}. \tag{2}$$

Equation (1) can be traced back to Landau [Lan27], and improving the constant C is an active field of research [Lou15]. Equation (2) is also classical and follows from Dirichlet's class number formula (see, e.g., [MV06, Section 4.4]). We note that under the Generalized Riemann Hypothesis, the bound in Eq. (1) can be improved to $\ell(f) = C \ln \ln f$, and holds for both quadratic and non-quadratic characters (see, e.g., [LLS15]).

2.4 Cyclotomic Number Fields and the Log-Unit Lattice

Cyclotomic number fields. Let L be a field. An element $\zeta \in L$ is a root of unity if $\zeta^m = 1$ for some positive integer m. The order of a root of unity $\zeta \in L$ is the order of the finite multiplicative subgroup of L^* generated by ζ. A primitive mth root of unity in L is a root of unity $\zeta \in L$ of order m. Note that if $\zeta \in L$ is a primitive mth root of unity, then the polynomial $X^m - 1 \in L[X]$ factors as $\prod_{i=0}^{m-1}(X - \zeta^i)$ over $L[X]$. Also note that the complete set of primitive mth roots in L consists of the powers ζ^j for $j \in \mathbb{Z}_m^*$.

An algebraic number field K is an extension field of the rationals $\mathbb{Q}$ such that its dimension $[K : \mathbb{Q}]$ as a $\mathbb{Q}$-vector space (i.e., its degree) is finite. If $\Omega \supset K$ is an extension field such that Ω is algebraically closed over $\mathbb{Q}$, then there are exactly $[K : \mathbb{Q}]$ field embeddings of K into Ω.[4] An algebraic number field is Galois if the order of its automorphism group equals its degree.[5] A number field K is cyclotomic if $K = \mathbb{Q}(\zeta)$ for some root of unity $\zeta \in K$. Its degree is $\varphi(m)$, where $\varphi(\cdot)$ is the Euler totient function and m is the order of ζ, and its ring of integers R is monogenic, i.e., $R = \mathbb{Z}[\zeta]$. We let U denote the cyclic (multiplicative) subgroup of mth roots of unity, which is generated by ζ.

A cyclotomic number field is Galois. If $K = \mathbb{Q}(\zeta)$ is a cyclotomic number field with $\zeta \in K$ an mth primitive root of unity then each automorphism is characterized by the assignment $\zeta \mapsto \zeta^j$ for some $j \in \mathbb{Z}_m^*$. As a consequence, if L is an extension field of a cyclotomic field K, then K is situated uniquely in L. For concreteness, we situate cyclotomic number fields in the complex numbers $\mathbb{C}$. Let m be a positive integer and define $\omega = \omega_m = \exp(2\pi\imath/m) \in \mathbb{C}$. Then ω is a primitive mth root of unity and $K = \mathbb{Q}(\omega)$ is the mth cyclotomic number field. The embeddings of K into the complex numbers (i.e., the automorphisms of K) are denoted σ_j for $j \in \mathbb{Z}_m^*$, where σ_j sends ω to ω^j. The concatenation $\sigma(a) = (\sigma_j(a))_{j \in \mathbb{Z}_m^*}$ of these embeddings is known as the *canonical embedding*, and is used to endow K with a geometry, e.g., $\|a\| := \|\sigma(a)\|$ for any $a \in K$.

Logarithmic embedding. The embeddings σ_i of K, being complex, come in conjugate pairs, i.e., $\sigma_j(x) = \overline{\sigma_{-j}(x)}$. We will mainly be concerned with their *magnitudes*, so we identify the pairs by indexing over the multiplicative quotient group $G := \mathbb{Z}_m^*/\{\pm 1\}$. We then have the *logarithmic embedding*, defined as

$$\begin{aligned} \mathrm{Log}\colon K &\to \mathbb{R}^{\varphi(m)/2} \\ a &\mapsto \big(\log|\sigma_i(a)|\big)_{i \in G}\ . \end{aligned}$$

The logarithmic embedding defines a group morphism, mapping the multiplicative group K^* to an additive subgroup of $\mathbb{R}^{\varphi(m)/2}$. The kernel of Log restricted to R^* is $\{\pm 1\} \cdot U$. The Dirichlet Unit Theorem (see [Sam70, Chapter 4.4, Theorem 1]) implies that $\Lambda = \mathrm{Log}(R^*)$, the image of the multiplicative unit group of R under the logarithmic embedding, is a full-rank lattice in the linear subspace of $\mathbb{R}^{\varphi(m)/2}$ orthogonal to the all-1s vector $\mathbf{1}$. We refer to Λ as the *log-unit lattice*.

Cyclotomic units. Let A be the multiplicative subgroup of K^* generated by $\pm\zeta$ and

$$z_j := \zeta^j - 1, \quad j \in \mathbb{Z}_m \setminus \{0\}.$$

Notice that $z_j = -\zeta^j \cdot z_{-j}$, so z_j and z_{-j} are equivalent modulo $\pm U$; in particular, $\mathrm{Log}(z_j) = \mathrm{Log}(z_{-j})$. The group of *cyclotomic units*, denoted C, is defined by

$$C = A \cap R^*\ .$$

The z_j given above are not necessarily units in R, and thus do not generate C. However, a closely related generating set, which we call the *canonical generators*, is given by the following lemma. Recall that $G = \mathbb{Z}_m^*/\{\pm 1\}$, and identify it with some canonical set of representatives in $\mathbb{Z}_m^*$.

[4] These embeddings are merely ring morphisms $\psi : K \to \Omega$. Each such ψ is automatically injective because K is a field. Also note that any such ψ fixes $\mathbb{Q}$ pointwise.

[5] An automorphism of a field L is a ring isomorphism $\psi\colon L \to L$. The automorphisms of L form a group with functional composition as the group operation.

Lemma 2.7 (Lemma 8.1 of [Was97]). *Let m be a prime power, and define $b_j := z_j/z_1 = (\zeta^j - 1)/(\zeta - 1)$. The group C of cyclotomic units is generated by $\pm\zeta$ and b_j for $j \in G \setminus \{1\}$.*

Notice that $\log C$ is a sublattice of Λ. As shown below, the index of Λ over $\log C$ is finite. In fact, it is $h^+(m)$, the *class number* of the real subfield $K^+ = \mathbb{Q}(\zeta + \bar{\zeta})$, defined as the index of the subgroup of principal fractional ideals in the multiplicative group of all fractional ideals (in K^+). The proof of this theorem is left as Exercise 8.5 in [Was97]. For completeness, we sketch the solution in Appendix A.

Theorem 2.8. *For a prime power $m > 2$, the index of the log-unit lattice Λ over $\log C$ is*

$$[\Lambda : \log C] = h^+(m).$$

Some facts and conjectures concerning h^+. For our purposes, we need $h^+(m)$ not to be very big. For all power-of-two m up to $m = 256$, and also for $m = 512$ under GRH, it is known that $h^+(m) = 1$ (see [Mil14]). Whether $h^+(m) = 1$ for all power-of-two m is known as Weber's class number problem, and is presented in the literature as a reasonable conjecture.

In the case of odd primes, it also appears that h^+ is quite small. Computations of Schoof [Sch03] and Miller [Mil15] show that $h^+(p) \leq 11$ for all primes $p \leq 241$. For powers of odd primes it has been conjectured (with support of the Cohen-Lenstra heuristic) that, for all but finitely many pairs (p, ℓ) where p is a prime, $h^+(p^{\ell+1}) = h^+(p^\ell)$ [BPR04]. A direct consequence is that $h^+(p^\ell)$ is bounded for a fixed p and increasing ℓ.

3 Geometry of the Canonical Generators

Throughout this section, let the cyclotomic index m be a prime power. Our goal here is to show that the canonical generators of the cyclotomic units, under the logarithmic embedding, are geometrically well-suited for bounded-distance decoding.

Recalling that $G = \mathbb{Z}_m^*/\{\pm 1\}$ is identified with some set of canonical representatives in $\mathbb{Z}_m^*$ and that $\log(b_j) = \log(b_{-j})$, define

$$\mathbf{b}_j = \log(b_j), \quad j \in G \setminus \{1\},$$

to be the log-embeddings of the canonical generators $b_j = (\zeta^j - 1)/(\zeta - 1)$ defined in Lemma 2.7. By Lemma 2.7, these $\mathbf{b}_j$ form a basis of the sublattice $\log C$, which by Theorem 2.8 has index $h^+(m)$ in Λ.

