Universiteit
Leiden
The Netherlands

Random walks on Arakelov class groups

Boer, K. de

Citation
Boer, K. de. (2022, September 22). *Random walks on Arakelov class groups*. Retrieved from https://hdl.handle.net/1887/3463719

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/3463719

Note: To cite this publication please use the final published version (if applicable).

5. A Worst-case to Average-case Reduction for Ideal Lattices

5.1. Summary

In this chapter we achieve a worst-case to average-case reduction for the Hermite Shortest Vector Problem (SVP) on ideal lattices of a fixed number field. Such a reduction allows to transform a fixed chosen instance of a problem (the worst case) to a sample of a fixed *distribution* over all instances of this problem (the average case). Slightly more formally said, a worst-case to average-case reduction consists of two parts: the first one being a *definition of the average-case distribution* and the second one being an *algorithm that reduces any input instance to a sample of that average-case distribution.*

In the reduction of this chapter, which concerns Hermite-SVP on ideal lattices of a fixed number field, this average-case distribution will be defined as something closely related to the *uniform distribution on the Arakelov class group*. This Arakelov class group is essentially the group of ideal lattices up to isometry.

The reduction algorithm in this chapter transforms any fixed input ideal lattice to a sample of the average-case distribution on the Arakelov class group by means of a *random walk*, as introduced in the previous chapter. This 'random walk' transformation of the input ideal lattice only slightly changes its geometry and is therefore compatible with the Hermite Shortest Vector Problem. More concretely, any short vector of the transformed ideal lattice can be reasonably *un*transformed to yield a short vector of the input lattice, with only a small loss in quality.

This particular approach for a worst-case to average-case reduction faces two challenges. The first challenge consists of finding a suitable *representation* of ideal lattices (or Arakelov classes), whereas the second one involves an appropriate treatment of the inherently continuous ideal lattices on finite precision machines.

Such a representation of ideal lattices suitable for the purposes of the worst-case to average-case reduction turns out to be doable by means of a *distribution* over the group of fractional ideals. More precisely, with any fixed ideal lattice we associate an algorithm that efficiently samples from a specific distribution, mainly consisting of fractional ideals that geometrically resemble the input ideal lattice – i.e., whose Arakelov class is close to that of the original ideal lattice. This specific distribution is then our representation of that fixed ideal lattice.

The appropriate treatment of the inherently continuous objects on finite machines happens by *discretization*. A considerable amount of this chapter is devoted to showing that this discretization does not have a significant effect on the overall worst-case to average-case reduction.

5.2. Introduction

The space of all ideal lattices (up to isometry) in a given number field forms naturally an abelian group, called the *Arakelov class group* – a fact well known to number theorists (e.g., [Sch08]). Yet this notion has never appeared explicitly in the literature on lattice-based cryptography. The relevance of this perspective is already illustrated by some previous work which implicitly exploit Arakelov ideals [Eis+14; BS16] and even the Arakelov class group [PHS19; Lee+19]. Beyond its direct result, this chapter aims at highlighting this powerful Arakelov class group formalism for finer and more rigorous analysis of computational problems in ideal lattices.

5.2.1. The Result

We exploit the random walk theorem of Chapter 4 to relate the average-case and the worst-case of Ideal-SVP, due to the interpretation of the Arakelov class group as the space of all ideal lattices up to isometry. Note that this reduction does not directly impact the security of existing schemes: there exists no modern cryptographic scheme based on the average-case version of Ideal-SVP. The value of our result lies in the introduction of a new tool, and an illustration of the cryptanalytic insights it offers.

As already mentioned, ideal lattices (up to isometry) of a given number field K can be identified with the elements of the Arakelov class group, also known as the degree zero part Pic_K^0 of the Picard group. There are two ways to move within this group: given an ideal, one can obtain a new one by 'distorting' it, or by 'sparsifying' it. In both cases, finding a short vector in the target ideal also allows to find a short vector in the source ideal, up to a certain loss of shortness. So, the quality (i.e., the shortness) of the short vector deteriorates with each extra step of the walk; therefore, we minimize the length of the random walk subject to the requirement that the target ideal is uniformly randomly distributed in the Arakelov class group.

This approach leads to a surprisingly tight reduction. In the case of cyclotomic number fields of prime power conductor $m = p^k$, under the Riemann Hypothesis for Hecke L-functions (which we abbreviate ERH for the Extended Riemann Hypothesis), and a mild assumption on the structure of the class groups, the loss of approximation factor is as small as $\widetilde{O}(\sqrt{m})$. In other words:

Main Theorem (informal). *Let $m = p^k$ be a prime power. If there exists a polynomial-time algorithm for solving Hermite-SVP with approximation factor γ over random ideal lattices of $\mathbb{Q}(\zeta_m)$, then there also exists a polynomial time algorithm that solves Hermite-SVP in* any *ideal lattice with approximation factor $\gamma' = \gamma \cdot \sqrt{m} \cdot \mathsf{poly}(\log m)$.*

In fact, this theorem generalizes to all number fields, but the loss in approximation factor needs to be expressed in more involved quantities. The precise

statement is the object of Theorem 5.9.

5.2.2. Overview

The Arakelov class group. Both the unit group [Cra+16] and the class group [CDW17] have been shown to play a key role in the cryptanalysis of ideal lattice problems. In these works of Cramer et al. [Cra+16; CDW17], these groups are exploited independently, in ways that nevertheless share strong similarities with each other. More recently, both groups have been used in combination for cryptanalytic purposes [PHS19; Lee+19]. It therefore seems natural to turn to a unifying theory.

The Arakelov class group (denoted Pic^0_K) is a combination of the unit torus $T = \mathrm{Log}\, K^0_{\mathbb{R}} / \mathrm{Log}(\mathcal{O}_K^\times)$ and of the class group Cl_K. The exponent 0 in $K^0_{\mathbb{R}}$ refers to elements of algebraic norm 1 (i.e., modulo renormalization), while the subscript $\mathbb{R}$ indicates that we are working in the topological completion of K. By 'a combination' we do not exactly mean that Pic^0_K is a direct product; we mean that there is a short exact sequence

$$0 \longrightarrow T \longrightarrow \mathrm{Pic}^0_K \longrightarrow \mathrm{Cl}_K \longrightarrow 0.$$

That is, T is (isomorphic to) a subgroup of Pic^0_K, and Cl_K is (isomorphic to) the quotient Pic^0_K / T. The Arakelov class group is an abelian group which combines an uncountable (yet compact) part T and a finite part Cl_K; topologically, it should be thought of as $|\mathrm{Cl}_K|$ many disconnected copies of the torus T (see Figure 4.1).

A worst-case to average-case reduction for ideal-SVP. An important aspect of the Arakelov class group for the present work is that this group has a geometric interpretation: it can essentially be understood as the group of all ideal lattices up to K-linear isometries. Furthermore, being equipped with a metric, it naturally induces a notion of near-isometry. Such a notion gives a new handle to elucidate the question of the hardness of ideal-SVP. Namely, knowing a short vector in $\mathfrak{a}$, and a near-isometry from $\mathfrak{a}$ to $\tilde{\mathfrak{a}}$, one can

deduce a short vector of $\tilde{\mathfrak{a}}$ up to a small loss induced by the distortion of the near-isometry. This suggests a strategy towards a worst-case to average-case reduction for ideal lattices, namely by randomly distorting a worst-case ideal to a random one (see Figure 5.2).

However, there are two issues with this strategy: first, this near-isometry keeps staying in a fixed class of Cl_K; i.e., one is stuck in one of the potentially many separated copies of the torus that constitute the Arakelov class group. Second, even if $|\mathrm{Cl}_K| = 1$, the unit torus T might be too large, and to reach the full torus from a given point, one may need near-isometry that are too distorted for our purposes.

In the language of algebraic geometry, distortion of ideal lattices corresponds to the 'infinite places' of the field K, while we can also exploit the 'finite places', i.e., the prime ideals. Indeed, if $\mathfrak{c}$ is an integral ideal of small norm and $\tilde{\mathfrak{a}} = \mathfrak{c}\mathfrak{a}$, then $\tilde{\mathfrak{a}}$ is a sublattice of $\mathfrak{a}$ and a short vector of $\tilde{\mathfrak{a}}$ is also a somewhat short vector of $\mathfrak{a}$, an idea already used in [CDW17; PHS19] (see Figure 5.1).

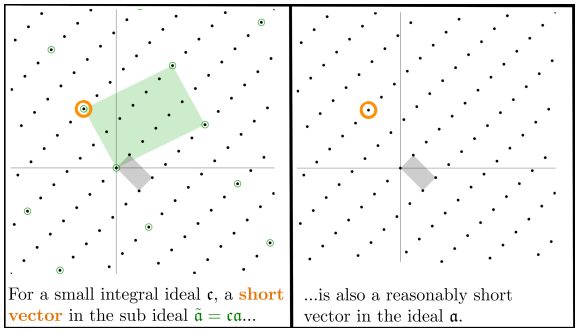

Figure 5.1.: If $\mathfrak{c}$ is an integral ideal of small norm and $\tilde{\mathfrak{a}} = \mathfrak{c}\mathfrak{a}$, then $\tilde{\mathfrak{a}}$ is a sublattice of $\mathfrak{a}$ and a short vector of $\tilde{\mathfrak{a}}$ is also a somewhat short vector of $\mathfrak{a}$.

5.2.3. Related work

Relation to recent cryptanalytic works. The general approach to this result was triggered by a heuristic observation made in [DPW19], suggesting that

the worst-case behavior of the quantum ideal-SVP algorithm built out of [Eis+14; BS16; Cra+16; CDW17] could be made not that far of the average-case behavior they studied experimentally. More specifically, we do achieve the hoped generalization of the class group mixing theorem of [JMV09; JW15] to Arakelov class groups.

Prior self-reduction via random walks. As already mentioned, our result shares strong similarities with a technique introduced by Jao, Miller and Venkatesan [JMV09] to study the discrete logarithm problem on elliptic curves. Just as ideal lattices can be seen as elements of the Arakelov class group, elliptic curves in certain families are in bijective correspondence with elements of the class group of a quadratic imaginary number field. In [JMV09], Jao et al. studied (discrete) random walks on class groups, and showed that they have a rapid mixing property. They deduced that from any elliptic curve, one can efficiently construct a random isogeny (a group homomorphism) to a uniformly random elliptic curve, allowing to transfer a worst case instance of the discrete logarithm problem to an average case instance. Instead of the finite class group, we studied random walks on the infinite Arakelov class group, which led to consequences in lattice-base cryptography, an area seemingly unrelated to elliptic curve cryptography.

Prior self-reduction for ideal lattices. Our self-reducibility result is not the first of its kind: in 2010, Gentry already proposed a self-reduction for an ideal lattice problem [Gen10], as part of his effort of basing Fully-Homomorphic Encryption on worst-case problems [Gen09]. Our result differs in several points.

- Our reduction does not rely on a factoring oracle, and is therefore classically efficient; this was already advertised as an open problem in [Gen10].
- The reduction of Gentry considers the Bounded Distance Decoding problem (BDD) in ideal lattices rather than a short vector problem. Note that this distinction is not significant with respect to quantum computers [Reg09].

- The definition of average case distribution is significantly different, and we view the one of [Gen10] as being somewhat ad-hoc. Given that the Arakelov class group captures exactly ideal lattices up to isometry, we consider the uniform distribution in the Arakelov class group as a much more natural and conceptually simpler choice.
- The worst case ideal input of [Gen10] has restrictions on the size of the norm, whereas our worst case ideal input is unrestricted.
- The loss on the approximation factor of our reduction is much more favorable than the one of Gentry [Gen10]. For example, in the case of cyclotomic number fields with prime-power conductor, Gentry's reduction (on BDD) seems to loose a factor at least $\Theta(n^{4.5})$, while our reduction (on Hermite-SVP) only loses a factor $\tilde{O}(\sqrt{n})$ making a mild assumption on plus-part h^+ of the class number.

Structure of this chapter

We start the remainder of this chapter by constructing an representation of Arakelov class elements that is appropriate to use in a worst-case to average-case reduction (Section 5.3).

After that, we describe a simplified version of the worst-case to average-case reduction; we leave out the difficulties concerning finite machine precision (Section 5.4). In the last part of this chapter, we will show by quite technical means that ignoring finite precision does not impact the reduction significantly (Section 5.5).

5.3. Representation of Ideal Lattices by Means of Distributions

Ideal lattices

Though the notion of ideal lattices is already given in this thesis (see Definition 2.19), we will restate the definition here.

Definition 5.1 (Ideal lattices)**.** *Let K be a number field with ring of integers $\mathcal{O}_K$. An ideal lattice of K is a $\mathcal{O}_K$-module $I \subseteq K_{\mathbb{R}}$, with the additional requirement that there exists an $x \in K_{\mathbb{R}} \backslash \{0\}$ such that $xI \subseteq \mathcal{O}_K$. We denote the group of ideal lattices by* IdLat_K*.*

In essence, the group of ideal lattices IdLat_K can be considered as a sort-of completion of the group of fractional ideals $\mathcal{I}_K$, in the same sense that the reals $\mathbb{R}$ are a completion of $\mathbb{Q}$. A straightforward way to imagine an ideal lattice $x\mathfrak{a} \subseteq K_{\mathbb{R}}$ is to think of an ideal $\mathfrak{a} \subseteq \mathcal{O}_K$ that is 'perturbed' by a vector $x \in K_{\mathbb{R}} = \{y \in \bigoplus_{\sigma:K\to\mathbb{C}} \mathbb{C} \mid y_{\bar{\sigma}} = \overline{y_\sigma}\}$ (see Figure 5.2).

