Lyubashevsky's Regularity Lemma

In this blogpost we exposit a result from Lyubashevsky et al. (2013) called the "regularity lemma" for Ring-LWE. Let m be an integer with $m \geq 3$. Let K be the m-th cyclotomic field. The field K can be either represented as $\mathbb{Q}[X]/\Phi(X)$ where $\Phi(X)$ is the m-th cyclotomic polynomial, or as $\mathbb{Q}(\zeta)$ where ζ is a root of $\Phi(X)$. The extension degree of K is $n = \varphi(m)$ where $\varphi(m)$ is the Euler totient function. Let R be the ring of algebraic integers in K. For cyclotomic fields, it is known that $R \cong \mathbb{Z}[\zeta]$ where ζ is a root of $\Phi(X)$. As such, we use polynomials in $\mathbb{Z}[\zeta]$ to represent elements of R. Let q be an integer with $q \geq 2$. We use R_q to denote the quotient ring R/qR. We represent elements of R_q with polynomials that, when seen as coefficient vectors, have each component within $[0, q-1]$. Thus R_q is a finite set of size q^n.

An informal, simplified statement of the regularity lemma is as follows. Let l be a positive integer with $l < 2^n$. Let $\mathcal{D}$ be a probability distribution over R_q. For suitable choices of $\mathcal{D}$, if we sample $a_1, \cdots, a_{l-1}$ uniformly from R_q, and sample $b_0, b_1, \cdots, b_{l-1}$ according to $\mathcal{D}$ from R_q, then the distribution of

$$b_0 + \sum_{i=1}^{l-1} a_i b_i$$

is close to the uniform distribution over R_q. This lemma is used to prove that the public keys of certain cryptographic systems are indistinguishable from random. In such systems one typically samples a vector $\mathbf{a} = (a_1, \cdots, a_{l-1})$ with uniform randomness, samples $\mathbf{s} = (s_1, \cdots, s_{l-1})$ and e according to some predefined distribution $\mathcal{D}$, and publishes $\mathbf{a}, \langle \mathbf{a}, \mathbf{s} \rangle + e$ as a public key.

1 Lattices

Definition 1. Let n be a positive integer. Let $\mathbf{v}_1, \cdots, \mathbf{v}_n$ be linearly independent vectors in $\mathbb{R}^n$. The set

$$L = \{k_1\mathbf{v}_1 + \cdots + k_n\mathbf{v}_n \mid k_1, \cdots, k_n \in \mathbb{Z}\}$$

is called a *lattice*. The set $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ is called a *basis* of L.

Example 2. The set of integer vectors $\mathbb{Z}^n$ forms a lattice, with basis $\{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$.

Remark 3. The basis of a lattice is not unique. For example, we can negate the sign of any element in the basis, and the resulting set is still a basis.

Definition 4. If L is a lattice with basis $\mathbf{v}_1, \cdots, \mathbf{v}_n$, then we define the *determinant* of L, denoted by $d(L)$, to be

$$d(L) = \det \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{bmatrix}.$$

Remark 5. The determinant of a lattice L is well-defined up to sign. If $\mathbf{v}_1, \cdots, \mathbf{v}_n$ and $\mathbf{w}_1, \cdots, \mathbf{w}_n$ are two bases for

L, then we can represent each $\mathbf{w}_i$ as

$$\mathbf{w}_i = \sum_{j=1}^{n} k_{ij}\mathbf{v}_j, \quad k_{ij} \in \mathbb{Z}.$$

Similarly, we have

$$\mathbf{v}_i = \sum_{j=1}^{n} k'_{ij}\mathbf{w}_j, \quad k'_{ij} \in \mathbb{Z}.$$

Now consider the matrices

$$\mathbf{V} = \begin{bmatrix}\mathbf{v}_1 & \cdots & \mathbf{v}_n\end{bmatrix}^\top, \quad \mathbf{W} = \begin{bmatrix}\mathbf{w}_1 & \cdots & \mathbf{w}_n\end{bmatrix}^\top, \quad \mathbf{A} = \begin{pmatrix} k_{11} & \cdots & k_{1n} \\ \vdots & \ddots & \vdots \\ k_{n1} & \cdots & k_{nn} \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} k'_{11} & \cdots & k'_{1n} \\ \vdots & \ddots & \vdots \\ k'_{n1} & \cdots & k'_{nn} \end{pmatrix}.$$