In order to apply the round-off algorithm and Claim 2.1 with this basis, we bound the norms $\|\mathbf{b}_j^\vee\|$ of the dual basis vectors. The remainder of this section is dedicated to proving the following theorem.

Theorem 3.1. *Let $m = p^k$ for a prime p, and let $\{\mathbf{b}_j^\vee\}_{j \in G\setminus\{1\}}$ denote the basis dual to $\{\mathbf{b}_j\}_{j\in G\setminus\{1\}}$. Then all $\|\mathbf{b}_j^\vee\|$ are equal, and*

$$\left\|\mathbf{b}_j^\vee\right\|^2 \leq 2k|G|^{-1} \cdot (\ell(m)^2 + O(1)) = O(m^{-1} \cdot \log^3 m)\ .$$

To prove the theorem we start by relating the basis vectors $\mathbf{b}_j$ to a certain G-circulant matrix. Recalling that $z_j = \zeta^j - 1$ is the numerator of b_j, define

$$\mathbf{z}_j := \log(z_j) = \mathbf{b}_j + \mathbf{z}_1\ . \tag{3}$$

Collect these vectors into a square G-by-G matrix $\mathbf{Z}$ whose jth column is $\mathbf{z}_{j^{-1}}$, and notice that its (i,j)th entry $\log|\omega^{i\cdot j^{-1}} - 1|$ is determined by $ij^{-1} \in G$ alone, so $\mathbf{Z}$ is the G-circulant matrix associated with $\mathbf{z}_1$. For each eigenvector $\chi \in \hat{G}$ of $\mathbf{Z}$, let $\lambda_\chi := \langle \mathbf{z}_1, \chi \rangle$ denote the corresponding eigenvalue.

Lemma 3.2. *For all $j \in G \setminus \{1\}$ we have*

$$\left\|\mathbf{b}_j^\vee\right\|^2 = |G|^{-1} \cdot \sum_{\chi \in \widehat{G} \setminus \{1\}} |\lambda_\chi|^{-2}. \tag{4}$$

Proof. Let $\mathbf{z}_j^\vee$ denote the vectors dual to the $\mathbf{z}_j$, i.e., the columns of $\mathbf{Z}^{-t}$. (As shown below in the proof of Theorem 3.1, $\mathbf{Z}^{-1}$ is indeed well defined because all eigenvalues λ_χ of $\mathbf{Z}$ are nonzero.)

We first claim that $\mathbf{b}_j^\vee$ is simply the projection of $\mathbf{z}_j^\vee$ orthogonal to $\mathbf{1}$, i.e., $\mathbf{b}_j^\vee = \mathbf{z}_j^\vee - |G|^{-1} \cdot \langle \mathbf{z}_j^\vee, \mathbf{1} \rangle \cdot \mathbf{1}$. Indeed, these vectors are all in $\mathrm{span}(\mathbf{b}_{j'})_{j'}$, the space orthogonal to $\mathbf{1}$, and moreover, for all $j, j' \in G \setminus \{1\}$ they satisfy

$$\langle \mathbf{z}_j^\vee - |G|^{-1} \cdot \langle \mathbf{z}_j^\vee, \mathbf{1} \rangle \cdot \mathbf{1}, \mathbf{b}_{j'} \rangle = \langle \mathbf{z}_j^\vee, \mathbf{b}_{j'} \rangle = \langle \mathbf{z}_j^\vee, \mathbf{z}_{j'} - \mathbf{z}_1 \rangle = \delta_{j,j'} - 0.$$

Now,

$$\left\|\mathbf{b}_j^\vee\right\|^2 = \left\|\mathbf{z}_j^\vee\right\|^2 - |G|^{-1} \cdot \langle \mathbf{z}_j^\vee, \mathbf{1} \rangle^2.$$

Recall by Lemma 2.4 that $\mathbf{Z}^{-t}$ is the G-circulant matrix associated with $\mathbf{z}_1^\vee$, which has eigenvalue $\lambda_\chi^{-1} = \langle \mathbf{z}_1^\vee, \chi \rangle$ for eigenvector $\chi \in \hat{G}$. By the remarks following Lemma 2.4, $\|\mathbf{z}_j^\vee\|^2 = |G|^{-1} \cdot \sum_{\chi \in \hat{G}} |\lambda_\chi|^{-2}$. The lemma follows by noting that $\langle \mathbf{z}_j^\vee, \mathbf{1} \rangle = \langle \mathbf{z}_1^\vee, \mathbf{1} \rangle = \lambda_1^{-1}$. □

We now provide an upper bound on the right-hand side of Equation (4). Our proof is similar to the proof that the cyclotomic units have finite index in the full group of units [Was97, Theorem 8.2].

Theorem 3.3 ([Was97, Lemma 4.8 and Theorem 4.9]). *Let χ be an even Dirichlet character of conductor $f > 1$, and let $\omega_f = \exp(2\pi\imath/f) \in \mathbb{C}$. Then*

$$\left| \sum_{a \in \mathbb{Z}_f^*} \overline{\chi(a)} \cdot \log|1 - \omega_f^a| \right| = \sqrt{f} \cdot |L(1, \chi)|.$$

For completeness, we briefly explain how the finite sum on the left hand side gives rise to an L-series, and refer to [Was97] for the details. Using the Taylor expansion

$$\log|1 - x| = -\sum_{k \geq 1} x^k/k \ ,$$

one gets a sum over finitely many a and infinitely many k of terms $\overline{\chi(a)} \cdot \omega_f^{ak}/k$. For a fixed k, the sum over a can easily be rewritten as $\tau(\chi) \cdot \chi(k)/k$, where $\tau(\chi)$ is a Gauss sum (see [Was97, Lemma 4.7]), which makes the Dirichlet L-function apparent.

Corollary 3.4. *Suppose $f > 1$ divides a prime power m. For any even Dirichlet character χ of conductor f,*

$$\left| \sum_{a \in \mathbb{Z}_m^*} \overline{\chi(a)} \cdot \log|1 - \omega_m^a| \right| = \sqrt{f} \cdot |L(1, \chi)|.$$

Proof. Let $\phi\colon \mathbb{Z}_m^* \to \mathbb{Z}_f^*$ be the map given by reduction modulo f. We have

$$\begin{aligned}\sum_{a\in\mathbb{Z}_m^*} \overline{\chi(a)}\cdot\log|1-\omega_m^a| &= \sum_{a\in\mathbb{Z}_f^*}\overline{\chi(a)}\sum_{\substack{b\in\mathbb{Z}_m^*\\ \phi(b)=a}}\log|1-\omega_m^b| \\ &= \sum_{a\in\mathbb{Z}_f^*}\overline{\chi(a)}\cdot\log\Big|\prod_{\substack{b\in\mathbb{Z}_m^*\\ \phi(b)=a}}(1-\omega_m^b)\Big| \\ &= \sum_{a\in\mathbb{Z}_f^*}\overline{\chi(a)}\cdot\log|1-\omega_f^a|,\end{aligned}$$

where in the last equality we have used the identity $\prod_{i\in\mathbb{Z}_n}(1-\omega_n^i Y) = 1-Y^n$ and $\omega_m^n = \omega_f$ with $n = m/f$. The claim follows by applying Theorem 3.3. □

We are now ready to complete the proof of the main theorem.

Proof of Theorem 3.1. Recall that the characters $\chi \in \hat{G}$ correspond to the even characters of $\mathbb{Z}_m^*$, because $\chi(\pm 1) = 1$. Also recall that by Lemma 2.4, the eigenvalues are

$$\lambda_\chi = \langle \mathbf{z}_1, \chi\rangle = \sum_{a\in G}\overline{\chi(a)}\cdot\log|1-\omega_m^a| = \frac{1}{2}\sum_{a\in\mathbb{Z}_m^*}\overline{\chi(\pm a)}\cdot\log|1-\omega_m^a|,$$

where the second equality holds because $|1-\omega_m^{-a}| = |1-\omega_m^a|$. Therefore, using Corollary 3.4 we have

$$|\lambda_\chi| = \frac{1}{2}\sqrt{f_\chi}\cdot|L(1,\chi)|\,, \tag{5}$$

and so by Lemma 3.2,

$$\left\|\mathbf{b}_j^\vee\right\|^2 = |G|^{-1}\cdot\sum_{\chi\in\widehat{G}\setminus\{1\}}|\lambda_\chi|^{-2} = 4|G|^{-1}\cdot\sum_{\chi\in\widehat{G}\setminus\{1\}} f_\chi^{-1}\cdot|L(1,\chi)|^{-2}\,.$$