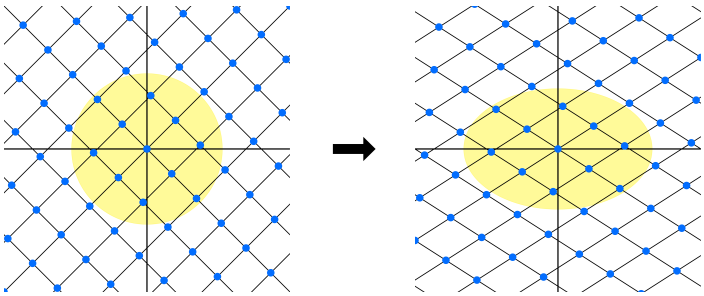

Figure 5.2.: In this two-dimensional example, the left ideal lattice is slightly stretched in the x-direction and slightly shrunk in the y-direction, leading to the perturbed ideal lattice on the right. The yellow circle functions as a visual aid, making the precise deformation of the lattice more explicit.

Representations

Above interpretation immediately gives a representation of the ideal lattice $x\mathfrak{a}$ by the pair $(x, \mathfrak{a}) \in K_{\mathbb{R}} \times \mathcal{I}_K$. But that representation is by no means unique; indeed, one can check that $(x\alpha^{-1}, (\alpha)\mathfrak{a}) \in K_{\mathbb{R}} \times \mathcal{I}_K$, for example, generates the same ideal lattice for any $\alpha \in K^*$. Here, $\alpha^{-1} \in K$ is seen as an element in $K_{\mathbb{R}}$ via the Minkowski embedding $K \hookrightarrow K_{\mathbb{R}}$.

Why do we need an efficient and canonical representation of an ideal lattice (or Arakelov class)?

As mentioned before, a worst-case to average-case distribution of a certain set of problem instances consists of two parts: the definition of a distribution on this set of instances, and an algorithm that reduces any fixed problem instance to this distribution.

Given an Arakelov divisor $\mathbf{a} \in \mathrm{Div}_K^0$, we know how to randomize it so that it is uniformly random in the quotient group Pic_K^0; namely, by the random walk procedure (see Chapter 4). So, for any $\mathbf{a} \in \mathrm{Div}_K^0$ we can efficiently compute a distribution $\mathbb{D}_{\mathbf{a}} \in L_1(\mathrm{Div}_K^0)$ that becomes an uniform distribution under the canonical map $L_1(\mathrm{Div}_K^0) \to L_1(\mathrm{Pic}_K^0), \mathbb{D} \mapsto \sum_{k \in K^*/\mu_K} \mathbb{D}(\cdot + k)$.

To obtain a worst-case to average-case reduction we need a *fixed* (average-case) distribution $\mathbb{D}_0$ on Div_K^0 and an efficient (reduction) map

$$\psi : L_1(\mathrm{Div}_K^0) \to L_1(\mathrm{Div}_K^0)$$

such that for all $\mathbf{a} \in \mathrm{Div}_K^0$, $\psi(\mathbb{D}_{\mathbf{a}}) = \mathbb{D}_0$. Also, this reduction map must be preserving certain geometric properties (be Hermite-SVP compatible) to be an actual useful reduction map.

Suppose for the moment that one has a canonical 'lift' $\mathbb{L} : \mathrm{Div}_K^0 \to \mathrm{Div}_K^0$ for which holds $[\mathbf{a}] = [\mathbf{b}] \Rightarrow \mathbb{L}(\mathbf{a}) = \mathbb{L}(\mathbf{b})$; i.e., it 'factors through' Pic_K^0. And suppose that this map is compatible with Hermite-SVP, i.e., solving Hermite-SVP in $\mathbb{L}(\mathbf{a})$ allows to solve Hermite-SVP in $\mathbf{a}$. Then this lift $\mathbb{L}$ serves as a reduction map, by sending the distribution $\mathbb{D}_{\mathbf{a}} \in L_1(\mathrm{Div}_K^0)$ to $\mathbb{L}(\mathbb{D}_{\mathbf{a}})$, with which we mean the distribution that samples $\mathbb{L}(\mathbf{b})$ with density $\mathbb{D}_{\mathbf{a}}(\mathbf{b})$. By the fact that $\mathbb{D}_{\mathbf{a}}$ maps to the uniform distribution under the canonical map $L_1(\mathrm{Div}_K^0) \to L_1(\mathrm{Pic}_K^0)$, the distribution $\mathbb{L}(\mathbb{D}_{\mathbf{a}}) = \mathbb{D}_0$ is the same for all $\mathbf{a} \in \mathrm{Div}_K^0$. So, such an efficient and Hermite-SVP compatible lift $\mathbb{L}$, which then computes an efficient and canonical representation of ideal lattices, is essentially what remains to construct to make the worst-case to average-case reduction work.

Algorithm 3: Sampling from the distribution $\mathcal{D}_{x\mathfrak{a}}$.

Require: A pair $(x, \mathfrak{a}) \in K_{\mathbb{R}} \times \mathcal{I}_K$ such that $\mathcal{N}(\mathfrak{a}) \prod_\sigma x_\sigma = 1$.

Ensure:

- A sample $\mathfrak{d}^{-1}$ from the distribution $\mathcal{D}_{x\mathfrak{a}}$
- A $v \in x\mathfrak{a}$ such that $\mathfrak{d}^{-1} = v^{-1}x\mathfrak{a}$.

1: Put $\varsigma = 2^{n+1} \cdot n \cdot |\Delta_K|$ and $M = 2\sqrt{n} \cdot \varsigma$.
2: Sample a center $c = (c_\sigma)_\sigma$ uniformly in $\mathcal{C}_M = \{(y_\sigma)_\sigma \mid |y_\sigma| = M \text{ for all embeddings } \sigma\}$.
3: Sample from the discrete Gaussian $\mathcal{G}_{x\mathfrak{a},\varsigma,c}$ with respect to the ideal lattice $x\mathfrak{a}$ with center $c = (c_\sigma)_\sigma$ and standard deviation ς, leading to some $v \in x\mathfrak{a}$.
4: **return** the inverse integral ideal $\mathfrak{d}^{-1} = v^{-1}x\mathfrak{a} \in \mathcal{I}_K$ and $v \in x\mathfrak{a}$.

We could not find such an efficient map $\mathbb{L}$ that is also compatible with Hermite-SVP – instead, we use a map $\mathbb{L} : \mathrm{Div}_K^0 \to L_1(\mathrm{Div}_K^0)$ that is sufficient for our needs. This is a canonical representation by means of a *distribution*. The map we chose has even codomain $L_1(\mathcal{I}_K)$, i.e., involves a discretization for efficiency. So the map $\mathbb{L} : \mathrm{Div}_K^0 \to L_1(\mathcal{I}_K)$ we construct, satisfies $\mathbb{L}(\mathbf{a}) = \mathbb{L}(\mathbf{b})$ for $[\mathbf{a}] = [\mathbf{b}]$ and is compatible with Hermite-SVP.

Concretely, $\mathbb{L}(\mathbf{a})$ consists of sampling a 'balanced' element $\alpha \in \mathbf{a}$ and outputting the ideal $\alpha^{-1} \cdot \mathbf{a} \in \mathcal{I}_K$. This ideal then quite resembles the geometry of $\mathbf{a}$ and lies in the same ideal class; so this ideal (when reduced to the Arakelov class group) must be close to $[\mathbf{a}]$.

Representation by means of a distribution

A representation that is both unique and (in some sense) classically efficiently computable can be made by means of a *distribution*. We will define a map $\mathrm{IdLat}_K \to L_1(\mathcal{I}_K), x\mathfrak{a} \mapsto \mathcal{D}_{x\mathfrak{a}}$ having the property that $\mathcal{D}_{x\mathfrak{a}}$ is an efficiently samplable distribution for any input ideal lattice $x\mathfrak{a} \in \mathrm{IdLat}_K$. The computation of this map $x\mathfrak{a} \mapsto \mathcal{D}_{x\mathfrak{a}}$ is described in Algorithm 3.

Remark 5.2. *It follows from the description of Algorithm 3 that the distribution $\mathcal{D}_{x\mathfrak{a}}$ indeed depends only on the ideal lattice $x\mathfrak{a}$ and not so much on the representation $(x, \mathfrak{a}) \in K_{\mathbb{R}} \times \mathcal{I}_K$ thereof. Hence the notation $\mathcal{D}_{x\mathfrak{a}}$, instead of, for example, $\mathcal{D}_{(x,\mathfrak{a})}$.*

The output of the element $v \in x\mathfrak{a}$ such that $\mathfrak{d}^{-1} = v^{-1}x\mathfrak{a}$ does not take any part in the distribution $\mathcal{D}_{x\mathfrak{a}}$. But it will have a major role in the worst-case to average-case reduction (see Algorithm 4), because it relates $\mathfrak{d}^{-1}$ to the input ideal lattice $x\mathfrak{a}$.

Equivalently, the distribution $\mathcal{D}_{x\mathfrak{a}}$ can be described by the following definition.

Definition 5.3 (Distribution representation of ideal lattices)**.** *Let $(x, \mathfrak{a}) \in K_{\mathbb{R}} \times \mathcal{I}_K$. The distribution $\mathcal{D}_{x\mathfrak{a}} \in L_1(\mathcal{I}_K)$ is supported only by inverse integral ideals. For integral ideals $\mathfrak{d} \in \mathcal{I}_K$ the probability is defined by the following rule.*

$$\mathcal{D}_{x\mathfrak{a}}[\mathfrak{d}^{-1}] = \frac{1}{\mathrm{Vol}(\mathcal{C}_M)} \int_{c \in \mathcal{C}_M} \frac{1}{\rho_\varsigma(x\mathfrak{a} - c)} \sum_{\substack{v \in x\mathfrak{a} \\ (v) = x\mathfrak{a}\mathfrak{d}}} \rho_\varsigma(v - c) dc, \tag{5.65}$$

where $\varsigma = 2^{n+1} \cdot n \cdot |\Delta_K|$, $M = 2\sqrt{n} \cdot \varsigma$ and $\mathcal{C}_M = \{(x_\sigma)_\sigma \in K_{\mathbb{R}} \mid |x_\sigma| = M\}$, the M-hypercircle in $K_{\mathbb{R}}$.

The fact that $\mathcal{D}_{x\mathfrak{a}}$ is a distribution follows by the following computation.

$$\begin{aligned}\sum_{\mathfrak{d} \in \mathcal{I}_K} \mathcal{D}_{x\mathfrak{a}}[\mathfrak{d}^{-1}] &= \frac{1}{\mathrm{Vol}(\mathcal{C}_M)} \int_{c \in \mathcal{C}_M} \frac{1}{\rho_\varsigma(x\mathfrak{a} - c)} \underbrace{\sum_{\mathfrak{d} \in \mathcal{I}_K} \sum_{\substack{v \in x\mathfrak{a} \\ (v) = x\mathfrak{a}\mathfrak{d}}} \rho_\varsigma(v - c)}_{\rho_\varsigma(x\mathfrak{a} - c)} dc \\ &= \frac{1}{\mathrm{Vol}(\mathcal{C}_M)} \int_{c \in \mathcal{C}_M} dc = 1.\end{aligned}$$

Remark 5.4. *The instantiation of $\varsigma \in \mathbb{R}_{>0}$ in Definition 5.3 is chosen this way because of the lower bound $\varsigma \geq 2^{n+1}\sqrt{n} \cdot |\Delta_K|^{1/(2n)} \cdot \lambda_n(\mathcal{O}_K)$ and $\lambda_n(\mathcal{O}_K) \geq n\sqrt{|\Delta_K|}$ (see Lemma 2.22). The first of these lower bounds arises*

from the size of an LLL-reduced basis of $x\mathfrak{a}$; and the standard deviation ς needs to be larger than this basis size for an efficient computation of the discrete Gaussian over $x\mathfrak{a}$ by Klein's algorithm [GPV08; Kle00].

The instantiation of $M = 2\sqrt{n} \cdot \varsigma$ (or larger) is required in order to have a balanced $v \in K_{\mathbb{R}}$ in line 3 of Algorithm 3. A balanced $v \in K_{\mathbb{R}}$ means that all entries v_σ of v are of roughly the same size, i.e., that $\frac{\max_\sigma |v_\sigma|}{\min_\sigma |v_\sigma|}$ is small. This has as a consequence that $\mathfrak{d}^{-1} \leftarrow \mathcal{D}_{x\mathfrak{a}}$ and $x\mathfrak{a}$ have a very similar geometry (see Lemma 5.5 part (iii)).

Properties of the distribution representation

Because the distribution $\mathcal{D}_{(x,\mathfrak{a})}$ in Definition 5.3 depends on the ideal lattice $x\mathfrak{a}$ and not on the representing pair $(x, \mathfrak{a})$, we can see the domain of the map $\mathcal{D}_\cdot$ as the group of ideal lattices IdLat_K, i.e., $\mathcal{D}_\cdot : \mathrm{IdLat}_K \to L_1(\mathcal{I}_K)$, $x\mathfrak{a} \to \mathcal{D}_{x\mathfrak{a}}$. Even more is true – two *isometric* ideal lattices $x\mathfrak{a}$ and $y\mathfrak{b}$ also have the same distribution $\mathcal{D}_{x\mathfrak{a}}$ and $\mathcal{D}_{y\mathfrak{b}}$. Two ideal lattices being isometric means that there exists an element $\xi = (\xi_\sigma)_\sigma \in \mathcal{C}_1 = \{(y_\sigma)_\sigma \mid |y_\sigma| = 1\}$ such that $x\mathfrak{a} = \xi y\mathfrak{b}$ (see Definition 2.20). So, this map can even be interpreted to have domain Pic_K^0.

Another two remarkable properties of the distribution $\mathcal{D}_{x\mathfrak{a}}$ are that $\mathfrak{d}^{-1} \leftarrow \mathcal{D}_{x\mathfrak{a}}$ always lies in the ideal class $[\mathfrak{a}]$ and has (with high probability) a geometry very similar to $x\mathfrak{a}$. So, in some sense, we may see a sample $\mathfrak{d}^{-1} \leftarrow \mathcal{D}_{x\mathfrak{a}}$ as a sort of 'discrete approximation' of the ideal lattice $x\mathfrak{a}$. These important properties of the distribution representation are spelled out in the following lemma.