We have $\mathbf{W} = \mathbf{AV}$ and $\mathbf{V} = \mathbf{BW}$, so $\det \mathbf{W} = \det \mathbf{A} \cdot \det \mathbf{V}, \det \mathbf{V} = \det \mathbf{B} \cdot \det \mathbf{W}$. Thus $\det \mathbf{A} \cdot \det \mathbf{B} = 1$. But both $\det \mathbf{A}, \det \mathbf{B}$ are integers, since each entry of $\mathbf{A}, \mathbf{B}$ is an integer. Therefore, both $\det \mathbf{A} = \pm 1, \det \mathbf{B} = \pm 1$, and $\det \mathbf{W} = \pm \det \mathbf{V}$. □

Definition 6. Let L be a lattice of rank n. The *dual lattice* of L, denoted by L^*, is the set

$$L^* = \{\mathbf{w} \in \mathbb{R}^n \mid \forall \mathbf{v} \in L, \langle \mathbf{v}, \mathbf{w} \rangle \in \mathbb{Z}\}.$$

Remark 7. Suppose that $\mathbf{v}_1, \cdots, \mathbf{v}_n$ is a basis of L. To see that the dual lattice of L is also a lattice of rank n, notice that we can find $\mathbf{w}_1, \cdots, \mathbf{w}_n$, such that

$$\langle \mathbf{w}_i, \mathbf{v}_j \rangle = \begin{cases} 1 & (i = j) \\ 0 & (i \neq j) \end{cases}.$$

Then it is easy to see that $\mathbf{w}_1, \cdots, \mathbf{w}_n$ is a basis of L^*. □

Lemma 8. Lattice duality is an involution. For each lattice L we have $(L^*)^* = L$.

Proof: Let $\mathbf{v}_1, \cdots, \mathbf{v}_n$ be a basis of L, and $\mathbf{w}_1, \cdots, \mathbf{w}_n$ be a basis of L^* as constructed above. Then it is easy to see that $\mathbf{v}_i \in (L^*)^*$, thus $L \subseteq (L^*)^*$. Now suppose $\mathbf{v}$ is a vector in $(L^*)^*$. We can write $\mathbf{v}$ as $\mathbf{v} = r_1\mathbf{v}_1 + \cdots + r_n\mathbf{v}_n$ with $r_i \in \mathbb{R}$. Then we have $\langle \mathbf{v}, \mathbf{w}_i \rangle = r_i$. But since $\mathbf{w}_i \in L^*$, we have $r_i \in \mathbb{Z}$. Hence $\mathbf{v} \in L$. □

Lemma 9. If L is a lattice then $d(L) = 1/d(L^*)$.

Proof: Let $\mathbf{v}_1, \cdots, \mathbf{v}_n$ be a basis of L, and $\mathbf{w}_1, \cdots, \mathbf{w}_n$ be a basis of L^* as constructed in Remark 7. Define the matrices

$$\mathbf{V} = \begin{bmatrix}\mathbf{v}_1 & \cdots & \mathbf{v}_n\end{bmatrix}, \quad \mathbf{W} = \begin{bmatrix}\mathbf{w}_1 & \cdots & \mathbf{w}_n\end{bmatrix}^\top.$$

Then we have $\mathbf{WV} = \mathbf{I}_n$. Hence $d(L) = \det \mathbf{V} = 1/\det \mathbf{W} = 1/d(L^*)$. □

Lemma 10. Let L, L' be two lattices of rank n with $L' \subseteq L$. Define an equivalence relation $\mathbf{v} \leftrightarrow \mathbf{v}'$ on L determined by

$$\mathbf{v} \leftrightarrow \mathbf{v}' \equiv \mathbf{v} - \mathbf{v}' \in L'.$$

Then the number of equivalence classes of this relation is exactly $|d(L')/d(L)|$. Furthermore, for each $\mathbf{v} \in L$, we have $|d(L')/d(L)|\,\mathbf{v} \in L'$.