We first consider the contribution to the sum coming from quadratic characters. When p is an odd prime, there is exactly one quadratic character (see [MV06, Section 9.3]), and it is of conductor p, hence by Equation (2) in Theorem 2.6, the contribution to the sum is $O(1)$ (assuming it is even; otherwise it does not participate in the sum). In the case $p = 2$ the contribution is also $O(1)$ since there are at most three quadratic characters (see again [MV06, Section 9.3]) and their conductor is bounded from above by an absolute constant. Finally, the contribution coming from non-quadratic characters is at most

$$\ell(m)^2\sum_{\chi\in\hat{G}\setminus\{1\}} f_\chi^{-1} \le \frac{k}{2}\cdot\ell(m)^2\,,$$

where we used Equation (1) in Theorem 2.6 and Claim 3.5 below. □

Claim 3.5. *Let $m = p^k$ for a prime p. Then, for $G = \mathbb{Z}_m^*/\{\pm 1\}$,*

$$\sum_{\chi\in\hat{G}\setminus\{1\}} f_\chi^{-1} \le \frac{k}{2}.$$

Proof. Notice that there are at most f Dirichlet characters of conductor f, at most half of which are even (when $f > 1$), so

$$\sum_{\chi \in \hat{G} \setminus \{1\}} f_\chi^{-1} \leq \sum_{\ell=1}^{k} \frac{p^\ell}{2} \cdot \frac{1}{p^\ell} = \frac{k}{2}.$$

□

4 Algorithmic Implications

The following is our main result about the decoding algorithm, showing that under mild restrictions on the distribution of the short generator, one can recover it from any generator that differs from it by a unit in C. Roughly speaking, the requirement from the distribution is that the *ratios* between its complex embeddings are not too large. We note that since the $\mathbf{v}_i$ below are assumed to be orthogonal to the all-1 vector, the *scale* of the distribution (or variance in the case of Gaussians) is irrelevant: this should not come as a surprise, since, e.g., one can normalize the input generator g' to have algebraic norm 1.

Theorem 4.1. *Let D be a distribution over $\mathbb{Q}(\zeta)$ with the property that for any tuple of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_{\varphi(m)/2-1} \in \mathbb{R}^{\varphi(m)/2}$ of Euclidean norm 1 that are orthogonal to the all-1 vector $\mathbf{1}$, the probability that $|\langle \mathrm{Log}(g), \mathbf{v}_i \rangle| < c\sqrt{m} \cdot (\log m)^{-3/2}$ holds for all i is at least some $\alpha > 0$, where g is chosen from D and c is a universal constant. Then there is an efficient algorithm that given $g' = g \cdot u$, where g is chosen from D and $u \in C$ is a cyclotomic unit, outputs an element of the form $\pm\zeta^j g$ with probability at least α.*

Proof. The algorithm applies the round-off algorithm from Claim 2.1 to $\mathrm{Log}(g') = \mathrm{Log}(g) + \mathrm{Log}(u)$, using the vectors $\mathbf{b}_j$ (defined and analyzed in Section 3) as the basis. By the assumption on D and Theorem 3.1, with probability at least α the output is $\mathrm{Log}(u) \in \mathrm{Log}(C)$. We next find integer coefficients a_j such that $\mathrm{Log}(u) = \sum a_j \mathbf{b}_j$, and compute $u' = \prod b_j^{a_j}$. Since $\mathrm{Log}(u') = \mathrm{Log}(u)$ it follows that u' must be of the form $\pm\zeta^j u$ for some sign and some j. Therefore, g'/u' is the desired element. □

In the next section we show that the condition on D in the theorem is satisfied by several natural distributions.

One possible concern with the above algorithm is that it expects as input $g \cdot u$ for a *cyclotomic unit* $u \in C$, whereas the first phase of the attack described in the introduction, i.e., a PIP algorithm, is only guaranteed to output $g \cdot u$ for an *arbitrary unit* $u \in R^*$. There are several reasons why this should not be an issue. First, as mentioned in Section 2, in some cases, e.g., for power-of-2 cyclotomic, it is conjectured that $C = R^*$. More generally, the index of C in R^*, which we recall is h^+, the class number of the totally real subfield, is often small. In such a case, if we have a list of coset representatives of C in R^*, we can enumerate over all of them and use the algorithm above to recover g, increasing the running time only by a factor of h^+. In order to obtain such a list of representatives, we can use an algorithm for computing the unit group, either classical [BF14] or quantum [EHKS14]. These algorithms are no slower than the known PIP algorithms and moreover, need only be applied once for a given cyclotomic field (as opposed to once for each public key). Alternatively, by running the PIP algorithm multiple times on a basis of a principal ideal with a *known* short generator chosen using the secret key generation algorithm, we can recover a list of representatives for all the cosets that show up as output of the PIP algorithm with non-negligible probability; we can then enumerate over that list.

In the above statement and proof we glossed over issues of precision and assumed for simplicity, as one often does, that the input g' is given exactly. To be fully rigorous, one needs to verify that the algorithm

can deal with inputs that are specified with finite precision, and still runs in time polynomial in its input size. Typically, by finite precision one means that the input is given in fixed-point representation, providing additive approximation to the true numbers. Here, however, it is more natural to assume that the input is given in (the strictly more general) floating-point representation, providing multiplicative approximation to the true numbers. Not only is this more natural, but also the known PIP algorithms [BF14, Bia14, BS15] generate an output in this format, or an output that can be easily converted to this format.[6] Luckily, dealing with floating-point inputs is straightforward. First notice that $\mathrm{Log}(g')$ can be written in standard fixed-point representation, and so can $\mathrm{Log}(u)$. The integer coefficients a_j can be stored exactly since they are at most exponential in the input size. Finally, by using a sufficiently good multiplicative approximation of b_j (with the multiplicative error being much less than $1/a_j$), we can obtain an arbitrarily good multiplicative approximation of u'. As a result we get a multiplicative approximation of the desired output g'/u' that can be made essentially as good as the multiplicative approximation of the input g'.

5 Tail Bounds

In this section we show that the condition on D in Theorem 4.1 is satisfied by two natural distributions: the continuous Gaussian and a wide enough discrete Gaussian (over any lattice). This section is independent of the other sections in this paper, and we avoid the use of notation from algebraic number theory. Instead, we identify elements of K with vectors in $\mathbb{R}^{\varphi(m)}$ by taking the real and the imaginary part of their $\varphi(m)/2$ complex embeddings, i.e., a is mapped to $(\Re(\sigma_j(a)), \Im(\sigma_j(a)))_{j\in G}$. As a result, all random variables appearing here are real. The results in this section should be easy to extend to other distributions.

We start with Lemma 5.2, a tail bound on the sum of subexponential random variables. The proof is based on a standard Bernstein argument, and follows the proof in [Ver12] apart from some minor modifications for convenience.

Definition 5.1. For $\alpha, \beta > 0$, a random variable X is (α, β)-*subexponential* if

$$\mathbb{E}[\cosh(\alpha X)] \le \beta \ ,$$

where recall that $\cosh(x) := (e^x + e^{-x})/2$.

Lemma 5.2 (Tail bound). *Let $X_1, \ldots, X_n$ be independent centered (i.e., expectation zero) (α, β)-subexponential random variables. Then, for any $\mathbf{a} = (a_1, \ldots, a_n) \in \mathbb{R}^n$ and every $t \ge 0$,*

$$\Pr\Big[\Big|\sum a_i X_i\Big| \ge t\Big] \le 2\exp\left(-\min\left(\frac{\alpha^2 t^2}{8\beta\|\mathbf{a}\|_2^2}, \frac{\alpha t}{2\|\mathbf{a}\|_\infty}\right)\right) .$$

Proof. By scaling, we can assume without loss of generality that $\alpha = 1$. Next, we use the inequality

$$e^{\delta x} - \delta x - 1 \le (e^{\delta x} - \delta x - 1) + (e^{-\delta x} + \delta x - 1) = 2(\cosh(\delta x) - 1) \le 2\delta^2(\cosh(x) - 1)$$

which holds for all $-1 \le \delta \le 1$ and all $x \in \mathbb{R}$, where the second inequality follows from the Taylor expansion. By applying this inequality to a $(1, \beta)$-subexponential centered random variable X, and taking expectations we see that for all $-1 \le \delta \le 1$,

$$\begin{aligned} \mathbb{E}[\exp(\delta X)] &\le 1 + 2\delta^2\,\mathbb{E}[\cosh(X) - 1] \\ &\le 1 + 2\delta^2(\beta - 1) \le \exp(2\delta^2\beta) \ . \end{aligned} \tag{6}$$

[6] In general number fields (in fact already in quadratic number fields), the use of floating point is *necessary*, since generators are typically doubly exponentially large and so would require exponential time to write down in fixed-point notation.