Lemma 5.5 (Properties of the distribution representation)**.** *The map $\mathcal{D}_\cdot : \mathrm{IdLat}_K \to L_1(\mathcal{I}_K)$ has the following properties.*

(i) (Isometric lattices have the same distribution) For all $x\mathfrak{a}, y\mathfrak{b} \in \mathrm{IdLat}_K$ which are isometric, i.e., $x\mathfrak{a} \sim y\mathfrak{b}$, we have $\mathcal{D}_{x\mathfrak{a}} = \mathcal{D}_{y\mathfrak{b}}$.

(ii) (Supported by a single ideal class) For all $x\mathfrak{a} \in \mathrm{IdLat}_K$*, the distribution* $\mathcal{D}_{x\mathfrak{a}}$ *on* $\mathcal{I}_K$ *is supported only by inverted integral ideals that lie in the ideal class* $[\mathfrak{a}]$.

(iii) (Bounded size) For all $x\mathfrak{a} \in \mathrm{IdLat}_K$ *with* $\mathcal{N}(\mathfrak{a}) \prod_\sigma x_\sigma = 1$*, the weight of the distribution* $\mathcal{D}_{x\mathfrak{a}}$ *is concentrated on inverted integral ideals* $\mathfrak{d}^{-1}$ *for which holds* $\mathcal{N}(\mathfrak{d}^{-1}) \geq (\varsigma + M)^{-n}$*. Concretely,*

$$\Pr_{\mathfrak{d}^{-1} \leftarrow \mathcal{D}_{x\mathfrak{a}}} \left[\mathcal{N}(\mathfrak{d}^{-1}) < (\varsigma + M)^{-n}\right] \leq 2e^{-n}.$$

(iv) (Similar geometry) For all $x\mathfrak{a} \in \mathrm{IdLat}_K$ *with* $\mathcal{N}(\mathfrak{a}) \prod_\sigma x_\sigma = 1$*, for almost all* $\mathfrak{d}^{-1} \leftarrow \mathcal{D}_{x\mathfrak{a}}$*, we have* $\mathfrak{d}^{-1} = v^{-1}x\mathfrak{a}$ *with* $\|v\|_\infty \|v^{-1}\|_\infty \leq 3$*, i.e.,* v *is* balanced*. Concretely,*

$$\Pr_{\mathfrak{d}^{-1} \leftarrow \mathcal{D}_{x\mathfrak{a}}} \left[\exists v \in K_\mathbb{R} : \mathfrak{d}^{-1} = v^{-1}x\mathfrak{a} \quad \text{and} \quad \|v\|_\infty \|v^{-1}\|_\infty \leq 3\right] \geq 1 - 2e^{-n}$$

Proof. (i) Write $x\mathfrak{a} = \xi y\mathfrak{b}$, use Definition 5.3 and use the fact that $|\xi_\sigma| = 1$ for all embeddings σ to deduce, for a fixed integral ideal $\mathfrak{d}$,

$$\begin{aligned}\frac{1}{\rho_\varsigma(x\mathfrak{a} - c)} \sum_{\substack{v \in x\mathfrak{a} \\ (v) = x\mathfrak{a}\mathfrak{d}}} \rho_\varsigma(v - c) &= \frac{1}{\rho_\varsigma(\xi y\mathfrak{b} - c)} \sum_{\substack{v \in y\mathfrak{b} \\ (v) = y\mathfrak{b}\mathfrak{d}}} \rho_\varsigma(\xi v - c) \\ &= \frac{1}{\rho_\varsigma(y\mathfrak{b} - \xi^{-1}c)} \sum_{\substack{v \in y\mathfrak{b} \\ (v) = y\mathfrak{b}\mathfrak{d}}} \rho_\varsigma(v - \xi^{-1}c).\end{aligned}$$

The map $c \mapsto \xi^{-1}c$ is an isometric smooth bijection on the hypercircle $\mathcal{C}_M$, so integrating with respect to the variable $\xi^{-1}c$ or c for $c \in \mathcal{C}_M$ doesn't change the value of the integral. Therefore, $\mathcal{D}_{(x,\mathfrak{a})}[\mathfrak{d}^{-1}] = \mathcal{D}_{(y,\mathfrak{b})}[\mathfrak{d}^{-1}]$ for all $\mathfrak{d} \in \mathcal{I}_K$.

(ii) From Equation (5.65) we see that $\mathfrak{d} = (v)/(x\mathfrak{a})$ for $v \in x\mathfrak{a}$ and therefore $\mathfrak{d}$ must be an integral ideal in the inverse class of $\mathfrak{a}$; so $\mathfrak{d}^{-1}$ lies in the ideal class $[\mathfrak{a}]$.

(iii) Use the fact that $\mathcal{N}(x\mathfrak{a}) = 1$ and the fact that $(v) = x\mathfrak{a}\mathfrak{d}$ to derive

$$\begin{aligned}\Pr_{\mathfrak{d}\leftarrow\mathcal{D}_{x\mathfrak{a}}}[\mathcal{N}(\mathfrak{d}) > (\varsigma + M)^n] &= \Pr_{\mathfrak{d}\leftarrow\mathcal{D}_{x\mathfrak{a}}}[\mathcal{N}(x\mathfrak{a}\mathfrak{d}) > (\varsigma + M)^n] \\ &\leq \Pr_{\substack{c\leftarrow\mathcal{C}_M \\ v\leftarrow\mathcal{G}_{x\mathfrak{a},\varsigma,c}}}[\|v\|_2 > \sqrt{n}\cdot\varsigma + \sqrt{n}\cdot M]. \\ &\leq \max_{c\in\mathcal{C}_M} \Pr_{v\leftarrow\mathcal{G}_{x\mathfrak{a},\varsigma,c}}[\|v - c\|_2 > \sqrt{n}\cdot\varsigma].\end{aligned}$$

Where the first inequality follows from norm inequalities; we have $n^{-1/2}\cdot\|v\|_2 \geq n^{-1}\cdot\|v\|_1 \geq \mathcal{N}(v)^{1/n} = \mathcal{N}(x\mathfrak{a}\mathfrak{d})^{1/n} \geq (\varsigma + M)$. The second inequality follows from the triangle inequality and the fact that $\|c\| = \sqrt{n}\cdot M$. By Banaszczyk's tail bound (see Lemma 2.25) and by smoothing arguments (see Lemma 2.31), we conclude

$$\Pr_{v\leftarrow\mathcal{G}_{x\mathfrak{a},\varsigma,c}}[\,\|v - c\| \geq \sqrt{n}\cdot\varsigma] \leq e^{-n}\cdot\frac{\rho_\varsigma(x\mathfrak{a})}{\rho_\varsigma(x\mathfrak{a} - c)} \leq 2e^{-n}.$$

For the smoothing argument we use the fact that $\varsigma = 2^{n+1}\cdot n\cdot|\Delta_K| \geq n\lambda_n(\mathcal{O}_K) \geq \eta_1(x\mathfrak{a})$ (see [MR07, Lm. 3.3 and 3.4]).

(iv) We have, by the norm inequalities, $\|v - c\| \geq \|v - c\|_\infty$, and therefore, by part (iii) of this lemma,

$$\Pr_{v\leftarrow\mathcal{G}_{x\mathfrak{a},\varsigma,c}}[\,\|v - c\|_\infty \geq \sqrt{n}\cdot\varsigma] \leq \Pr_{v\leftarrow\mathcal{G}_{x\mathfrak{a},\varsigma,c}}[\,\|v - c\| \geq \sqrt{n}\cdot\varsigma] \leq 2e^{-n}.$$

Since $|c_\sigma| = M$ for all embeddings σ, and since $M = 2\sqrt{n}\varsigma$, we have, except with probability $2e^{-n}$,

$$\|v\|_\infty\left\|v^{-1}\right\|_\infty = \frac{\max_\sigma|v_\sigma|}{\min_\sigma|v_\sigma|} \leq \frac{M + \sqrt{n}\varsigma}{M - \sqrt{n}\varsigma} = 3.$$

□

Remark 5.6. *The bound $\|v\|_\infty\|v^{-1}\| \leq 3$ w.h.p. in part (iv) of Lemma 5.5 can be tightened to $\|v\|_\infty\|v^{-1}\| \leq 1 + O(e^{-n})$ by taking $M = 2^n\cdot\varsigma$. Because this would only remove a rather non-significant constant 3 in the quality loss of the output in the worst-case to average-case reduction, we choose this 'constant bound' for simplicity.*

A consequence of Lemma 5.5 is that the map $\mathcal{D}_\cdot : \mathrm{IdLat}_K \to L_1(\mathcal{I}_K)$, that sends ideal lattices to distributions on $\mathcal{I}_K$ factors through the quotient group $\mathrm{IdLat}_K / \sim$, where '$\sim$' stands for factoring out by isometries. As the group of ideal lattices up to isometries is naturally isomorphic to the Arakelov class group Pic_K^0 (see Lemma 2.21), we might as well consider Pic_K^0 as the domain of the map $\mathcal{D}$.

5.4. The Worst-case to Average-case Reduction

Introduction

A worst-case to average-case reduction consists of two main parts: the definition of the average-case distribution and an algorithm that reduces a fixed problem instance to a sample of the average-case distribution.

We start this section with the definition of the average-case distribution, which is derived from the uniform distribution on the Arakelov class group. After that, we will describe the reduction algorithm. In this description of the worst-case to average-case reduction we temporarily ignore issues regarding real numbers. In the last part of this section we will prove the correctness of the reduction algorithm and examine the precise quality loss that occurs in the reduction.

Discussing and solving the issues regarding real numbers and finite precision in the distribution algorithm Algorithm 3 and the reduction algorithm Algorithm 4 is deferred to Section 5.5. In that section we will prove that both the distribution algorithm and the reduction algorithm can be run efficiently on a finite machine by means of appropriate discretization.

Definition of the average-case distribution

Knowing in advance that the reduction algorithm will make use of the random walk machinery of Chapter 4, which leads to a near-uniform distribution

on the Arakelov class group, the average-case distribution must be strongly tied to this distribution.

Indeed, we define the average-case distribution to be distribution on $L_1(\mathcal{I}_K)$ defined by the following rule.

$$\mathcal{D}_{\mathcal{U}(\mathrm{Pic}_K^0)}[\mathfrak{d}^{-1}] = \frac{1}{|\mathrm{Pic}_K^0|} \int_{\mathbf{a} \in \mathrm{Pic}_K^0} \mathcal{D}_{\mathbf{a}}[\mathfrak{d}^{-1}] d\mathbf{a}. \tag{5.66}$$

In essence this is just 'taking the average' of all distributions $\mathcal{D}_{\mathbf{a}}$ (as in Section 5.3) where $\mathbf{a}$ is taken uniformly from the Arakelov class group.

Reduction algorithm

The reduction algorithm essentially consists of taking an input ideal lattice $x\mathfrak{a}$, applying a specific random walk procedure on it as in Chapter 4, yielding $\tilde{x}\tilde{\mathfrak{a}}$, and sampling an ideal $\mathfrak{d}^{-1} \leftarrow \mathcal{D}_{\tilde{x}\tilde{\mathfrak{a}}}$. A rigorous, precise description of this procedure is spelled out in Algorithm 4.

Remark 5.7. *Algorithm 4, and also Algorithm 3, are strictly spoken not algorithms that can be run on a finite computer, because of the continuous distributions occurring in the algorithm descriptions. In Algorithm 3 it is the uniform sampling from the hypercircle $\mathcal{C}_M$ and in Algorithm 4 it is the Gaussian sampling that is inherently continuous.*

In Section 5.5 we will show that those continuous distributions can be efficiently discretized without a significant impact on the final result. Therefore, we just ignore these continuity issues for now, for the sake of clarity and brevity.

Explanation of the reduction algorithm

Algorithm 4: The worst-case to average-case reduction algorithm

Require:

- A pair $(x, \mathfrak{a}) \in K_{\mathbb{R}} \times \mathcal{I}_K$ satisfying $\mathcal{N}(\mathfrak{a}) \prod_\sigma x_\sigma = 1$.
- The values $[\Lambda_K : C]$ and $\eta_1(C^*)$ of a suitable sublattice $C \subseteq \Lambda_K$ of the logarithmic unit lattice,
- An oracle $\mathcal{A}$ that solves γ-Hermite SVP in $\mathfrak{d}^{-1}$ whenever $\mathfrak{d}^{-1} \leftarrow \mathcal{D}_{\mathcal{U}(\mathrm{Pic}_K^0)}$.

Ensure: A vector $\alpha \in x\mathfrak{a}$ that is a solution to $B^{1/n} \cdot \gamma$-Hermite SVP in the ideal lattice $x\mathfrak{a}$, i.e.,

$$\|\alpha\| \leq \gamma \cdot B^{1/n} \cdot \det(x\mathfrak{a})^{1/n},$$

where $B = \widetilde{O}(4^n \cdot \eta_1(C^*)^{\mathrm{r}} \cdot [\Lambda_K : C] \cdot |\operatorname{Pic}_K^0| \cdot (\log|\Delta_K|)^2)$, or, *failure*.