Proof: Without loss of generality, we may assume $L = \mathbb{Z}^n$. Let $\mathbf{v}_1, \cdots, \mathbf{v}_n$ be a basis of L'. Let $\mathbf{V} = \begin{bmatrix}\mathbf{v}_1 & \cdots & \mathbf{v}_n\end{bmatrix}$. For any $\mathbf{v} \in \mathbb{Z}^n$, $\mathbf{v} \in L'$ if and only if $\mathbf{V}^{-1}\mathbf{v} \in \mathbb{Z}^n$.

Now recall that every matrix in $\mathbb{Z}^{n\times n}$ has a *Smith normal form*. There exists matrices $\mathbf{S}, \mathbf{A}, \mathbf{T}$ with the following properties:

1. $\mathbf{S}, \mathbf{A}, \mathbf{T} \in \mathbb{Z}^{n\times n}$;
2. $\mathbf{S}^{-1}, \mathbf{T}^{-1} \in \mathbb{Z}^{n\times n}$;
3. $\mathbf{V} = \mathbf{SAT}$;
4. Matrix $\mathbf{A}$ is diagonal.

Since the matrices $\mathbf{S}, \mathbf{T}$ and their inverses all contain only integer entires, we must have $\det \mathbf{S} = \pm 1, \det \mathbf{T} = \pm 1$. Let $a_1, \cdots, a_n$ be the diagonal entries of $\mathbf{A}$. Then we have $|\det \mathbf{V}| = |a_1 \cdots a_n|$. Now $\mathbf{V}^{-1} = \mathbf{T}^{-1}\mathbf{A}^{-1}\mathbf{S}^{-1}$. For a given vector $\mathbf{v} \in \mathbb{Z}^n$, let $\mathbf{w} = \mathbf{S}^{-1}\mathbf{v}$. Then it is easy to see that $\mathbf{V}^{-1}\mathbf{v} \in \mathbb{Z}^n$ if and only if each w_i is a multiple of a_i. Hence, $\mathbf{v} \leftrightarrow \mathbf{v}'$ iff $w_1 = w_1' \bmod a_1, w_2 = w_2' \bmod a_2$, etc. From this it is easy to see that the number of equivalence classes is at most $|\det \mathbf{V}|$. To see that it is exactly $|\det \mathbf{V}|$, note that for each $\mathbf{w} \in \mathbb{Z}^n$ we have $\mathbf{w} = \mathbf{S}^{-1}(\mathbf{Sw})$, and so every equivalence class has at least one member.

For any given $\mathbf{v} \in \mathbb{Z}^n$, every component of $|d(L')/d(L)|\,\mathbf{v}$ is a multiple of $|d(L')/d(L)|$, hence a multiple of $a_1, \cdots, a_n$. Hence $\mathbf{V}^{-1}(|d(L')/d(L)|\,\mathbf{v}) \in \mathbb{Z}^n$ and $|d(L')/d(L)|\,\mathbf{v} \in L'$. □

Lemma 11. If L_1, L_2 are two lattices such that $L_1 \subseteq L_2$, then $L_2^* \subseteq L_1^*$.

Proof: If $\mathbf{w}$ is a vector such that $\langle \mathbf{w}, \mathbf{v}\rangle \in \mathbb{Z}$ for any $\mathbf{v} \in L_2$, then also $\langle \mathbf{w}, \mathbf{v}\rangle \in \mathbb{Z}$ for any $\mathbf{v} \in L_1$. □

Lemma 12 (Minkowski's Theorem)**.** Let L be a lattice of rank n. Let V be a centrally symmetric convex subset of $\mathbb{R}^n$. If $\mathrm{vol}(V) > 2^n \cdot |d(L)|$ then V contains at least one non-zero vector in L.

Proof: Jarvis (2014, Theorem 7.8, p. 152). □

2 Harmonic Analysis

Let $f(\mathbf{x})$ be a continuous function $\mathbb{R}^n \mapsto \mathbb{C}$, such that $\int_{\mathbb{R}^n} |f(\mathbf{x})|\ \mathrm{d}\mathbf{x}$ converges. Let S be a countable subset of $\mathbb{R}^n$ such that $\sum_S |f(\mathbf{x})|$ also converges. We can define a generalized function $\mu(\mathbf{x})$ on $\mathbb{R}^n$ as

$$\mu(\mathbf{x}) = \sum_{\mathbf{x}_0 \in S} f(\mathbf{x}_0)\, \delta(\mathbf{x} - \mathbf{x}_0),$$

where $\delta(\mathbf{x})$ is Dirac's delta function.