Using Markov's inequality, we can bound the upper tail probability for any $\lambda > 0$ as

$$\begin{aligned}\Pr\Big[\sum a_i X_i \geq t\Big] &= \Pr\Big[\exp\Big(\lambda \sum a_i X_i\Big) \geq \exp(\lambda t)\Big] \\ &\leq \exp(-\lambda t) \cdot \mathbb{E}\Big[\exp\Big(\lambda \sum a_i X_i\Big)\Big] \\ &= \exp(-\lambda t) \cdot \prod \mathbb{E}[\exp(\lambda a_i X_i)] \\ &\leq \exp(-\lambda t + 2\beta\lambda^2 \|\mathbf{a}\|_2^2) \ ,\end{aligned}$$

where in the second inequality we used (6) and assumed that $\lambda\|\mathbf{a}\|_\infty \leq 1$. Taking $\lambda = \min(t/(4\beta\|\mathbf{a}\|_2^2), 1/\|\mathbf{a}\|_\infty)$ this bound becomes at most

$$\exp\left(-\min\left(\frac{t^2}{8\beta\|\mathbf{a}\|_2^2}, \frac{t}{2\|\mathbf{a}\|_\infty}\right)\right) .$$

We complete the proof by applying the same argument with $-\mathbf{a}$. □

The next claim follows immediately from Definition 5.1.

Claim 5.3. *If Y is a non-negative random variable such that both $\mathbb{E}[Y]$ and $\mathbb{E}[Y^{-1}]$ are finite, then $\log Y$ is a $(1, \beta)$-subexponential random variable for some $\beta > 0$.*

The following is an immediate corollary of the tail bound. It shows that the condition in Theorem 4.1 holds with overwhelming probability for a continuous Gaussian distribution of any radius that is spherical in the embedding basis. Notice that the parameter r plays no role in the conclusion of the statement.

Lemma 5.4. *Let $X_1, \ldots, X_n, X'_1, \ldots, X'_n$ be i.i.d. $N(0, r)$ variables for some $r > 0$, and let $\hat{X}_i = (X_i^2 + X_i'^2)^{1/2}$. Then, for any vectors $\mathbf{a}^{(1)}, \ldots, \mathbf{a}^{(\ell)} \in \mathbb{R}^n$ of Euclidean norm 1 that are orthogonal to the all-1 vector, and every $t \geq C$ for some universal constant C,*

$$\Pr\Big[\exists j, \ \Big|\sum_i a_i^{(j)} \log(\hat{X}_i)\Big| \geq t\Big] \leq 2\ell \exp(-t/2) \ .$$

Proof. By union bound, it suffices to prove the lemma for the case $\ell = 1$, and we let $\mathbf{a} = \mathbf{a}^{(1)}$. Since $\sum a_i = 0$, we can assume without loss of generality that $r = 1$. Notice that $\hat{X}_i$ has a chi distribution with 2 degrees of freedom (also known as a Rayleigh distribution) whose density function is given by $xe^{-x^2/2}$ for $x > 0$ and zero otherwise. In particular, it is easy to see that both $\mathbb{E}[\hat{X}_i]$ and $\mathbb{E}[\hat{X}_i^{-1}]$ are finite (both are $\sqrt{\pi/2}$). Therefore, by Claim 5.3, $\log \hat{X}_i$ is $(1, \beta)$ subexponential for some constant $\beta > 0$. From this it follows that $\hat{X}_i = \log \hat{X}_i - \mathbb{E}[\log \hat{X}_i]$ are *centered* $(1, \beta')$ subexponential random variables for some constant $\beta' > 0$. The result now follows by applying Lemma 5.2 to $\hat{X}_1, \ldots, \hat{X}_n$, using the bound $\|\mathbf{a}\|_\infty \leq 1$, and the observation that $\sum_i a_i \mathbb{E}[\log \hat{X}_i] = 0$. □

In the next lemma we show that small perturbations of the continuous Gaussian distribution still satisfy the condition in Theorem 4.1.

Lemma 5.5. *Let $X = (X_1, \ldots, X_n, X'_1, \ldots, X'_n)$ be i.i.d. $N(0, r)$ variables for some $r > 0$, and let $Y = (Y_1, \ldots, Y_n, Y'_1, \ldots, Y'_n)$ be a (not necessarily independent) random vector satisfying $\|Y\|_2 \leq u$ with probability 1 for some $u \leq r/(20\sqrt{n})$. Let $Z = X + Y$ and define $\hat{X}_i$, $\hat{Y}_i$, $\hat{Z}_i$ as before. Then for any vectors $\mathbf{a}^{(1)}, \ldots, \mathbf{a}^{(\ell)} \in \mathbb{R}^n$ of Euclidean norm 1 that are orthogonal to the all-1 vector, it holds with constant probability that for all j,*

$$\Big|\sum_i a_i^{(j)} \log(\hat{Z}_i)\Big| \leq 1 + 10 \log \ell \ .$$

Proof. By Lemma 5.4 we have that with some constant probability close to 1,

$$\forall j,\ \left|\sum_i a_i^{(j)} \log(\hat{X}_i)\right| < 10 \log \ell\ . \tag{7}$$

Moreover, since $\hat{X}_i < r/(10\sqrt{n})$ implies that both $|X_i|$ and $|X_i'|$ are smaller than $r/(10\sqrt{n})$, we see that by independence of X_i, X_i', the probability of the former event is at most c/n for some small constant c. As a result we have that with constant probability close to 1,

$$\forall i,\ \hat{X}_i > r/(10\sqrt{n})\ .$$

In the following we assume that these two conditions hold (which happens with constant probability close to 1 by union bound), and bound the effect of Y. Now let $\mathbf{a}$ be one of the vectors in the statement of the lemma. Then,

$$\begin{aligned}\left|\sum_i a_i \log(\hat{Z}_i)\right| &\le \left|\sum_i a_i \log(\hat{X}_i)\right| + \left|\sum_i a_i \log(\hat{Z}_i/\hat{X}_i)\right| \\ &\le 10 \log \ell + \left|\sum_i a_i \log(\hat{Z}_i/\hat{X}_i)\right|\ ,\end{aligned}$$

where we used Eq. (7). Notice that by the triangle inequality (for two-dimensional Euclidean space),

$$\hat{X}_i - \hat{Y}_i \le \hat{Z}_i \le \hat{X}_i + \hat{Y}_i\ .$$

Since $\hat{Y}_i \le \|Y\|_2 \le u \le r/(20\sqrt{n}) \le \hat{X}_i/2$, and using the inequality $|\log(1+\delta)| \le 2|\delta|$ valid for all $\delta \in [-1/2, 1/2]$,

$$\begin{aligned}\left|\sum_i a_i \log(\hat{Z}_i/\hat{X}_i)\right| &\le \Big(\sum_i (\log(\hat{Z}_i/\hat{X}_i))^2\Big)^{1/2} \\ &\le \Big(\sum_i (2\hat{Y}_i/\hat{X}_i)^2\Big)^{1/2} \\ &\le 20\sqrt{n}/r \cdot \Big(\sum_i \hat{Y}_i^2\Big)^{1/2} \\ &\le 20\sqrt{n}u/r \le 1\ ,\end{aligned}$$

where the first inequality follows from Cauchy-Schwarz. □

Finally, we consider the spherical (in the embedding basis) discrete Gaussian distribution over an arbitrary lattice $L \subseteq \mathbb{R}^{2n}$. Such distributions show up often in cryptographic constructions (see, e.g., [LPR13]), and often that lattice is the (embedding of the) ring of integers R. For background on the discrete Gaussian distribution and the smoothing parameter, see, e.g., [MR04]. In order to apply Lemma 5.5 to this distribution, take X to be the continuous Gaussian D_r for some $r \ge 100n\eta_\varepsilon(L)$, and Y the discrete Gaussian $D_{L-X,s}$ over the coset $L - X$ of parameter $s = \eta_\varepsilon(L)$ for some negligible parameter ε. Using Banaszczyk's result [Ban93] we have that with all but exponentially small probability in n, $\|Y\|_2 \le \sqrt{2n}\eta_\varepsilon(L) \le r/(60\sqrt{n})$. Moreover, by the lemma below, the distribution of $Z = X + Y$ is within negligible statistical distance of the discrete Gaussian distribution $D_{L,r'}$ for $r' = (r^2 + \eta_\varepsilon(L)^2)^{1/2}$. We therefore see that the condition in Theorem 4.1 holds for the discrete Gaussian distribution $D_{L,r'}$ for any lattice L and any $r' > 200n\eta_\varepsilon(L)$.