1: Put $s = \max(\sqrt{2} \cdot \eta_1(C^*), (\log n)^2)$ and $B = \widetilde{O}(4^n \cdot \eta_1(C^*)^{\mathrm{r}} \cdot [\Lambda_K : C] \cdot |\operatorname{Pic}_K^0| \cdot (\log|\Delta_K|)^2)$ as in Corollary 5.8.
2: Multiply the ideal $\mathfrak{a}$ by a prime ideal $\mathfrak{p}$ uniformly sampled from the set $\{\mathfrak{p} \text{ prime ideal of } \mathcal{O}_K \mid \mathcal{N}(\mathfrak{p}) \leq B\}$, yielding $\mathfrak{a}\mathfrak{p}$.
3: Sample a Gaussian distributed $y \leftarrow \mathcal{G}_{s,H}$, where H is the hyperplane where the logarithmic unit lattice lives in.
4: Put $p = \mathcal{N}(\mathfrak{p})^{1/n}$, so that $e^y x\mathfrak{a}\mathfrak{p}/p$ has norm 1, where $e^y \in K_{\mathbb{R}}$ is the component-wise exponentiation of $y \in H$.
5: Sample $\mathfrak{d}^{-1} \leftarrow \mathcal{D}_{e^y \cdot x \cdot \mathfrak{a}\mathfrak{p}/p}$ using Algorithm 3, and let $v \in e^y x\mathfrak{a}\mathfrak{p}/p$ be the additional output of Algorithm 3 that satisfies $\mathfrak{d}^{-1} = v^{-1} e^y x\mathfrak{a}\mathfrak{p}/p$.
6: Invoke the γ-Hermite SVP oracle $\mathcal{A}$ on $\mathcal{D}_{\mathcal{U}(\mathrm{Pic}_K^0)}$ to find a $\kappa \in \mathfrak{d}^{-1}$ for which holds $\|\kappa\| \leq \gamma \cdot \det(\mathfrak{d}^{-1})^{1/n}$
7: **return** $p \cdot e^{-y} \cdot v \cdot \kappa \in x\mathfrak{a}$.

Randomize the input ideal lattice $x\mathfrak{a}$. The first four steps of Algorithm 4 actually applies a random walk on the input ideal lattice $x\mathfrak{a}$, resulting in a randomized ideal lattice $\tilde{x}\tilde{\mathfrak{a}}$. By the results of Chapter 4, this ideal lattice is nearly uniformly distributed in the Arakelov class group. Therefore, sampling from the distribution $\mathcal{D}_{\tilde{x}\tilde{\mathfrak{a}}} \approx \mathcal{D}_{\mathcal{U}(\mathrm{Pic}_K^0)}$ associated with this random ideal lattice $\tilde{x}\tilde{\mathfrak{a}}$ then yields an ideal $\mathfrak{d}^{-1}$ that must be closely distributed as $\mathcal{D}_{\mathcal{U}(\mathrm{Pic}_K^0)}$. So, that is an intuitive idea of why the output ideal $\mathfrak{d}^{-1}$ is almost distributed as the *average-case distribution* as in Equation (5.66).

Sample $\mathfrak{d}^{-1} \leftarrow \mathcal{D}_{\tilde{x}\tilde{\mathfrak{a}}}$ and apply the Hermite-SVP oracle $\mathcal{A}$ on $\mathfrak{d}^{-1}$. Because $\mathfrak{d}^{-1} \leftarrow \mathcal{D}_{\tilde{x}\tilde{\mathfrak{a}}}$ is so close to the average-case distribution, we can actually invoke the oracle $\mathcal{A}$ to find a short vector in the ideal $\mathfrak{d}^{-1} \leftarrow \mathcal{D}_{\tilde{x}\tilde{\mathfrak{a}}}$. The sampling is done in step 5 and calling the oracle in step 6 of Algorithm 4.

Transform the short vector $\gamma \in \mathfrak{d}^{-1}$ into a short vector in the randomized ideal lattice $\tilde{x}\tilde{\mathfrak{a}}$. Recall that Algorithm 3 on input $\tilde{x}\tilde{\mathfrak{a}}$ outputs both $\mathfrak{d}^{-1} \leftarrow \mathcal{D}_{\tilde{x}\tilde{\mathfrak{a}}}$ and a $v \in \tilde{x}\tilde{\mathfrak{a}}$ such that $\mathfrak{d}^{-1}v = \tilde{x}\tilde{\mathfrak{a}}$.

So, any short vector $\kappa \in \mathfrak{d}^{-1}$ can be transformed into a short vector $v\kappa \in \tilde{x}\tilde{\mathfrak{a}}$. Because this $v \in \tilde{x}\tilde{\mathfrak{a}}$ is *balanced*, this does not affect the shortness of the vector κ much; in a way one might say that the ideal lattices $\tilde{x}\tilde{\mathfrak{a}}$ and $\mathfrak{d}^{-1}$ geometrically very much resemble each other.

Transform a short vector in $\tilde{x}\tilde{\mathfrak{a}}$ to a short vector in the input ideal lattice $x\mathfrak{a}$. By construction, $\tilde{x}\tilde{\mathfrak{a}} = e^y/p \cdot x\mathfrak{a}\mathfrak{p}$, i.e., the randomized ideal lattice is just the input ideal lattice multiplied by a prime ideal, slightly disturbed a renormalized. By undoing the disturbance (i.e., dividing by e^y) and undoing the renormalization (i.e., multiplying by $p = \mathcal{N}(\mathfrak{p})^{1/n}$) on the short vector $v\kappa \in \tilde{x}\tilde{\mathfrak{a}}$, we obtain a short vector in $x\mathfrak{a}\mathfrak{p} \subseteq x\mathfrak{a}$. More precisely: because $v\kappa \in \tilde{x}\tilde{\mathfrak{a}}$, we have that $p \cdot e^{-y} \cdot (v\kappa) \in x\mathfrak{a}\mathfrak{p} \subseteq x\mathfrak{a}$.

Reason for quality loss. Note that the reduction algorithm only ensures to find a vector solving $B^{1/n} \cdot \gamma$-Hermite SVP, whereas the oracle $\mathcal{A}$ in

Algorithm 4 is assumed to be able to find a vector satisfying γ-Hermite SVP on an 'average case' ideal lattice (see Equation (5.66)).

This particular loss $B^{1/n}$ comes from the fact that we cannot not reasonably 'undo' the part of the random walk where we multiply the input ideal lattice $x\mathfrak{a}$ by a random prime ideal $\mathfrak{p}$. So, this reduction algorithm actually finds a γ-Hermite short vector in $x\mathfrak{a}\mathfrak{p}$, a slightly wider ideal than $x\mathfrak{a}$. As $x\mathfrak{a}\mathfrak{p} \subseteq x\mathfrak{a}$, and the root determinants of these ideal lattices differ with a factor $p = \mathcal{N}(\mathfrak{p})^{1/n}$, a γ-Hermite short vector in $x\mathfrak{a}\mathfrak{p}$ is a $B^{1/n} \cdot \gamma$-Hermite short vector in $x\mathfrak{a}$, as $\mathcal{N}(\mathfrak{p})^{1/n} \leq B^{1/n}$.

Proof of correctness and quantification of the quality loss

In order to prove the result of this chapter, we need the following specialization of the random walk theorem of Chapter 4, which is specifically tailored to the worst-case to average-case reduction.

Corollary 5.8 (Random walk in the Arakelov class group, simplified)**.** *Let K be a number field, and let $C \subseteq \Lambda_K$ be a sublattice of the logarithmic unit lattice. Assuming the Extended Riemann Hypothesis, there exists a bound $B = \widetilde{O}(4^n \cdot \eta_1(C^*)^{\mathrm{r}} \cdot [\Lambda_K : C] \cdot |\operatorname{Pic}_K^0| \cdot (\log|\Delta_K|)^2)$ such that the random walk distribution with* one *step $\mathcal{W}_{\operatorname{Pic}_K^0}(B, 1, s)$ is exponentially close to uniform in $L_1(\operatorname{Pic}_K^0)$.*

$$\|\mathcal{W}_{\operatorname{Pic}_K^0}(B, 1, s) - \mathcal{U}(\operatorname{Pic}_K^0)\|_1 \leq 2^{-n}$$

Proof. Apply Theorem 4.18 from Chapter 4 with

- $k = \frac{1}{2\log n} \cdot (\mathbb{r} \cdot \log(1/\tilde{s}) + \log(\operatorname{Vol}(\operatorname{Pic}_K^0)) + 2\log(1/\varepsilon) + \log[\Lambda_K : C] + 2)$, so that taking $N = 1$ satisfies the requirements of the theorem.
- $s = 1/(\sqrt{2} \cdot \eta_1(C^*))$, so that $\tilde{s} = 1/\eta_1(C^*)$;
- $\varepsilon = 2^{-n}$.

Appropriately substituting above instantiations in the formula for B in Theorem 4.18, noting that $n^{2k} = O(\eta_1(C^*)^{\mathbb{r}} \cdot |\operatorname{Pic}_K^0| \cdot 4^n \cdot [\Lambda_K : C])$, we obtain

$$B = \widetilde{O}\Big(n^{2k}[n^2(\log\log(1/\varepsilon))^2 + n^2(\log(1/\tilde{s}))^2 + n^2\log([\Lambda_K : C])^2 + (\log|\Delta_K|)^2]\Big)$$

$$= \widetilde{O}(4^n \cdot \eta_1(C^*)^{\mathbb{r}} \cdot [\Lambda_K : C] \cdot |\operatorname{Pic}_K^0| \cdot (\log|\Delta_K|)^2).$$

□

Using above specialized random walk theorem, we can prove the main theorem of this chapter.

Theorem 5.9. *Let K be a number field with logarithmic unit lattice Λ_K, let $C \subseteq \Lambda_K$ be any sublattice, and denote its dual lattice by C^*. Put $1/s = \max(\sqrt{2} \cdot \eta_1(C^*), \log(n)^2)$.*

Assume we have a (possibly randomized) algorithm $\mathcal{A}$ that solves γ-Hermite-SVP within an approximation factor $\gamma \geq 1$ and probability[1] at least $q > 0$ when given an input $\mathfrak{a}$ with $\mathfrak{a} \leftarrow \mathcal{D}_{\mathcal{U}(\operatorname{Pic}_K^0)}$.

Then there exists a randomized algorithm $\mathcal{B}$ solving $(O(B^{1/n}) \cdot \gamma)$-Hermite-SVP in any ideal lattice $x\mathfrak{a} \in \operatorname{IdLat}_K$, with probability[2] at least $q - n^{-\omega(1)}$, where $B = \widetilde{O}(4^n \cdot s^{-\mathbb{r}} \cdot [\Lambda_K : C] \cdot |\operatorname{Pic}_K^0| \cdot (\log|\Delta_K|)^2)$. The algorithm $\mathcal{B}$ runs within time polynomial in $\log|\Delta_K|, \log[\Lambda_K : C], \operatorname{size}(x)$ and $\operatorname{size}(M_{\mathfrak{a}})$ on input $(x, \mathfrak{a}) \in K_{\mathbb{R}} \times \mathcal{I}_K$ and needs one call to the algorithm $\mathcal{A}$.

[1] Here, the probability q is taken over the random choice of $\mathfrak{a} \leftarrow \mathcal{D}_{\mathcal{U}(\operatorname{Pic}_K^0)}$ and over the possible internal randomness of the algorithm $\mathcal{A}$

[2] Here, the probability is taken over the internal randomness of $\mathcal{B}$

Furthermore, the loss $B^{1/n}$ in the approximation factor of Hermite-SVP in the reduction can be upper bounded as follows.

$$B^{1/n} = \begin{cases} \widetilde{O}(\sqrt{n}) & \text{if } K = \mathbb{Q}(\zeta_{p^k}), \text{ a prime power cyclotomic field, assuming that } h_K^+ = \log(n)^{O(n)}. \\ \widetilde{O}(n^{1-n_{\mathbb{C}}/n} \cdot |\Delta_K|^{1/(2n)}) & \text{otherwise} \end{cases} \tag{5.67}$$

Proof. Proof of the running time. The random walk process and the distribution representation of the reduction have inherently continuous aspects, that need to be discretized in order to be suitable for an actual computer. The discretized version of the reduction is treated in Section 5.5, in which also its running time and its discretization error is studied. In Theorem 5.11 we show that the reduction can be approximated within a negligible error margin, using time polynomial in $\log|\Delta_K|, \log[\Lambda_K : C], \mathrm{size}(x)$; here we take $\varepsilon = 2^{-n}$ to have exponentially small error.

Success probability. By the choice of parameters in reduction Algorithm 4, the Arakelov class of $xe^y/p \cdot \mathfrak{ap}$ (where $\mathfrak{p}$ and $y \in H$ are randomly chosen as in Algorithm 4) must be exponentially close to uniform in Pic_K^0 in total variation distance (see Corollary 5.8). By the data processing inequality [CT06, §2.8], $\mathcal{D}_{xe^y/p\cdot\mathfrak{ap}}$ is exponentially close to $\mathcal{D}_{\mathcal{U}(\mathrm{Pic}_K^0)}$ as well. Therefore, the algorithm $\mathcal{A}$ cannot distinguish reasonably between the two distributions and outputs with probability at least $q - 2^{-n}$ a solution of γ-Hermite SVP in $\mathfrak{d}^{-1} \leftarrow \mathcal{D}_{xe^y/p\cdot\mathfrak{ap}}$.

Quality of the output. Let us assume that algorithm $\mathcal{A}$ indeed found a solution to γ-Hermite SVP, i.e., a vector $\kappa \in \mathfrak{d}^{-1}$ which satisfies $\|\kappa\| \leq \gamma \cdot \det(\mathfrak{d}^{-1})^{1/n}$, where $\mathfrak{d}^{-1} \leftarrow \mathcal{D}_{xe^y/p\cdot\mathfrak{ap}}$.

As $\kappa \in \mathfrak{d}^{-1} = v^{-1}e^y x/p \cdot \mathfrak{ap}$ (see Algorithm 3), we must have that[3] $\kappa = v^{-1}e^y/p \cdot \alpha$ for some $\alpha \in x\mathfrak{ap}$. *This* particular $\alpha \in x\mathfrak{ap} \subset x\mathfrak{a}$ is a solution for $O(B^{1/n}) \cdot \gamma$-Hermite SVP in $x\mathfrak{a}$, which can be seen by the following

[3] Note that $p = \mathcal{N}(\mathfrak{p})^{1/n}$

lines of reasoning. We have the following bound on $\|\alpha\|$, where we write out $p = \mathcal{N}(\mathfrak{p})^{1/n}$,

$$\begin{aligned}\|\alpha\| = \|ve^{-y}\kappa\| \cdot \mathcal{N}(\mathfrak{p})^{1/n} &\leq \|v\|_\infty \|e^{-y}\|_\infty \|\kappa\| \cdot \mathcal{N}(\mathfrak{p})^{1/n} \\ &\leq \|v\|_\infty \|e^{-y}\|_\infty \cdot \gamma \cdot \det(\mathfrak{d}^{-1})^{1/n} \cdot \mathcal{N}(\mathfrak{p})^{1/n},\end{aligned} \tag{5.68}$$

But also, by the fact that multiplication by e^y doesn't change the determinant and $\det(x/p \cdot \mathfrak{a}\mathfrak{p}) = \det(x\mathfrak{a})$ (by definition of $p = \mathcal{N}(\mathfrak{p})^{1/n}$), we have

$$\det(\mathfrak{d}^{-1}) = \det(v^{-1}e^y x/p \cdot \mathfrak{a}\mathfrak{p}) \leq \|v^{-1}\|_\infty^n \cdot \det(x\mathfrak{a}). \tag{5.69}$$

Combining Equation (5.68) and Equation (5.69), using the fact that $\mathcal{N}(\mathfrak{p}) \leq B$, and $\|v^{-1}\|_\infty \|v\|_\infty \leq 3$ with high probability (see Lemma 5.5, 'v is balanced'), we obtain

$$\begin{aligned}\|\alpha\| &\leq \underbrace{\|v\|_\infty \cdot \|v^{-1}\|_\infty}_{\leq 3 \ (w.h.p.)} \cdot \underbrace{\|e^{-y}\|_\infty}_{\leq 3 \ (w.h.p.)} \cdot \underbrace{\mathcal{N}(\mathfrak{p})^{1/n}}_{\leq B^{1/n}} \cdot \gamma \cdot \det(x\mathfrak{a})^{1/n} \\ &\leq 9 \cdot B^{1/n} \cdot \gamma \cdot \det(x\mathfrak{a})^{1/n}.\end{aligned}$$

Here, the bound on $\|e^y\|_\infty$ can be obtained by the fact that $y \leftarrow \mathcal{G}_{H,s}$ is from a Gaussian distribution, with[4] $s \leq 1/\log(n)^2$. Namely, $\|y\|_\infty \leq (\log n)^2 s \leq 1$ except with probability at most $2^{-\Omega((\log n)^2)} = n^{-\omega(1)}$. Therefore $\|e^y\|_\infty \leq e^{\|y\|_\infty} \leq 3$ except with probability $n^{-\omega(1)}$.

Conclusion. So, with probability $q - n^{-\omega(1)}$, algorithm $\mathcal{B}$ solves $9 \cdot B^{1/n} \cdot \gamma$-Hermite SVP in the input ideal lattice $x\mathfrak{a} \in \mathrm{IdLat}_K$, within polynomial time in $\log|\Delta_K|, \log[\Lambda_K : C]$, size$(x)$ and size$(M_\mathfrak{a})$, and using one call to the algorithm $\mathcal{A}$. The explicit bounds on $B^{1/n}$ in Equation (5.67) are proved in Proposition 5.10. □

[4] Note that $s \leq 1/(\log n)^2$, by the instantiation $1/s = \max(\sqrt{2} \cdot \eta_1(C^*), \log(n)^2)$ in the theorem.

Proposition 5.10. *The loss $B^{1/n}$ in the approximation factor of Hermite-SVP in the reduction of Theorem 5.9 can be upper bounded as follows.*

$$B^{1/n} = \begin{cases} \widetilde{O}(\sqrt{n}) & \text{if } K = \mathbb{Q}(\zeta_{p^k}), \text{ a prime power cyclotomic field, assuming that } h_K^+ = \log(n)^{O(n)}. \\ \widetilde{O}(n^{1-n_{\mathbb{C}}/n} \cdot |\Delta_K|^{1/(2n)}) & \text{otherwise} \end{cases}$$

Proof. The difference between the upper bounds of $B^{1/n}$ for different types of number fields depends on the choice of the sublattice $C \subseteq \Lambda_K$ of the logarithmic unit lattice. Because $1/s = \max(\sqrt{2}\eta_1(C^*), \log(n)^2)$, the product $s^{-\mathbb{r}} \cdot [\Lambda_K : C]$ is the only part of B that depends on the choice of this sublattice.

For general number fields, we will choose $C = \Lambda_K$, and use a general upper bound $\eta_1(\Lambda_K^*) \leq O(n(\log n)^3)$ due to Kessler and Dobrowolski [Kes91; Dob79] to obtain $s^{-\mathbb{r}} \cdot [\Lambda_K : C] \leq O(n^{\mathbb{r}} \log(n)^{3\mathbb{r}})$.

For cyclotomic number fields with prime power conductor, we choose $C \subseteq \Lambda_K$ to be the sublattice of Λ_K consisting of the logarithmic image of the cyclotomic units [Was12, Ch. 8]. For this sublattice it is known that $[\Lambda_K : C] = h_K^+$, the class number of the maximal totally real subfield of K, and $\eta_1(C^*) \leq O(1)$, so that $s^{-\mathbb{r}} \cdot [\Lambda_K : C] \leq O(\log(n)^{2\mathbb{r}} \cdot h_K^+) = \log(n)^{O(n)}$ for these prime power cyclotomic number fields, under the assumption that $h_K^+ = \log(n)^{O(n)}$. The precise derivation of these bounds follow later in this proof.

Plugging these bounds into the value of B in Theorem 5.9, using $\mathbb{r} = n - n_{\mathbb{C}} - 1 \leq n$, $|\operatorname{Pic}_K^0|^{1/n} = \widetilde{O}(|\Delta_K|^{1/(2n)})$ (see Lemma 2.17), $|\Delta_K|^{1/(2n)} \leq \sqrt{n}$

for cyclotomic fields, and suppressing polylogarithmic factors, we obtain

$$B^{1/n} = \widetilde{O}\Big(\underbrace{\overbrace{s^{-\mathbb{r}/n}}^{=\widetilde{O}(n^{\mathbb{r}/n})\text{ for general number fields}} \cdot [\Lambda_K : C]^{1/n}}_{=\log(n)^{O(1)}\text{ for prime power cyclotomic fields}} \cdot \overbrace{|\operatorname{Pic}_K^0|^{1/n}}^{\widetilde{O}(|\Delta_K|^{1/(2n)})} \cdot \overbrace{(\log|\Delta_K|)^{2/n}}^{\text{polylog. factor}} \Big)$$

$$= \begin{cases} \widetilde{O}(\sqrt{n}) & \text{for prime power cyclotomic fields,} \\ & \text{assuming that } h_K^+ = \log(n)^{O(n)} \\ \widetilde{O}(n^{1-n_{\mathbb{C}}/n} \cdot |\Delta_K|^{1/(2n)}) & \text{for general number fields} \end{cases}$$

General number fields. We take $C = \Lambda_K$, so that $[\Lambda_K : C] = 1$. By the fact that $\eta_1(\Lambda_K^*) \leq \frac{\sqrt{\mathbb{r}}}{\lambda_1(\Lambda_K)}$ [MR07, Lm. 3.2] and by the general upper bound $1/\lambda_1(\Lambda_K) \leq 1000\sqrt{\mathbb{r}+1}\log(\mathbb{r})^3$ [Kes91; Dob79], we obtain $\eta_1(\Lambda_K^*) \leq \sqrt{\mathbb{r}}/\lambda_1(\Lambda_K) \leq 2000 \cdot \mathbb{r} \cdot \log(\mathbb{r})^3$. Therefore, since $1/s = \max(\sqrt{2} \cdot \eta_1(C^*), (\log n)^2)$,

$$s^{-\mathbb{r}} \cdot [\Lambda_K : C] \leq O(n^{\mathbb{r}} \log(n)^{3\mathbb{r}}) \text{ for general number fields } K$$

Prime power cyclotomic number fields. We take C to be the logarithmic image of the *group of cyclotomic units*, which are units that have a specific compact shape [Was12, Ch. 8] . For this *logarithmic cyclotomic unit lattice* $C \subseteq \Lambda_K$, holds $[\Lambda_K : C] = h_K^+$, the class number of the maximal real field in the cyclotomic field K [Was12, Thm. 8.2]. Due to a result of Cramer et al. [Cra+16, Thm. 3.1] we have an upper bound on the last successive minimum $\lambda_{\mathbb{r}}(C^*)$ of the dual logarithmic cyclotomic unit lattice. Combined with a general smoothing parameter bound for lattices [MR07, Lm. 3.3], this yields the following bound on the smoothing parameter of the dual logarithmic cyclotomic unit lattice: $\eta_1(C^*) \leq \log(4\mathbb{r})\lambda_{\mathbb{r}}(C^*) \leq O(\log(\mathbb{r})^{5/2} \cdot \mathbb{r}^{-1/2}) = O(1)$. Therefore, with the instantiation $1/s = \max(\sqrt{2} \cdot \eta_1(C^*), (\log n)^2)$,

$$s^{-\mathbb{r}} \cdot [\Lambda_K : C] \leq O(\log(n)^{2\mathbb{r}} \cdot h_K^+) \text{ for prime power cyclotomic fields } K.$$

□

5.5. Discretizing the Reduction Algorithm

5.5.1. Introduction

In the reduction algorithm of Section 5.4 (see Algorithm 4), we saw that the random walk procedure is inherently continuous, due to its continuous Gaussian walk. On top of that, the computation of the distribution representation $\mathcal{D}$ also has a continuous aspect, namely the sampling of a vector on a large circle $\mathcal{C}_M$.

The purpose of this section is to show that the result of applying the random walk procedure and the distribution representation using only *finite precision* doesn't differ too much from the result when one would use infinite precision instead. In other words, actually computing the random walk and the distribution on a finite machine (as in Algorithm 6 and Algorithm 5) doesn't spoil the end result. In particular, none of the operations in this section involves real numbers; it is all floating point arithmetic.

Additionally, this section also provides an upper bound on the running time of this discretized reduction algorithm.

We define $\mathcal{D}_{\mathcal{W}(\mathrm{Pic}_K^0)+x\mathfrak{a}}$ by the distribution of $\mathfrak{d}^{-1}$ in step 5 of Algorithm 4, and $\ddot{\mathcal{D}}_{\ddot{\mathcal{W}}(\mathrm{Pic}_K^0)+x\mathfrak{a}}$ by the distribution of $\mathfrak{d}^{-1}$ in step 6 of Algorithm 6. A precise description of these distributions for the case $x\mathfrak{a} = \mathcal{O}_K$ can be found in Definition 5.14 and Definition 5.16, respectively. These distributions are only being described for the case $x\mathfrak{a} = \mathcal{O}_K$, as the general case is a mere translation of this base case. Note the dots above $\ddot{\mathcal{D}}$ and $\ddot{\mathcal{W}}$ to indicate discreteness.

Theorem 5.11. *Let $x\mathfrak{a} \in \mathrm{IdLat}_K^0$ be a norm-one ideal lattice, where $\mathfrak{a}$ is represented by a finite-precision matrix $M_\mathfrak{a}$ and $x \in K_\mathbb{R}$ is represented by a finite-precision vector. Then, Algorithm 6 approximates the distribution of Algorithm 4 within a total variation distance of $23 \cdot \varepsilon$, i.e.,*

$$\|\mathcal{D}_{\mathcal{W}(\mathrm{Pic}_K^0)+x\mathfrak{a}} - \ddot{\mathcal{D}}_{\ddot{\mathcal{W}}(\mathrm{Pic}_K^0)+x\mathfrak{a}}\| \leq 23 \cdot \varepsilon,$$

and runs within time polynomial in $\log|\Delta_K|, \mathrm{size}(x), \mathrm{size}(M_{\mathfrak{a}})$ *(see Section 2.1) and* $\log(1/\varepsilon)$.

Roadmap of the proof

Introduction. In this proof, we show that the the random walk distribution $\mathcal{D}_{\mathcal{W}(\mathrm{Pic}^0_K)+x\mathfrak{a}}$ from $\mathfrak{d}^{-1}$ in line 5 of Algorithm 4 and the *discretized* random walk distribution $\ddot{\mathcal{D}}_{\ddot{\mathcal{W}}(\mathrm{Pic}^0_K)+x\mathfrak{a}}$ from $\mathfrak{d}^{-1}$ in line 5 of Algorithm 6 are close to each other in the total variation distance.

In the proof we will, without loss of generality, assume that $x\mathfrak{a} = \mathcal{O}_K$. The case of general $x\mathfrak{a}$ consists of a mere translation of the distributions involved and does not affect the proof structure. Therefore, we resort to proving closeness of $\mathcal{D}_{\mathcal{W}(\mathrm{Pic}^0_K)}$ and $\ddot{\mathcal{D}}_{\ddot{\mathcal{W}}(\mathrm{Pic}^0_K)}$.

The proof of closeness in total variation distance proceeds by two steps; the first step discretizes the continuous Gaussian sampling in the reduction Algorithm 4, whereas the second step discretizes the uniform sampling on the M-circle in the distribution Algorithm 3 which is used in the reduction.

Sampling the Gaussian walk in a discrete manner doesn't spoil the resulting distribution. In the random walk procedure, a Gaussian distribution is sampled in the logarithmic unit lattice ambient vector space and subsequently exponentiated component-wise to act on the processed input ideal lattice. This part is referred to as the 'continuous walk' of the random walk procedure. A finite computer cannot sample from continuous distributions, so in the actual algorithmic implementation a *discrete Gaussian* is sampled on a sufficiently fine grid – a lattice – on the ambient vector space.