Lemma 5.6 (Special case of [Pei10, Theorem 3.1]). *Let L be a lattice and $r, s > 0$ be such that $s \ge \eta_\varepsilon(L)$ for some $\varepsilon \le 1/2$. Then if we choose $\mathbf{x}$ from the continuous Gaussian D_r and then choose $\mathbf{y}$ from the discrete Gaussian $D_{L-\mathbf{x},s}$ then $\mathbf{x} + \mathbf{y}$ is within statistical distance 8ε of the discrete Gaussian $D_{L,(r^2+s^2)^{1/2}}$.*

6 Shortest Generators of Principal Ideals and an SVP Algorithm

In a principal ideal $\mathcal{I}$, how long (in the Euclidean norm) can the shortest generator be, relative to its algebraic norm? In this section we provide lower and upper bounds showing that for a cyclotomic ring R of prime-power index m, the answer is $\exp(\tilde{\Theta}(\sqrt{m})) \cdot S(\mathcal{I})$, where $S(\mathcal{I}) = \mathrm{N}(\mathcal{I})^{1/\varphi(m)}$ is the dimension-normalized algebraic norm of $\mathcal{I}$, and $\tilde{\Theta}$ hides polylogarithmic factors. (To be precise, the lower bound is under the mild conjecture that $h^+(m) = 2^{O(m)}$; see the end of Section 2.4.) By contrast, it is well known (see, e.g., [PR07, Lemmas 6.1 and 6.2]) that the minimum distance (i.e., the length of a shortest nonzero vector) of any ideal is bounded by $\Omega(\sqrt{m}) \cdot S(\mathcal{I})$ and $O(m) \cdot S(\mathcal{I})$, by the arithmetic-mean/geometric-mean inequality and Minkowski's theorem, respectively. Therefore, any algorithm that always outputs a generator when given a principal ideal (e.g., the algorithm analyzed in the previous sections) obtains no better than a $\exp(\tilde{\Omega}(\sqrt{m}))$ approximation factor for the Shortest Vector Problem, in the worst case.

We first show in Section 6.1 that upper and lower bounds on shortest generators follow directly from an analysis of the *covering radius* of the log-unit lattice Λ (and its sublattice $\mathrm{Log}\, C$), in the ℓ_∞ and ℓ_1 norms (respectively). Sections 6.2 and 6.3 then prove upper and lower bounds on these covering radii. In fact, the proofs demonstrate more: the lower bound holds for "almost all" principal ideals, and the upper bound is algorithmic in the following sense: given an arbitrary generator (which can be found using the quantum PIP algorithm of [BS15, BS16]), we can efficiently find a generator satisfying the bound, which in particular is a $\exp(\tilde{O}(\sqrt{m}))$-approximate shortest vector in the ideal.

Throughout this section we let $m > 2$ be a prime power, and let $n := |G| = \varphi(m)/2 = \Theta(m)$. Let H be the subspace of $\mathbb{R}^n$ spanned by $\Lambda = \mathrm{Log}\, R^*$ (and by $\mathrm{Log}\, C$, the log embedding of the cyclotomic units), which is the subspace orthogonal to $\mathbf{1}$, the all-1s vector. Define the covering radius of a lattice $\mathcal{L}$ with respect to the ℓ_p norm as

$$\mu^{(p)}(\mathcal{L}) = \max_{\mathbf{x} \in \mathrm{span}(\mathcal{L})} \min_{\mathbf{v} \in \mathcal{L}} \|\mathbf{x} - \mathbf{v}\|_p = \max_{\mathbf{x} \in \mathrm{span}(\mathcal{L})} \min_{\mathbf{v} \in \mathbf{x} + \mathcal{L}} \|\mathbf{v}\|_p \,.$$

6.1 Relation to Covering Radius

For any $g \in R$, let $\mathcal{I} = gR$. Also let $\mathbf{g} = \mathrm{Log}(g)$ and write it as $\mathbf{g} = s\mathbf{1} + \mathbf{g}_H$ where $\mathbf{g}_H \in H$. Observe that $s = \log S(\mathcal{I})$, because

$$\mathrm{N}(\mathcal{I}) = \mathrm{N}(g) = \prod_{i \in \mathbb{Z}_m^*} \sigma_i(g) = \prod_{i \in G} |\sigma_i(g)|^2 = \exp(2\langle \mathbf{g}, \mathbf{1} \rangle) = \exp(s \cdot \varphi(m)).$$

Lemma 6.1. *Let g, $\mathcal{I}$, s, and $\mathbf{g}_H$ be as above. There exists an efficient algorithm that, given g and any $\mathbf{h}_H \in \mathbf{g}_H + \mathrm{Log}\, C$, outputs a generator h of $\mathcal{I}$ such that*

$$\|h\| \le \sqrt{\varphi(m)} \cdot \exp(\|\mathbf{h}_H\|_\infty) \cdot S(\mathcal{I}).$$

In particular, there exists a generator of Euclidean norm at most $\sqrt{\varphi(m)} \cdot \exp(\mu^{(\infty)}(\mathrm{Log}\, C)) \cdot S(\mathcal{I})$.

Proof. As in the proof of Theorem 4.1, for simplicity we ignore issues of precision; see the discussion at the end of Section 4. The algorithm lets $\mathbf{u} = \mathbf{h}_H - \mathbf{g}_H \in \mathrm{Log}\, C$, computes the coefficients $a_j \in \mathbb{Z}$ such that $\mathbf{u} = \sum a_j \mathbf{b}_j$, and outputs $h = g \cdot \prod b_j^{a_j}$. Because $\mathbf{h} := \mathrm{Log}(h) = \mathrm{Log}(g) + \mathbf{u} = s\mathbf{1} + \mathbf{h}_H$, we have

$$\|h\|^2 = \sum_{i \in \mathbb{Z}_m^*} |\sigma_i(h)|^2 \le \varphi(m) \cdot \exp(\|\mathbf{h}\|_\infty)^2 = \varphi(m) \cdot \exp(\|\mathbf{h}_H\|_\infty)^2 \cdot S(\mathcal{I})^2.$$

□

Lemma 6.2. *There exists a principal ideal $\mathcal{I} \subseteq R$ for which every generator has Euclidean norm at least $\exp(\Omega(\mu^{(1)}(\Lambda)/m)) \cdot S(\mathcal{I})$.*

In fact, the proof shows that a "random principal ideal," i.e., one whose generators correspond to a uniformly random coset of the log-unit lattice, satisfies the above bound with overwhelming probability. (Formalizing this requires a bit more effort; we omit the details.)

Proof. Let $\mathbf{x} + \Lambda \subset H$ be a "deep hole" coset of Λ in the ℓ_1 norm, i.e., one for which $\|\mathbf{v}\|_1 \geq \mu^{(1)}(\Lambda)$ for every $\mathbf{v} \in \mathbf{x} + \Lambda \subset H$. Because the n coordinates of any such $\mathbf{v}$ sum to zero, the sum of the positive coordinates must be exactly $\|\mathbf{v}\|_1/2$, and therefore there must be a coordinate that is at least $\mu^{(1)}(\Lambda)/(2n) = \Omega(\mu^{(1)}(\Lambda)/m)$.

Next, assume for a moment that there exists $g \in R$ for which $\mathbf{g}_H = \mathbf{x}$, where as before we write $\mathbf{g} = \mathrm{Log}(g) = s\mathbf{1} + \mathbf{g}_H$. Then any generator h of the ideal $\mathcal{I} = gR$ satisfies $\mathrm{Log}(h) \in \mathrm{Log}(g) + \Lambda = s\mathbf{1} + \mathbf{x} + \Lambda$, so by the observation above, it must have the claimed Euclidean norm.

To complete the proof, notice that even if there does not exist a g as above, one can find g so as to make $\mathbf{g}_H$ arbitrarily close to $\mathbf{x}$, which suffices for the above analysis. To see this, consider $x = M \cdot \mathrm{Exp}(\mathbf{x})$, where M is a sufficiently large integer and $\mathrm{Exp}(\mathbf{x}) \in \mathrm{Log}^{-1}(\mathbf{x})$ denotes an arbitrary preimage in $K_{\mathbb{R}} := K \otimes_{\mathbb{Q}} \mathbb{R}$ of $\mathbf{x}$ under the log embedding (extended to $K_{\mathbb{R}}$). Then rounding x to a nearest $g \in R$ yields the claim. □

6.2 Covering Radius Upper Bound and an SVP Algorithm

Theorem 6.3. *There is an efficient randomized algorithm that given any vector $\mathbf{x} \in H$ outputs a vector $\mathbf{v} \in \mathrm{Log}\, C$ such that $\|\mathbf{x} - \mathbf{v}\|_\infty = O(\sqrt{m \log m})$ with high probability.*

Before giving the proof, we mention some implications of the theorem. First, using the fact that $\mathrm{Log}\, C$ is a sublattice of Λ, we immediately get the following corollary regarding the covering radii of these lattices.

Corollary 6.4. *For a prime power m, we have $\mu^{(\infty)}(\Lambda) \leq \mu^{(\infty)}(\mathrm{Log}\, C) \leq O(\sqrt{m \log m})$.*

We remark that this corollary can also be obtained directly from Lemma 6.7 below and the non-trivial result of Banaszczyk and Szarek [BS97] (see also [Ban98]). We also note that if the Komlós conjecture is true, then the $\sqrt{\log m}$ factor in the corollary can be removed.

It follows immediately from the corollary and Lemma 6.1 that any principal ideal $\mathcal{I}$ has a generator whose Euclidean norm is at most $\exp(O(\sqrt{m \log m})) \cdot S(\mathcal{I})$. This also leads to an efficient quantum algorithm providing a non-trivial approximation to SVP in principal ideals, as described in the following theorem.

Theorem 6.5. *There is an efficient quantum algorithm that approximates SVP on principal ideal lattices in cyclotomics of prime-power index m to within approximation factor $2^{O(\sqrt{m \log m})}$.*

Proof. Given a principal ideal $\mathcal{I}$, first use the efficient quantum algorithm of Biasse and Song [BS15] to recover a generator g of $\mathcal{I}$, and as above, write $\mathrm{Log}(g) = s\mathbf{1} + \mathbf{g}_H$ for $\mathbf{g}_H \in H$. Next, apply Theorem 6.3 to $\mathbf{g}_H$ and let $\mathbf{v} \in \mathrm{Log}\, C$ be the output. Finally, apply the algorithm from Lemma 6.1 with g and $\mathbf{h}_H := \mathbf{g}_H - \mathbf{v}$ to find a generator h whose Euclidean norm is at most $\exp(O(\sqrt{m \log m})) \cdot S(\mathcal{I})$, and output h. It is sufficiently short since, as mentioned at the start of the section, $\lambda_1(\mathcal{I}) = \Omega(\sqrt{m}) \cdot S(\mathcal{I})$ by the arithmetic mean-geometric mean inequality. □

For the proof of Theorem 6.3, we need a simple probabilistic lemma, as well as a bound on the norm of the $\mathbf{b}_j$. For $\alpha \in [0, 1]$, define $S(\alpha)$ as the unique probability distribution on support $\{\alpha, \alpha - 1\}$ with expectation 0 (i.e., it assigns probability $1 - \alpha$ to α and probability α to $\alpha - 1$).

Lemma 6.6. *Let $\mathbf{A}$ be an $n \times n$ matrix all of whose rows have Euclidean norm at most $T > 0$, and let $\alpha_1, \ldots, \alpha_n \in [0,1]$ be arbitrary. Let $x_1, \ldots, x_n$ be independent with x_i distributed as $S(\alpha_i)$, and let $\mathbf{x} = (x_1, \ldots, x_n)$. Then with probability $\Omega(1/\sqrt{n})$, both*

$$\|\mathbf{A}\mathbf{x}\|_\infty \leq O(T\sqrt{\log n}) \qquad \textit{and} \qquad \left|\sum x_i\right| \leq O(1) \; .$$

Proof. Since $S(\alpha)$ is bounded, it is a subgaussian random variable of constant subgaussian norm. (See [Ver12, Section 5.2.3] for the definition and properties of subgaussian random variables.) Because the sum of independent subgaussian random variables is also subgaussian (see [Ver12, Lemma 5.9]), $(\mathbf{A}\mathbf{x})_i$ has subgaussian norm $O(T)$ for every $i = 1, \ldots, n$. Therefore, for a large enough universal constant $C > 0$,

$$\Pr\left[|(\mathbf{A}\mathbf{x})_i| > CT\sqrt{\log n}\right] = O(1/n^2) \; ,$$

and by a union bound we get

$$\Pr\left[\|\mathbf{A}\mathbf{x}\|_\infty > CT\sqrt{\log n}\right] = O(1/n) \; . \tag{8}$$

Next, by the Berry-Esseen theorem (see, e.g., [O'D14, Section 5.2]), since the x_i have expectation 0 and bounded second and third moments, the probability that $|\sum x_i| = O(1)$ is $\Omega(1/\sqrt{n})$. Together with Eq. (8) and the union bound, this completes the proof. □

Lemma 6.7. *Let m be a prime power. Then for all $j \in G$, $\|\mathbf{z}_j\| = O(\sqrt{m})$, where $\mathbf{z}_j$ are the vectors defined in Eq.* (3)*.*

Proof. Notice that

$$\begin{aligned}
\|\mathbf{z}_j\|^2 &= \sum_{i \in G} \log^2|\omega^{ij} - 1| = \sum_{i \in G} \log^2|\omega^i - 1| \\
&= \sum_{i \in G} \log^2|2\sin(\pi i/m)| \leq \sum_{i=1}^{\lfloor m/2 \rfloor} \log^2(2\sin(\pi i/m)) \\
&= \sum_{i=1}^{\lfloor m/2 \rfloor} f(i/m) \; ,
\end{aligned} \tag{9}$$

where $f : [0, 1/2] \to \mathbb{R}$ is given by $f(x) = \log^2(2\sin(\pi x))$. Since $f(x) \leq \log 2$ for $1/6 \leq x \leq 1/2$ (recall that $\sin(\pi/6) = 1/2$), the contribution to the sum in Eq. (9) coming from $i > \lfloor m/6 \rfloor$ is at most $O(m)$. It therefore suffices to consider the contribution coming from $i \in \{1, \ldots, \lfloor m/6 \rfloor\}$. Since $\sin(\pi x) \geq 2x$ for $0 \leq x \leq 1/2$ (as follows from the concavity of sine on $[0, \pi/2]$), that contribution satisfies

$$\sum_{i=1}^{\lfloor m/6 \rfloor} f(i/m) \leq \sum_{i=1}^{\lfloor m/6 \rfloor} \log^2(4i/m) \leq m \int_0^{1/6} \log^2(4x)dx = O(m) \; ,$$

the last equality following from

$$\int_0^y \log^2(x)dx = y(\log^2 y - 2\log y + 2).$$

□

Proof of Theorem 6.3. Given any $\mathbf{y} \in H$, find real coefficients $(a_j)_{j \in G \setminus \{1\}}$ such that $\mathbf{y} = \sum a_j \mathbf{b}_j$. For $j \in G \setminus \{1\}$, let $\alpha_j = (a_j \bmod 1) \in [0,1)$ be the fractional part of a_j, and let x_j be independent random variables distributed like $S(\alpha_j)$. The algorithm outputs $\mathbf{u} = \sum (a_j - x_j)\mathbf{b}_j$. Notice that $\mathbf{u} \in \operatorname{Log} C$ as desired. To analyze the distance of $\mathbf{u}$ from $\mathbf{y}$, for convenience let x_1 be an independent random variable distributed like $S(0)$ (so $x_1 = 0$ always). Recalling that $\mathbf{b}_j = \mathbf{z}_j - \mathbf{z}_1$, write

$$\mathbf{y} - \mathbf{u} = \sum_{j \in G} x_j(\mathbf{z}_j - \mathbf{z}_1) = \sum_{j \in G} x_j \mathbf{z}_j - \left(\sum_{j \in G} x_j\right)\mathbf{z}_1 \ ,$$

and so by the triangle inequality

$$\begin{aligned} \|\mathbf{y} - \mathbf{u}\|_\infty &\le \left\|\sum_{j \in G} x_j \mathbf{z}_j\right\|_\infty + \left|\sum_{j \in G} x_j\right| \cdot \|\mathbf{z}_1\|_\infty \\ &\le \left\|\sum_{j \in G} x_j \mathbf{z}_j\right\|_\infty + \left|\sum_{j \in G} x_j\right| \cdot O(\sqrt{m}) \ , \end{aligned}$$

where we used the trivial bound $\|\mathbf{z}_1\|_\infty \le \|\mathbf{z}_1\|_2$ and applied Lemma 6.7.[7] We now apply Lemma 6.6 to the matrix $\mathbf{Z}$ whose columns are the $\mathbf{z}_j$. Since $\mathbf{Z}$ is G-circulant, the Euclidean norms of all its rows and columns are the same, and by Lemma 6.7 are $O(\sqrt{m})$. We therefore obtain that with probability $\Omega(1/\sqrt{n})$,

$$\|\mathbf{y} - \mathbf{u}\|_\infty \le O(\sqrt{m \log n}) + O(\sqrt{m}) = O(\sqrt{m \log m}) \ ,$$

as desired. The success probability can be amplified by repetition. □

6.3 Covering Radius Lower Bound

Let $h' := (h^+)^{1/(n-1)}$, which we recall is conjectured to be constant. Combined with Lemma 6.2, the theorem below shows that there exists a principal ideal $\mathcal{I} \subseteq R$ for which every generator has Euclidean norm at least $\exp(\Omega(\sqrt{m}/(h' \log m))) \cdot S(\mathcal{I})$.