The discrete random walk distribution resulting from sampling the Gaussian walk in this discrete way, whereas keeping the rest of the random walk procedure the same, is what we will call $\ddot{\mathcal{W}}(\mathrm{Pic}^0_K)$. By a technical computation, we will show that $\mathcal{D}_{\mathcal{W}(\mathrm{Pic}^0_K)} \approx \mathcal{D}_{\ddot{\mathcal{W}}(\mathrm{Pic}^0_K)}$.

Sampling from a discrete circle doesn't change the map $\mathcal{D}$ much. In the beginning of the distribution representation $\mathcal{D}$, a vector is uniformly sampled from a M-circle $\mathcal{C}_M$ in $K_\mathbb{R}$. In reality, on a finite computer, we need to sample from this M-circle in a discrete manner, while keeping the rest of the distribution computation the same. This particular map is called $\ddot{\mathcal{D}} : \mathrm{Pic}_K^0 \to L_1(\mathcal{I}_K)$.

By showing that $\ddot{\mathcal{D}}$ and $\mathcal{D}$ are close for any $\mathbf{a} \in \mathrm{Pic}_K^0$, we draw the conclusion that for any distribution $\mathcal{P}$ on Pic_K^0, $\ddot{\mathcal{D}}_\mathcal{P}$ and $\mathcal{D}_\mathcal{P}$ are close as well. In particular, $\ddot{\mathcal{D}}_{\mathcal{W}(\mathrm{Pic}_K^0)} \approx \mathcal{D}_{\mathcal{W}(\mathrm{Pic}_K^0)}$.

Finalizing. By using the above two parts, we can show that the following three distributions are actually close.

$$\mathcal{D}_{\mathcal{W}(\mathrm{Pic}_K^0)} \underbrace{\approx}_{\text{First part}} \mathcal{D}_{\ddot{\mathcal{W}}(\mathrm{Pic}_K^0)} \underbrace{\approx}_{\text{Second part}} \ddot{\mathcal{D}}_{\ddot{\mathcal{W}}(\mathrm{Pic}_K^0)}$$

By observing that the latter distribution can actually be computed by a classical finite machine, we finish the proof.

5.5.2. Precise Definition of the Distributions $\mathcal{D}_{\mathcal{W}(\mathrm{Pic}_K^0)}$, $\mathcal{D}_{\ddot{\mathcal{W}}(\mathrm{Pic}_K^0)}$ and $\ddot{\mathcal{D}}_{\ddot{\mathcal{W}}(\mathrm{Pic}_K^0)}$

Before defining the three relevant distributions, we first need to define the discretization of the Gaussian (in the random walk procedure) and of the circle (in the distribution procedure). The discretization of the continuous Gaussian happens by sampling a discrete Gaussian on a square grid of the log-unit hyperplane and the discretization of the hypercircle happens by taking equidistant points on this hypercircle.

Definition 5.12 (Orthogonal lattice in the log-unit hyperplane H)**.** *By choosing an orthonormal basis $(\mathbf{b}_1, \ldots, \mathbf{b}_\mathbb{r})$ of the $\mathbb{r}$-dimensional vector space $H = \{(x_\sigma)_\sigma \in \log K_\mathbb{R} \mid \sum_\sigma x_\sigma = 0\}$, we define $\mathbb{Z}_H = \mathbf{b}_1\mathbb{Z} + \ldots + \mathbf{b}_\mathbb{r}\mathbb{Z}$.*

The actual choice of the orthonormal basis doesn't matter in the proofs, so we will just work with the lattice $\mathbb{Z}_H \subseteq H$ and leaving the basis choice implicit. For $D \in \mathbb{N}_{>0}$, we denote $\frac{1}{D}\mathbb{Z}_H$ for the scaling of $\mathbb{Z}_H$ by $\frac{1}{D}$, i.e., $\frac{1}{D}\mathbb{Z}_H = \frac{1}{D}\cdot(\mathbf{b}_1\mathbb{Z}+\ldots+\mathbf{b}_r\mathbb{Z})$. To make the random walk procedure efficiently computable on a finite machine, we discretize the continuous Gaussian walk over H by sampling from a discrete Gaussian over $\frac{1}{D}\mathbb{Z}_H$.

Definition 5.13 (Sampling in the finite set $\ddot{\mathcal{C}}_M \subseteq \mathcal{C}_M \in K_{\mathbb{R}}$)**.** *For a small discretization parameter $\varepsilon > 0$, we put $k = \sqrt{n} \cdot M \cdot \lceil 1/\varepsilon \rceil$,*

$$\ddot{\mathcal{C}}_M^{(\varepsilon)} = \{(x_\sigma)_\sigma \in \mathcal{C}_M \mid x_\sigma = \pm M e^{2\pi i j/k} \text{ for some } j \in \mathbb{N} \}.$$

Recall that for real embeddings σ we have $x_\sigma = \pm M$, and for complex embeddings $x_{\bar{\sigma}} = \overline{x_\sigma}$, due to the fact that $\mathcal{C}_M \subseteq K_{\mathbb{R}}$. We often suppress the notation of ε in $\ddot{\mathcal{C}}_M$.

For most purposes, the precise definition of $\ddot{\mathcal{C}}_M^{(\varepsilon)}$ is not so important; what matters more is the fact that any point in $\mathcal{C}_M$ is ε-close to $\ddot{\mathcal{C}}_M^{(\varepsilon)}$ (see Figure 5.3).

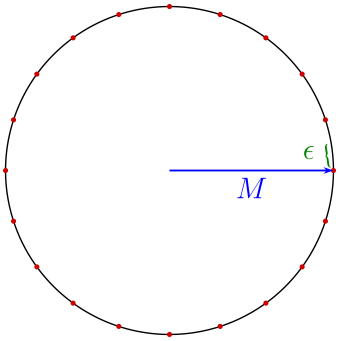

Figure 5.3.: Any point on the circle $\mathcal{C}_M$ is ε-close to the red discretized circle $\ddot{\mathcal{C}}_M$.

Now we are ready to rigorously define the three distributions involved. We start with the distribution involving a continuous Gaussian and a continuous circle, Definition 5.14. The algorithm associated with this distribution is Algorithm 4, with Algorithm 3 as a subroutine.

Then we proceed by the definition of a intermediate distribution, which has a discrete Gaussian sampling in the random walk procedure, but still has a continuous sample from the circle in the distribution procedure, see Definition 5.15. The difference between these two distributions is marked with the color blue. The algorithm associated with this distribution is Algorithm 6, still with Algorithm 3 as a subroutine. Also in this algorithm description, the differences are marked in the color blue.

The last distribution is the one that can be run on a finite computer and has both a discretized Gaussian and a discretized circle, see Definition 5.16. The difference between this distribution and the intermediate distribution is marked with the color green. The algorithm associated with this distribution is Algorithm 6, with the discrete sampling on the circle Algorithm 5 as a subroutine, where the differences with the original (Algorithm 3) is marked with the color green as well.

Definition 5.14 (Continuous Gaussian, continuous circle)**.** *Denoting* $P_B = \{\mathfrak{p} \text{ prime ideal of } \mathcal{O}_K \mid \mathcal{N}(\mathfrak{p}) \leq B\}$*, the output distribution* $\mathcal{D}_{\mathcal{W}(\mathrm{Pic}_K^0)}$ *of Algorithm 4 can be described by the following rule, for any integral ideal* $\mathfrak{d} \in \mathcal{I}_K$*.*

$$\mathcal{D}_{\mathcal{W}(\mathrm{Pic}_K^0)}[\mathfrak{d}^{-1}] = \frac{1}{|P_B| \cdot \mathrm{Vol}(\mathcal{C}_M)} \int_{c \in \mathcal{C}_M} \sum_{\mathfrak{p} \in P_B} \int_{y \in H} \sum_{\substack{v \in \mathfrak{p} \\ (v) = \mathfrak{p}\mathfrak{d}}} \frac{\rho_\varsigma(e^y v/p - c)}{\rho_\varsigma(e^y \mathfrak{p}/p - c)} s^{-\mathbb{r}} \rho_s(y) dy,$$

where $p = \mathcal{N}(\mathfrak{p})^{-1/n}$*.*

Definition 5.15 (Discrete Gaussian, continuous circle)**.** *Denoting* $P_B = \{\mathfrak{p} \text{ prime ideal of } \mathcal{O}_K \mid \mathcal{N}(\mathfrak{p}) \leq B\}$*, the output distribution* $\mathcal{D}_{\ddot{\mathcal{W}}(\mathrm{Pic}_K^0)}$*, where the continuous Gaussian* $\mathcal{G}_{H,s}$ *in Algorithm 4 is replaced by a discrete Gaussian* $\mathcal{G}_{\frac{1}{D}\mathbb{Z}_H, s}$*, can be described by the following rule, for any integral ideal* $\mathfrak{d} \in \mathcal{I}_K$*.*

$$\mathcal{D}_{\ddot{\mathcal{W}}(\mathrm{Pic}_K^0)}[\mathfrak{d}^{-1}] =$$

$$\frac{1}{|P_B| \cdot \mathrm{Vol}(\mathcal{C}_M)} \int_{c \in \mathcal{C}_M} \sum_{\mathfrak{p} \in P_B} \sum_{\ddot{y} \in \frac{1}{D}\mathbb{Z}_H} \sum_{\substack{v \in \mathfrak{p} \\ (v) = \mathfrak{p}\mathfrak{d}}} \frac{\rho_\varsigma(\lceil e^{\ddot{y}} \rfloor v/p - c)}{\rho_\varsigma(\lceil e^{\ddot{y}} \rfloor \mathfrak{p}/p - c)} \frac{\rho_s(\ddot{y})}{\rho_s(\frac{1}{D}\mathbb{Z}_H)},$$

where $\lceil e^{\ddot{y}} \rfloor$ *means that* $e^{\ddot{y}}$ *is computed with* $\lceil \log_2 D \rceil$ *bits of precision in all coordinates, and where* $p = \mathcal{N}(\mathfrak{p})^{-1/n}$.

Definition 5.16 (Discrete Gaussian, discrete circle)**.** *Denoting* $P_B = \{\mathfrak{p} \text{ prime ideal of } \mathcal{O}_K \mid \mathcal{N}(\mathfrak{p}) \leq B\}$, *the output distribution* $\mathcal{D}_{\ddot{\mathcal{W}}(\mathrm{Pic}^0_K)}$, *where the continuous Gaussian* $\mathcal{G}_{H,s}$ *in Algorithm 4 is replaced by a discrete Gaussian* $\mathcal{G}_{\frac{1}{D}\mathbb{Z}_H,s}$, *and the continuous uniform distribution on* $\mathcal{C}_M$ *in* $\mathcal{D} : \mathrm{Pic}^0_K \to L_1(\mathcal{I}_K)$ *is replaced by a uniform distribution over the finite set* $\ddot{\mathcal{C}}_M$ *can be described by the following rule, for any integral ideal* $\mathfrak{d} \in \mathcal{I}_K$.

$$\ddot{\mathcal{D}}_{\ddot{\mathcal{W}}(\mathrm{Pic}^0_K)}[\mathfrak{d}^{-1}] = \frac{1}{|P_B| \cdot |\ddot{\mathcal{C}}_M|} \sum_{\ddot{c} \in \ddot{\mathcal{C}}_M} \sum_{\mathfrak{p} \in P_B} \sum_{\ddot{y} \in \frac{1}{D}\mathbb{Z}_H} \sum_{\substack{v \in \mathfrak{p} \\ (v)=\mathfrak{p}\mathfrak{d}}} \frac{\rho_\varsigma(\lceil e^{\ddot{y}} \rfloor v/p - \ddot{c})}{\rho_\varsigma(\lceil e^{\ddot{y}} \rfloor \mathfrak{p}/p - \ddot{c})} \frac{\rho_s(\ddot{y})}{\rho_s(\frac{1}{D}\mathbb{Z}_H)},$$

where $\lceil e^{\ddot{y}} \rfloor$ *means that* $e^{\ddot{y}}$ *is computed with* $\lceil \log_2 D \rceil$ *bits of precision in all coordinates, and where* $p = \mathcal{N}(\mathfrak{p})^{-1/n}$.

5.5.3. Discretized Algorithm Analogues

In the following text we treat the discrete analogues of Algorithm 4 and Algorithm 3. We show that these discretized algorithms (Algorithm 6 and Algorithm 5) run in polynomial time with respect to the input size and that their output distribution does not differ significantly from their continuous counterparts.

We start with defining the algorithms and showing that they run in polynomial time. The remainder of this chapter, Section 5.5.4, is devoted to showing that the discretized and non-discretized algorithms indeed yield almost the same distribution.

Lemma 5.17. *Algorithm 5 is correct and runs within time polynomial in* $\log|\Delta_K|$, $\mathrm{size}(M_\mathfrak{a})$, $\log(1/\varepsilon)$ *and* $\mathrm{size}(x)$.

Proof. The input of this algorithm is given by the vector $x \in K_\mathbb{R}$ (given in a finite precision representation) and a basis matrix $B_\mathfrak{a}$ of the ideal $\mathfrak{a}$.

Algorithm 5: Sampling efficiently from a distribution very close to $\mathcal{D}_{x\mathfrak{a}}$, discretized

Require: A pair $(x, \mathfrak{a}) \in K_{\mathbb{R}} \times \mathcal{I}_K$ such that $\mathcal{N}(\mathfrak{a}) \prod_\sigma x_\sigma = 1$.

Ensure: An sample from a distribution (14ε)-close to the distribution $\mathcal{D}_{x\mathfrak{a}}$ in the total variation distance.

1: Put $\varsigma = 2^{n+1} \cdot n \cdot |\Delta_K|$ and $M = 2\sqrt{n} \cdot \varsigma$.