Theorem 6.8. *For a prime power m, the log-unit lattice satisfies*

$$\mu^{(1)}(\Lambda) \ge \Omega\big(m^{3/2}/(h' \log m)\big) \ .$$

Proof. Using Lemma 6.9 below,

$$(\det(\operatorname{Log} C))^{1/(n-1)} = \Omega(m^{1/2}/\log m) \ .$$

Since $\det(\Lambda) = \det(\operatorname{Log} C)/h^+$,[8]

$$(\det(\Lambda))^{1/(n-1)} = \Omega(m^{1/2}/(h' \log m)) \ .$$

The theorem now follows from the fact that

$$\operatorname{vol}(B_1^n \cap H) \le \sqrt{n} \cdot 2^{n-1}/(n-1)! = O(1/n)^{n-1} \ ,$$

where $B_1^n := \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_1 \le 1\}$. To prove this inequality, notice that (1) the volume of B_1^{n-1} is $2^{n-1}/(n-1)!$, (2) the projection of $B_1^n \cap H$ on the first $n-1$ coordinates is contained in B_1^{n-1}, and (3) this projection shrinks volumes by $\sqrt{n}$, as can be seen by computing its Jacobian. □

[7] In fact, $\|\mathbf{z}_1\|_\infty = O(\log m)$, but we do not need this.

[8] We note that $2^{n-1} \det(\Lambda)/\sqrt{n}$ is known as the *regulator*, and a bound similar to what we obtain here can be derived from the Brauer-Siegel theorem [Was97, Page 43]. This leads to a bound that is both somewhat weaker and ineffective.

Lemma 6.9. *The determinant of* $\operatorname{Log} C$ *satisfies*

$$\det(\operatorname{Log} C)^{1/(n-1)} = \Omega(\sqrt{m}/\log m).$$

Proof. Recall from the proof of Lemma 3.2 that $\mathbf{b}_j^\vee$ is the projection of $\mathbf{z}_j^\vee$ orthogonal to $\mathbf{1}$. The $|G|$-dimensional full-rank lattice generated by $\{\mathbf{z}_j^\vee\}_{j\in G}$ has determinant

$$|\det(\mathbf{Z}^{-t})| = \prod_{\chi\in\widehat{G}} |\lambda_\chi^{-1}| \,.$$

Next, notice that the shortest vector in the intersection of this lattice with the span of $\mathbf{1}$ is $\mathbf{Z}^{-t}\mathbf{1} = \lambda_1^{-1}\mathbf{1}$, whose Euclidean norm is $\lambda_1^{-1}\sqrt{|G|}$. Therefore, the dual of $\operatorname{Log} C$, which is the projection of this lattice orthogonally to $\mathbf{1}$, has determinant

$$|G|^{-1/2} \prod_{\chi\in\widehat{G}\setminus\{1\}} |\lambda_\chi^{-1}| \,,$$

and therefore

$$\begin{aligned}\det(\operatorname{Log} C) &= |G|^{1/2} \prod_{\chi\in\widehat{G}\setminus\{1\}} |\lambda_\chi| \\ &= |G|^{1/2} \prod_{\chi\in\widehat{G}\setminus\{1\}} \left|\frac{1}{2}\sqrt{f_\chi}\cdot L(1,\chi)\right| , \end{aligned} \tag{10}$$

where we used Eq. (5). Letting $m = p^k$ for a prime p, and using Theorem 2.6, we get that

$$L := \prod_{\chi\in\widehat{G}\setminus\{1\}} |L(1,\chi)| = \Omega((\log m)^{-(n-1-q)}\cdot p^{-q/2})$$

where q denotes the number of even quadratic characters modulo m, which is at most 3 (see [MV06, Section 9.3]). We conclude that

$$L^{1/(n-1)} = \Omega(1/\log m). \tag{11}$$

Next, consider $F = \prod_{\chi\in\hat{G}\setminus\{1\}} f_\chi$. For each $0 < j \le k$, there are exactly $\varphi(p^j) - \varphi(p^{j-1})$ characters of conductor $f_\chi = p^j$. Exactly half are these are even when p is odd and $j > 1$, and also when $p = 2$ and $j > 2$. When p is odd and $j = 1$ there are $\varphi(p)/2 - 1$ even characters of conductor p, and when $p = 2$ there are no even characters of conductor 2 or 4. Assuming p is odd (the case $p = 2$ being very similar), this leads to

$$\begin{aligned}\log_p F &= \sum_{j=1}^{k} j\cdot\frac{\varphi(p^j)-\varphi(p^{j-1})}{2} - \frac{1}{2} \\ &= \frac{k}{2}\cdot\varphi(p^k) - \frac{1}{2}\sum_{j=0}^{k-1}\varphi(p^j) - \frac{1}{2} \\ &= kn - \frac{p-1}{2}\sum_{j=0}^{k-2} p^j - 1 = kn - \frac{p^{k-1}}{2} - \frac{1}{2} \,, \end{aligned}$$

and we conclude that

$$F = m^{n\left(1-\frac{1}{2k(p-1)}-\frac{1}{2kn}\right)} = \Omega(m)^n \,. \tag{12}$$

Plugging (11) and (12) into (10) completes the proof. □

References

[ADS15] D. Aggarwal, D. Dadush, and N. Stephens-Davidowitz. Solving the closest vector problem in 2^n time - the discrete Gaussian strikes again! In *FOCS*, pages 563–582. 2015.

[Bab85] L. Babai. On Lovász' lattice reduction and the nearest lattice point problem. *Combinatorica*, 6(1):1–13, 1986. Preliminary version in STACS 1985.

[Ban93] W. Banaszczyk. New bounds in some transference theorems in the geometry of numbers. *Mathematische Annalen*, 296(4):625–635, 1993.

[Ban98] W. Banaszczyk. Balancing vectors and Gaussian measures of n-dimensional convex bodies. *Random Structures Algorithms*, 12(4):351–360, 1998.

[Ber14a] D. Bernstein, June 2014. Personal communication.

[Ber14b] D. Bernstein. A subfield-logarithm attack against ideal lattices. `http://blog.cr.yp.to/20140213-ideal.html`, Febuary 2014.

[BF14] J.-F. Biasse and C. Fieker. Subexponential class group and unit group computation in large degree number fields. *LMS J. Comput. Math.*, 17(suppl. A):385–403, 2014.

[Bia14] J.-F. Biasse. Subexponential time relations in the class group of large degree number fields. *Adv. Math. Commun.*, 8(4):407–425, 2014.

[BPR04] J. Buhler, C. Pomerance, and L. Robertson. Heuristics for class numbers of prime-power real cyclotomic fields. *Fields Inst. Commun*, 41:149–157, 2004.

[BS97] W. Banaszczyk and S. J. Szarek. Lattice coverings and Gaussian measures of n-dimensional convex bodies. *Discrete Comput. Geom.*, 17(3):283–286, 1997.

[BS15] J.-F. Biasse and F. Song. A note on the quantum attacks against schemes relying on the hardness of finding a short generator of an ideal in $\mathbb{Q}(\zeta_{2^n})$. Technical Report 2015-12, The University of Waterloo, 2015. Revision of September 28th 2015.

[BS16] J.-F. Biasse and F. Song. A polynomial time quantum algorithm for computing class groups and solving the principal ideal problem in arbitrary degree number fields. In *SODA*. 2016.