2: Sample a center $\ddot{c} = (\ddot{c}_\sigma)_\sigma$ uniformly in the finite subset $\ddot{\mathcal{C}}_M := \ddot{\mathcal{C}}_M^{(\varepsilon/n)} \subseteq \mathcal{C}_M = \{(y_\sigma)_\sigma \mid |y_\sigma| = M \text{ for all embeddings } \sigma\}$. Where $\ddot{\mathcal{C}}_M$ is such that any point in $\mathcal{C}_M$ is ε/n-close to $\ddot{\mathcal{C}}_M$ (see Definition 5.13)

3: Sample from the discrete Gaussian $\mathcal{G}_{x\mathfrak{a},\varsigma,\ddot{c}}$ with respect to the ideal lattice $x\mathfrak{a}$ with center $\ddot{c} = (\ddot{c}_\sigma)_\sigma$ and standard deviation ς, leading to some $v \in x\mathfrak{a}$.

4: **return** the inverse integral ideal $\mathfrak{d}^{-1} = v^{-1} x\mathfrak{a} \in \mathcal{I}_K$

We denote with $\text{size}(x)$ the number of bits needed to represent $x \in K_{\mathbb{R}}$ and with $\text{size}(B_\mathfrak{a})$ the number of bits needed to represent the basis $B_\mathfrak{a}$ (see Section 2.1).

We go through the lines of Algorithm 5 to examine the running time. Line 1 can clearly be done in linear time in $\log(|\Delta_K|)$ and n. Line 2 samples from in set $\ddot{\mathcal{C}}_M^{(\varepsilon)}$, which are essentially at most[5] $n/2$ independent samples of the discretized circle $\{Me^{2\pi ij/D} \mid j \in \mathbb{N}\}$, with $D = \sqrt{n}M\lceil 1/\varepsilon \rceil$. One such sample takes time linear in $\log M = O(\log |\Delta_K|)$ and $\log(1/\varepsilon)$, so a sample from $\ddot{\mathcal{C}}_M^{(\varepsilon)}$ costs $O(n(\log |\Delta_K| + \log(1/\varepsilon)))$. Line 3 uses Klein's algorithm [GPV08; Kle00] to sample from the discrete Gaussian $\mathcal{G}_{x\mathfrak{a},\varsigma,\ddot{c}}$, which runs in time polynomial in $\text{size}(B_\mathfrak{a})$ and $\text{size}(x)$, by an adaptation of [GPV08, Thm. 4.1] for an exponentially small statistical distance. An additional property of Klein's algorithm is that the output $v \leftarrow \mathcal{G}_{x\mathfrak{a},\varsigma,\ddot{c}}$ is actually polynomially bounded by $\text{size}(x)$ and $\text{size}(B_\mathfrak{a})$. The last line, line 4, uses ideal division and multiplication, which (naively) takes the time to solve a system of equations involving a $n^2 \times n^2$ matrix (see [Coh99, §4.8.4]) having

[5] At most half of n, because of the complex conjugate embeddings

entry sizes comparable to that of size($B_\mathfrak{a}$) and size(x); therefore this can be done within polynomial time in $\log|\Delta_K|$, size($B_\mathfrak{a}$) and size(x). As all lines can be computed in polynomial time of size($B_\mathfrak{a}$), size(x) and $\log|\Delta_K|$, the result follows.

The correctness is proven later, in Lemma 5.20. □

Lemma 5.18. *Algorithm 6 is correct and runs within time polynomial in* $\log|\Delta_K|$, size($M_\mathfrak{a}$), $\log[\Lambda_K : C]$ *and* size(x), *and uses one call to a* γ*-Hermite SVP oracle.*

Proof. The input of this algorithm is given by the vector $x \in K_\mathbb{R}$ (given in a finite precision representation) and a basis matrix $B_\mathfrak{a}$ of the ideal $\mathfrak{a}$. We denote with size(x) the number of bits needed to represent $x \in K_\mathbb{R}$ and with size($B_\mathfrak{a}$) the number of bits needed to represent the basis $B_\mathfrak{a}$.

Since $\log|\operatorname{Pic}_K^0| = O(\log|\Delta_K|)$ (see Lemma 2.17), $n = O(\log|\Delta_K|)$ and $\eta_1(C^*) \leq \eta_1(\Lambda_K^*) \leq \widetilde{O}(n)$ (see the proof of Proposition 5.10) the quantity $\log B$ is polynomially bounded in $\log|\Delta_K|$ and $\log[\Lambda_K : C]$. Similarly, $\log D$, the logarithm of the discretization parameter of the Gaussian, is polynomially bounded by $\log|\Delta_K|$ and $\log(\varepsilon^{-1})$.

We go through all steps of Algorithm 6 to estimate the running time. Step 1 of Algorithm 6 runs within time quasi-linear in $\log B$. Step 2 involves the sampling a random prime ideal $\mathfrak{p}$ and the multiplication of ideals $\mathfrak{a}$ and $\mathfrak{p}$. The random sampling can be done within polynomial time (see Lemma 2.14). The product $\mathfrak{pa}$ can be computed by reducing the $n^2 \times n$ matrix consisting of the products of the respective $\mathbb{Z}$-generators of $\mathfrak{a}$ and $\mathfrak{p}$ which runs in time polynomial in n, size($\mathfrak{a}$) and $\log B$ (where B is the maximum norm of $\mathfrak{p}$). Step 3 consists of discrete Gaussian sampling in the lattice $\frac{1}{D}\mathbb{Z}_H$ with standard deviation s satisfying $\widetilde{O}(n) \leq 1/s \leq \log(n)^2$. An adaptation of [GPV08, Thm. 4.1] shows that this can be done in time polynomially bounded by $\log D$ and n, i.e. bounded by $\log|\Delta_K|$ and $\log(\varepsilon^{-1})$. An additional property of this sampling is that the output is polynomially bounded as well. Step 4 is just rescaling, which has no serious impact on

Algorithm 6: The worst-case to average-case reduction algorithm, discretized

Require:

- A pair $(x, \mathfrak{a}) \in K_{\mathbb{R}} \times \mathcal{I}_K$ satisfying $\mathcal{N}(\mathfrak{a}) \prod_\sigma x_\sigma = 1$.
- The values $[\Lambda_K : C]$ and $\eta_1(C^*)$ of a suitable sublattice $C \subseteq \Lambda_K$ of the logarithmic unit lattice,
- An oracle $\mathcal{A}$ that solves γ-Hermite SVP in $\mathfrak{d}^{-1}$ whenever $\mathfrak{d}^{-1} \leftarrow \mathcal{D}_{\mathcal{U}(\mathrm{Pic}_K^0)}$.
- An error parameter $\varepsilon > 0$

Ensure: A vector $\alpha \in x\mathfrak{a}$ that is a solution to $B^{1/n} \cdot \gamma$-Hermite SVP in the ideal lattice $x\mathfrak{a}$, i.e.,

$$\|\alpha\| \leq \gamma \cdot B^{1/n} \cdot \det(x\mathfrak{a})^{1/n},$$

where $B = \widetilde{O}(4^n \cdot \eta_1(C^*)^{\mathrm{r}} \cdot [\Lambda_K : C] \cdot |\operatorname{Pic}_K^0| \cdot (\log|\Delta_K|)^2)$, or, *failure*.

1: Put $s = \max(\sqrt{2} \cdot \eta_1(C^*), (\log n)^2)$ and $B = \widetilde{O}(4^n \cdot \eta_1(C^*)^{\mathrm{r}} \cdot [\Lambda_K : C] \cdot |\operatorname{Pic}_K^0| \cdot (\log|\Delta_K|)^2)$ as in Corollary 5.8.

2: Multiply the ideal $\mathfrak{a}$ by a prime ideal $\mathfrak{p}$ uniformly sampled from the set $\{\mathfrak{p} \text{ prime ideal of } \mathcal{O}_K \mid \mathcal{N}(\mathfrak{p}) \leq B\}$, yielding $\mathfrak{ap}$.

3: Sample $\ddot{y} \leftarrow \mathcal{G}_{s, \frac{1}{D}\mathbb{Z}_H}$, where $D = 2^{n+2} \cdot n^4 \cdot \lceil |\Delta_K| \cdot \varepsilon^{-1} \rceil$ and $\mathbb{Z}_H$ is an orthonormal basis of the hyperplane H where the logarithmic unit lattice lives in (see Definition 5.12).

4: Put $p = \mathcal{N}(\mathfrak{p})^{1/n}$, so that $e^y x\mathfrak{ap}/p$ has norm 1, where $e^y \in K_{\mathbb{R}}$ is the component-wise exponentiation of $y \in H$.

5: Sample $\mathfrak{d}^{-1} \leftarrow \mathcal{D}_{(\lfloor e^{\ddot{y}} \rceil \cdot x/p, \mathfrak{ap})}$ using Algorithm 5, where $\lfloor e^{\ddot{y}} \rceil \in K_{\mathbb{R}}$ is the component-wise exponentiation of $\ddot{y} \in H$, computed with $\lceil \log_2 D \rceil$ bits of precision in all coordinates. Furthermore, let $v \in e^y x\mathfrak{ap}/p$ be the additional output of Algorithm 5 that satisfies $\mathfrak{d}^{-1} = v^{-1} \lfloor e^{\ddot{y}} \rceil x\mathfrak{ap}/p$.

6: Invoke the γ-Hermite SVP oracle $\mathcal{A}$ on $\mathcal{D}_{\mathcal{U}(\mathrm{Pic}_K^0)}$ to find a $\kappa \in \mathfrak{d}^{-1}$ for which holds $\|\kappa\| \leq \gamma \cdot \det(\mathfrak{d}^{-1})^{1/n}$

7: **return** $p \cdot (\lfloor e^{\ddot{y}} \rceil)^{-1} \cdot v \cdot \kappa \in x\mathfrak{a}$.

the running time. Step 5 uses Algorithm 5, which runs in time polynomially bounded by size($\lfloor e^{\ddot{y}} \rceil \cdot x/p$), size($M_{\mathfrak{a}}$) and $\log B$. As size($\lfloor e^{\ddot{y}} \rceil \cdot x/p$) can be linearly bounded by $\log B$, size(x) and $\log D$ (because $\lfloor e^{\ddot{y}} \rceil$ is computed with relative bit precision $\log_2 D$), this step is polynomially bounded as well in $\log |\Delta_K|$, $\log \varepsilon^{-1}$ and size(x). Step 6 invokes the γ-Hermite SVP oracle once. Step 7 just rescales the element $\kappa \in \mathfrak{d}^{-1}$ without a serious impact on the running time.

Later in this section we prove two closeness lemmas, namely Lemma 5.19 and Lemma 5.20. From those two lemmas, one obtains the desired closeness of distributions of the sampling mechanism of $\mathfrak{d}^{-1}$; this proves the correctness. □

5.5.4. Closeness Proofs

Sampling the Gaussian walk in a discrete manner doesn't spoil the resulting distribution

Lemma 5.19. *Let K be a number field and let $1 > s > 0$ be a given Gaussian spread parameter for the continuous part of the random walk, let $\varepsilon > 0$ be a given error parameter and let $M = 2 \cdot n^{3/2} \cdot 2^{n+1} \cdot |\Delta_K|$ as in Algorithm 5. Let $\frac{1}{D}\mathbb{Z}_H \subseteq H$ be the discretization of the Log-unit space to compute the discrete Gaussian analogue of the continuous part of the random walk, with $D \in \mathbb{N}$ such that $D \geq \lceil (4 \cdot s^{-2}\sqrt{n} + 100 \cdot n^2 M) \cdot 1/\varepsilon \rceil$.*

Then

$$\| \mathcal{D}_{\ddot{\mathcal{W}}(\mathrm{Pic}^0_K)} - \mathcal{D}_{\mathcal{W}(\mathrm{Pic}^0_K)} \|_1 \leq 18 \cdot \varepsilon$$

Proof. Examining the definitions of the distributions $\mathcal{D}_{\mathcal{W}(\mathrm{Pic}^0_K)}$ and $\mathcal{D}_{\ddot{\mathcal{W}}(\mathrm{Pic}^0_K)}$ (see Definitions 5.14 and 5.15), we can apply the triangle inequality and a

norm inequality, to directly deduce

$$\|\mathcal{D}_{\mathcal{W}(\mathrm{Pic}_K^0)} - \mathcal{D}_{\ddot{\mathcal{W}}(\mathrm{Pic}_K^0)}\|_1 \tag{5.70}$$
$$\leq \max_{\substack{c\in\mathcal{C}_M \\ \mathfrak{p}\in P_B}} \underbrace{\sum_{\mathfrak{d}} \sum_{\substack{v\in\mathfrak{p} \\ (v)=\mathfrak{p}\mathfrak{d}}} \left| \int_{y\in H} \frac{\rho_\varsigma(e^y v/p - c)}{\rho_\varsigma(e^y\mathfrak{p}/p - c)} s^{-\mathbb{r}}\rho_s(y)dy - \sum_{\ddot{y}\in\frac{1}{D}\mathbb{Z}_H} \frac{\rho_\varsigma(\lfloor e^{\ddot{y}}\rceil v/p - c)}{\rho_\varsigma(\lfloor e^{\ddot{y}}\rceil\mathfrak{p}/p - c)} \frac{\rho_s(\ddot{y})}{\rho_s(\frac{1}{D}\mathbb{Z}_H)} \right|}.$$

Therefore, we can focus on the quantity in the bracket of Equation (5.70) for a fixed prime ideal $\mathfrak{p} \in P_B$ and a fixed center $c \in \mathcal{C}_M$ from the M-circle. We rewrite the term within the absolute value signs by using a block tiling of the orthonormal lattice $\frac{1}{D}\mathbb{Z}_H \subseteq H$ (see Definition 5.12) with fundamental domain F_H satisfying $\mathrm{Vol}(F_H) = D^{-\mathbb{r}}$. Observing that we can collapse the summation $\sum_{\mathfrak{d}} \sum_{\substack{v\in\mathfrak{p} \\ (v)=\mathfrak{p}\mathfrak{d}}}$ to $\sum_{v\in\mathfrak{p}}$ (as the sum with $\mathfrak{d}$ is over integral ideals), we obtain that the quantity in the bracket of Equation (5.70) is at most

$$\sum_{v\in\mathfrak{p}} \left| \sum_{\ddot{y}\in\frac{1}{D}\mathbb{Z}_H} \int_{y\in\ddot{y}+F_H} \underbrace{\frac{\rho_\varsigma(e^y v/p - c)}{\rho_\varsigma(e^y\mathfrak{p}/p - c)}}_{A} \underbrace{s^{-\mathbb{r}}\rho_s(y)}_{B} - \underbrace{\frac{\rho_\varsigma(\lfloor e^{\ddot{y}}\rceil v/p - c)}{\rho_\varsigma(\lfloor e^{\ddot{y}}\rceil\mathfrak{p}/p - c)}}_{A'} \underbrace{\frac{D^{\mathbb{r}}\rho_s(\ddot{y})}{\rho_s(\frac{1}{D}\mathbb{Z}_H)}}_{B'} dy \right|$$

Applying the triangle inequality, switching integrals and sums, and using the identity $AB - A'B' = B(A - A') + (B - B')A'$, above equation is at most

$$\sum_{\ddot{y}\in\frac{1}{D}\mathbb{Z}_H} \int_{y\in\ddot{y}+F_H} s^{-\mathbb{r}}\rho_s(y) \underbrace{\sum_{v\in\mathfrak{p}} \left| \frac{\rho_\varsigma(e^y v/p - c)}{\rho_\varsigma(e^y\mathfrak{p}/p - c)} - \frac{\rho_\varsigma(\lfloor e^{\ddot{y}}\rceil v/p - c)}{\rho_\varsigma(\lfloor e^{\ddot{y}}\rceil\mathfrak{p}/p - c)} \right|}_{\|\mathcal{G}_{\mathfrak{p}/p,\varsigma/e^y,c} - \mathcal{G}_{\mathfrak{p}/p,\varsigma/\lfloor e^{\ddot{y}}\rceil,c}\|} dy \tag{5.71}$$

$$+ \sum_{\ddot{y}\in\frac{1}{D}\mathbb{Z}_H} \int_{y\in\ddot{y}+F_H} \left| s^{-\mathbb{r}}\rho_s(y) - \frac{D^{\mathbb{r}}\rho_s(\ddot{y})}{\rho_s(\frac{1}{D}\mathbb{Z}_H)} \right| \underbrace{\sum_{v\in\mathfrak{p}} \frac{\rho_\varsigma(e^{\ddot{y}} v/p - c)}{\rho_\varsigma(e^{\ddot{y}}\mathfrak{p}/p - c)}}_{=1} dy \tag{5.72}$$

First part of the sum, Equation (5.71). Apply Lemma A.39 to show that the two Gaussians are reasonably close to each other. Writing $e^{\tilde{y}} = \lfloor e^{\ddot{y}}\rceil$, we have $\|\tilde{y} - \ddot{y}\| \leq \|e^{\tilde{y}-\ddot{y}} - 1\| \leq \frac{\sqrt{n}}{D}$ because $e^{\tilde{y}}$ is the $\log_2(D)$-bit precision

relative approximation of $e^{\ddot{y}}$. By construction, we have $\|y - \ddot{y}\| \leq \frac{\sqrt{n}}{D}$ as well, because $y \in \ddot{y} + F_H$, therefore,

$$\|y - \tilde{y}\| \leq \frac{2\sqrt{n}}{D}.$$

Because $\varsigma > \eta_\varepsilon(x\mathfrak{a})$ for all ideal lattices $x\mathfrak{a} \in \mathrm{IdLat}_K$, we can apply Lemma A.39 with $\delta = \frac{2\sqrt{n}}{D}$. Since $M > \varsigma > 1$, $\|c\| = \sqrt{n} \cdot M$ (because $c \in \mathcal{C}_M$) and $D \geq 100 \cdot M \cdot n^2/\varepsilon$, we obtain

$$\begin{aligned}\|\mathcal{G}_{\mathfrak{p},\varsigma/e^y,c} - \mathcal{G}_{\mathfrak{p},\varsigma/\lfloor e^{\ddot{y}} \rceil,c}\| &\leq 8\varepsilon + 4\pi(\tfrac{1}{\varsigma^2} + n + 2n\|c\|) \cdot \|y - \tilde{y}\| \\ &\leq 8\varepsilon + \frac{100 \cdot M \cdot n^2}{D} = 9 \cdot \varepsilon. \end{aligned} \tag{5.73}$$

Since this bound is independent of $y \in H$ and $\ddot{y} \in \frac{1}{D}\mathbb{Z}_H$, and since

$$\sum_{\ddot{y} \in \frac{1}{D}\mathbb{Z}_H} \int_{y \in \ddot{y} + F_H} s^{-\mathbb{r}} \rho_s(y) dy = 1,$$

we deduce that Equation (5.71) must also be bounded by $9 \cdot \varepsilon$.

Second part of the sum, Equation (5.72). One can apply smoothing arguments; since $s < 1$, we have $s \geq s^2 \geq \frac{\sqrt{n} \cdot \varepsilon^{-1}}{D} \geq \frac{\log(2n(1+\varepsilon^{-1}))}{D} \geq \log(2n(1+\varepsilon^{-1}))\lambda_{\mathbb{r}}(\frac{1}{D}\mathbb{Z}_H) \geq \eta_\varepsilon(\frac{1}{D}\mathbb{Z}_H)$ (see [MR07, Lm. 3.3]). Therefore,

$$s^{-\mathbb{r}} \in (1 - 2\varepsilon, 1 + 2\varepsilon) \cdot \frac{D^{\mathbb{r}}}{\rho_s(\frac{1}{D}\mathbb{Z}_H + y)} \quad \text{for all } y \in H.$$

Putting this into Equation (5.72), we obtain, using Lemma A.37, using the lower bound on D and $\mathrm{Vol}(F_H) = D^{-\mathbb{r}}$,

$$\begin{aligned} &\sum_{\ddot{y} \in \frac{1}{D}\mathbb{Z}_H} \int_{y \in F_H} \left| s^{-\mathbb{r}} \rho_s(\ddot{y} + y) - \frac{D^{\mathbb{r}} \rho_s(\ddot{y})}{\rho_s(\frac{1}{D}\mathbb{Z}_H)} \right| dy \\ &\leq 4\varepsilon + \max_{y \in F_H} \underbrace{\sum_{\ddot{y} \in \frac{1}{D}\mathbb{Z}_H} \left| \frac{\rho_s(\ddot{y} + y)}{\rho_s(\frac{1}{D}\mathbb{Z}_H + y)} - \frac{\rho_s(\ddot{y})}{\rho_s(\frac{1}{D}\mathbb{Z}_H)} \right|}_{\|D_{\frac{1}{D}\mathbb{Z}_H, s, y} - D_{\frac{1}{D}\mathbb{Z}_H, s, 0}\|} \\ &\leq 8\varepsilon + (\tfrac{\pi}{s^2} + 2\pi n) \max_{y \in F_H} \|y\| \leq 8\varepsilon + (\tfrac{\pi}{s^2} + 2\pi n)\frac{\sqrt{n}}{D} \leq 9 \cdot \varepsilon \end{aligned} \tag{5.74}$$

Combining the upper bound on the first and the second part of the sum (see Equations (5.73) and (5.74)), we obtain the result. □

The difference between $\mathcal{D}_{\mathcal{P}(\mathrm{Pic}_K^0)}$ and $\ddot{\mathcal{D}}_{\mathcal{P}(\mathrm{Pic}_K^0)}$, the one with a discretized circle, is negligible for all distributions $\mathcal{P}$

Lemma 5.20. *Let K be a number field and let $1 > \varepsilon > 0$ be a given error parameter. Let $\ddot{\mathcal{C}}_M \subseteq \mathcal{C}_M$ be a discretization of $\mathcal{C}_M$ as in Definition 5.13. Let furthermore $\mathcal{P} \in L_1(\mathrm{Pic}_K^0)$ be any distribution on Pic_K^0 (i.e., $\int_{\mathbf{a}\in\mathrm{Pic}_K^0} \mathcal{P}(\mathbf{a})d\mathbf{a} = 1$). Then we have*

$$\|\mathcal{D}_{\mathcal{P}(\mathrm{Pic}_K^0)} - \ddot{\mathcal{D}}_{\mathcal{P}(\mathrm{Pic}_K^0)}\| \leq 14\varepsilon$$

Proof. The definitions of the two distributions read as follows, for integral ideals $\mathfrak{d} \in \mathcal{I}_K$.

$$\mathcal{D}_{\mathcal{P}(\mathrm{Pic}_K^0)}[\mathfrak{d}^{-1}] = \int_{c\in\mathcal{C}_M} \int_{\mathbf{a}\in\mathrm{Pic}_K^0} \mathcal{P}(\mathbf{a}) \sum_{\substack{v\in\mathbf{a}\\(v)=\mathbf{a}\mathfrak{d}}} \frac{\rho_\varsigma(v-c)}{\rho_\varsigma(\mathbf{a}-c)} d\mathbf{a}dc$$

$$\ddot{\mathcal{D}}_{\mathcal{P}(\mathrm{Pic}_K^0)}[\mathfrak{d}^{-1}] = \frac{1}{|\ddot{\mathcal{C}}_M|} \sum_{\ddot{c}\in\ddot{\mathcal{C}}_M} \int_{\mathbf{a}\in\mathrm{Pic}_K^0} \mathcal{P}(\mathbf{a}) \sum_{\substack{v\in\mathbf{a}\\(v)=\mathbf{a}\mathfrak{d}}} \frac{\rho_\varsigma(v-\ddot{c})}{\rho_\varsigma(\mathbf{a}-\ddot{c})} d\mathbf{a}$$

By grouping integrals and summation signs, and splitting up the integral over $\mathcal{C}_M$ over multiple 'arcs' $A_{\ddot{c}}$ for $\ddot{c} \in \ddot{\mathcal{C}}_M$ (that satisfy $\|c - \ddot{c}\| < \varepsilon/n$ for all $c \in A_{\ddot{c}}$), we obtain

$$\|\mathcal{D}_{\mathcal{P}(\mathrm{Pic}_K^0)} - \ddot{\mathcal{D}}_{\mathcal{P}(\mathrm{Pic}_K^0)}\|$$

$$= \sum_{\mathfrak{d}\in\mathcal{I}_K} \left| \int_{\mathbf{a}\in\mathrm{Pic}_K^0} \mathcal{P}(\mathbf{a}) \sum_{\substack{v\in\mathbf{a}\\(v)=\mathbf{a}\mathfrak{d}}} \left(\int_{c\in\mathcal{C}_M} \frac{\rho_\varsigma(v-c)}{\rho_\varsigma(\mathbf{a}-c)} dc - \frac{1}{|\ddot{\mathcal{C}}_M|} \sum_{\ddot{c}\in\ddot{\mathcal{C}}_M} \frac{\rho_\varsigma(v-\ddot{c})}{\rho_\varsigma(\mathbf{a}-\ddot{c})} \right) d\mathbf{a} \right|$$

$$= \sum_{\mathfrak{d}\in\mathcal{I}_K} \left| \int_{\mathbf{a}\in\mathrm{Pic}_K^0} \mathcal{P}(\mathbf{a}) \sum_{\substack{v\in\mathbf{a}\\(v)=\mathbf{a}\mathfrak{d}}} \left(\sum_{\ddot{c}\in\ddot{\mathcal{C}}_M} \int_{c\in A_{\ddot{c}}} \left(\frac{\rho_\varsigma(v-c)}{\rho_\varsigma(\mathbf{a}-c)} - \frac{\rho_\varsigma(v-\ddot{c})}{\rho_\varsigma(\mathbf{a}-\ddot{c})} \right) dc \right) d\mathbf{a} \right|. \tag{5.75}$$

Applying the triangle inequality, switching integral and summation signs appropriately, collapsing the summation $\sum_{\mathfrak{d}\in\mathcal{I}_K} \sum_{\substack{v\in\mathbf{a}\\(v)=\mathbf{a}\mathfrak{d}}}$ to $\sum_{v\in\mathbf{a}}$ (as $\mathfrak{d}$ ranges

over integral ideals) and replacing the integral over Pic_K^0 by the maximum, we obtain that Equation (5.75) must be bounded by

$$\max_{\mathbf{a}\in\mathrm{Pic}_K^0}\left(\sum_{\ddot{c}\in\ddot{\mathcal{C}}_M}\int_{c\in A_{\ddot{c}}}\underbrace{\sum_{v\in\mathbf{a}}\left|\frac{\rho_\varsigma(v-c)}{\rho_\varsigma(\mathbf{a}-c)}-\frac{\rho_\varsigma(v-\ddot{c})}{\rho_\varsigma(\mathbf{a}-\ddot{c})}\right|}_{\|D_{\mathbf{a},\varsigma,c}-D_{\mathbf{a},\varsigma,\ddot{c}}\|}dc\right)\leq(4+\tfrac{\pi}{\varsigma^2}+2\pi n)\varepsilon/n$$
$$\leq 14\varepsilon. \tag{5.76}$$

This holds by the fact that $\|c-\ddot{c}\|<\varepsilon/n$ and $\varsigma>1$, together with Lemma A.37, which bounds the total variation distance between two discrete Gaussians with different centers. □

Conclusion

Applying Lemma 5.19 and Lemma 5.20 with $\mathcal{P}(\mathrm{Pic}_K^0)=\ddot{\mathcal{W}}(\mathrm{Pic}_K^0)$, and using Algorithm 6 for the running time, we obtain Theorem 5.11.