[CGS14] P. Campbell, M. Groves, and D. Shepherd. Soliloquy: A cautionary tale. ETSI 2nd Quantum-Safe Crypto Workshop, 2014. Available at `http://docbox.etsi.org/Workshop/2014/201410_CRYPTO/S07_Systems_and_Attacks/S07_Groves_Annex.pdf`.

[EHKS14] K. Eisenträger, S. Hallgren, A. Kitaev, and F. Song. A quantum algorithm for computing the unit group of an arbitrary degree number field. In *STOC*, pages 293–302. ACM, 2014.

[Gen09] C. Gentry. Fully homomorphic encryption using ideal lattices. In *STOC*, pages 169–178. 2009.

[GGH13] S. Garg, C. Gentry, and S. Halevi. Candidate multilinear maps from ideal lattices. In *EUROCRYPT*, pages 1–17. 2013.

[GS02] C. Gentry and M. Szydlo. Cryptanalysis of the revised NTRU signature scheme. In *EUROCRYPT*, pages 299–320. 2002.

[HPS98] J. Hoffstein, J. Pipher, and J. H. Silverman. NTRU: A ring-based public key cryptosystem. In *ANTS*, pages 267–288. 1998.

[Lan27] E. Landau. Über Dirichletsche Reihen mit komplexen Charakteren. *Journal für die reine und angewandte Mathematik*, 157:26–32, 1927.

[Lan02] S. Lang. *Algebra*, volume 211 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, third edition, 2002.

[Len82] A. K. Lenstra. Lattices and factorization of polynomials over algebraic number fields. In *Computer Algebra*, pages 32–39. Springer, 1982.

[LLL82] A. K. Lenstra, H. W. Lenstra, Jr., and L. Lovász. Factoring polynomials with rational coefficients. *Mathematische Annalen*, 261(4):515–534, December 1982.

[LLS15] Y. Lamzouri, X. Li, and K. Soundararajan. Conditional bounds for the least quadratic non-residue and related problems. *Math. Comp.*, 84(295):2391–2412, 2015.

[LMPR08] V. Lyubashevsky, D. Micciancio, C. Peikert, and A. Rosen. SWIFFT: A modest proposal for FFT hashing. In *FSE*, pages 54–72. 2008.

[Lou15] S. Louboutin. An explicit lower bound on moduli of Dirichlet L-functions at $s = 1$. *J. Ramanujan Math. Soc.*, 30(1):101–113, 2015.

[LPR10] V. Lyubashevsky, C. Peikert, and O. Regev. On ideal lattices and learning with errors over rings. *Journal of the ACM*, 60(6):43:1–43:35, November 2013. Preliminary version in Eurocrypt 2010.

[LPR13] V. Lyubashevsky, C. Peikert, and O. Regev. A toolkit for ring-LWE cryptography. In *EUROCRYPT*, pages 35–54. 2013.

[LSS14] A. Langlois, D. Stehlé, and R. Steinfeld. GGHLite: More efficient multilinear maps from ideal lattices. In *Advances in Cryptology–EUROCRYPT 2014*, pages 239–256. Springer, 2014.

[Mic02] D. Micciancio. Generalized compact knapsacks, cyclic lattices, and efficient one-way functions. *Computational Complexity*, 16(4):365–411, 2007. Preliminary version in FOCS 2002.

[Mil14] J. C. Miller. Class numbers of totally real fields and applications to the Weber class number problem. *Acta Arith.*, 164(4):381–398, 2014.

[Mil15] J. C. Miller. Real cyclotomic fields of prime conductor and their class numbers. *Math. Comp.*, 84(295):2459–2469, 2015.

[MR04] D. Micciancio and O. Regev. Worst-case to average-case reductions based on Gaussian measures. *SIAM J. Comput.*, 37(1):267–302, 2007. Preliminary version in FOCS 2004.

[MV06] H. L. Montgomery and R. C. Vaughan. *Multiplicative Number Theory I*. Cambridge University Press, 2006.

[MV10] D. Micciancio and P. Voulgaris. A deterministic single exponential time algorithm for most lattice problems based on Voronoi cell computations. In *STOC*, pages 351–358. 2010.

[O'D14] R. O'Donnell. *Analysis of Boolean Functions*. Cambridge University Press, 2014.

[Pei10] C. Peikert. An efficient and parallel Gaussian sampler for lattices. In *CRYPTO*, pages 80–97. 2010.

[PR07] C. Peikert and A. Rosen. Lattices that admit logarithmic worst-case to average-case connection factors. In *STOC*, pages 478–487. 2007.

[Sam70] P. Samuel. *Algebraic Theory of Numbers*. Hermann, Paris, 1970.

[Sch87] C.-P. Schnorr. A hierarchy of polynomial time lattice basis reduction algorithms. *Theor. Comput. Sci.*, 53:201–224, 1987.

[Sch03] R. Schoof. Class numbers of real cyclotomic fields of prime conductor. *Mathematics of computation*, 72(242):913–937, 2003.

[Sch15] J. Schank. LOGCVP, Pari implementation of CVP in $\mathrm{Log}\,\mathbb{Z}[\zeta_{2^n}]^*$. `https://github.com/jschanck-si/logcvp`, March 2015.

[She14] D. Shepherd, December 2014. Personal communication.

[SV10] N. P. Smart and F. Vercauteren. Fully homomorphic encryption with relatively small key and ciphertext sizes. In *Public Key Cryptography*, pages 420–443. 2010.

[Ver12] R. Vershynin. Introduction to the non-asymptotic analysis of random matrices. In *Compressed sensing*, pages 210–268. Cambridge Univ. Press, Cambridge, 2012. Available at `http://www-personal.umich.edu/~romanv/papers/non-asymptotic-rmt-plain.pdf`.

[Was97] L. Washington. *Introduction to Cyclotomic Fields*. Graduate Texts in Mathematics. Springer New York, 1997.

A Proof of Theorem 2.8

Proof. First, Corollary 4.13 of [Was97] gives that $\mathbb{Z}[\zeta]^*$ is generated by $\mathbb{Z}[\zeta+\overline{\zeta}]^*$ and ζ, so it follows that

$$\Lambda = \mathrm{Log}\,\mathbb{Z}[\zeta]^* = \mathrm{Log}\,\mathbb{Z}[\zeta+\overline{\zeta}]^*,$$

since the kernel of Log is the group $\{\pm 1\}\cdot U$.

Next, recall that the group of *cyclotomic units* is defined as $C = A \cap R^*$. We define the group of *real* cyclotomic units as $C^+ = A \cap \mathbb{Z}[\zeta+\overline{\zeta}]^*$. The analogue of Lemma 2.7 for the real cyclotomic units, also included in Lemma 8.1 of [Was97], says that the group C^+ of real cyclotomic units is generated by -1 and $\zeta^{(1-j)/2}\cdot b_j$. So as above, we obtain that

$$\mathrm{Log}\,C = \mathrm{Log}\,C^+ .$$

The theorem then follows from the sequence of equalities

$$[\Lambda : \mathrm{Log}\,C] = \left[\mathrm{Log}\,\mathbb{Z}\left[\zeta+\overline{\zeta}\right]^* : \mathrm{Log}\,C^+\right] = \left[\mathbb{Z}\left[\zeta+\overline{\zeta}\right]^* : C^+\right] = h^+ ,$$

where the second equality follows from $\ker(\mathrm{Log}) \cap C^+ = \ker(\mathrm{Log}) \cap \mathbb{Z}[\zeta+\overline{\zeta}]^*$ $(= \{\pm 1\})$, and the third equality is Theorem 8.2 of [Was97]. □

B Numeric Data

The previous sections established *asymptotic* bounds related to the log-embeddings of the cyclotomic units. Figure 1 gives concrete numeric data for several practical (and even impractical) choices of cyclotomic fields. This data confirms that the method works in practice.

k ($m = 2^k$)	6	7	8	9	≥ 10
$\|\mathbf{b}_j^\vee\|^{-2}$	5.04	8.56	14.69	25.71	≥ 45.85

k ($m = 3^k$)	4	5	≥ 6
$\|\mathbf{b}_j^\vee\|^{-2}$	5.72	13.65	≥ 34.04

k ($m = 5^k$)	3	4	≥ 5
$\|\mathbf{b}_j^\vee\|^{-2}$	10.04	36.43	≥ 143

Figure 1: Lower bounds on the inverse lengths of the dual vectors $\mathbf{b}_j^\vee$ defined in Section 3, for various cyclotomics of prime-power index. Larger values correspond to larger decoding distances for the log-embedding of the cyclotomic units.