Definition 13 (Fourier transform for continuous functions over $\mathbb{R}^n$)**.** The *Fourier transform* of $f(\mathbf{x})$ is a function $\widehat{f}(\mathbf{y})$ defined as

$$\widehat{f}(\mathbf{y}) = \int_{\mathbb{R}^n} e^{2\pi i \langle \mathbf{x}, \mathbf{y}\rangle}\, f(\mathbf{x})\ \mathrm{d}\mathbf{x}.$$

Definition 14 (Fourier transform for discrete sums over S)**.** The Fourier transform of $\mu(\mathbf{x})$ is a function $\widehat{\mu}(\mathbf{y}) : \mathbb{R}^n \mapsto \mathbb{C}$, defined as

$$\widehat{\mu}(\mathbf{y}) = \sum_{\mathbf{x} \in S} e^{2\pi i \langle \mathbf{x}, \mathbf{y}\rangle}\, f(\mathbf{x}).$$

In this section we follow Grafakos (2014) to establish Poisson's summation formula:

Theorem 15. If there exists $C, \delta \in \mathbb{R}$ with $C > 0, \delta > 0$, and

$$\forall \mathbf{x} \in \mathbb{R}^n,\ |f(\mathbf{x})| \leq C(1+|x|)^{-n-\delta},$$

$$\sum_{\mathbf{m}\in\mathbb{Z}^n} |\widehat{f}(\mathbf{y}+\mathbf{m})| < \infty,$$

then for every $\mathbf{u}, \mathbf{y} \in \mathbb{R}^n$ we have

$$\sum_{\mathbf{x}\in\mathbb{Z}^n+\mathbf{u}} e^{2\pi i\langle \mathbf{x},\mathbf{y}\rangle} f(\mathbf{x}) = \sum_{\mathbf{z}\in\mathbb{Z}^n} e^{-2\pi i\langle \mathbf{z},\mathbf{u}\rangle}\, \widehat{f}(\mathbf{y}+\mathbf{z}).$$

2.1 Fourier Series of Periodic Functions

Let us define

$$F(\mathbf{u}) = \sum_{\mathbf{x}\in\mathbb{Z}^n+\mathbf{u}} e^{2\pi i\langle \mathbf{x},\mathbf{y}\rangle} f(\mathbf{x}).$$

We have $F(\mathbf{u}) = \widehat{\mu}(\mathbf{y})$ with $S = \mathbb{Z}^n + \mathbf{u}$. Now $F(\mathbf{u})$ is periodic over $\mathbb{Z}^n$. If $\mathbf{u} - \mathbf{u}' \in \mathbb{Z}^n$ then $\mathbb{Z}^n + \mathbf{u} = \mathbb{Z}^n + \mathbf{u}'$.

We write $\mathbb{T}_n$ for the n-torus $\mathbb{R}^n/\mathbb{Z}^n$. Thus a function $f(\mathbf{x})$ is defined on $\mathbb{T}^n$ if it is defined on $\mathbb{R}^n$ and periodic over $\mathbb{Z}^n$. We represent each element in $\mathbb{T}^n$ with a point in the cube $[-1/2, 1/2]^n$.

Definition 16 (Fourier series for functions periodic over $\mathbb{Z}^n$)**.** The *Fourier series* of $F(\mathbf{u})$ is a function $\widehat{F}(\mathbf{m}) : \mathbb{Z}^n \mapsto \mathbb{C}$, defined as

$$\widehat{F}(\mathbf{m}) = \int_{[-1/2,1/2]^n} e^{2\pi i\langle \mathbf{u},\mathbf{m}\rangle} F(\mathbf{u})\, \mathrm{d}\mathbf{u}.$$

In this subsection we prove that:

Theorem 17. If $F(\mathbf{u}), G(\mathbf{u})$ are two continuous functions $\mathbb{T}^n \mapsto \mathbb{C}$, and $\widehat{F}(\mathbf{m}) = \widehat{G}(\mathbf{m})$ for every $\mathbf{m} \in \mathbb{Z}^n$, then $F(\mathbf{u}) = G(\mathbf{u})$ for every $\mathbf{u} \in \mathbb{T}^n$.

Definition 18. An *approximate identity* is a sequence of continuous functions $k_1, k_2, \cdots : \mathbb{T}^n \mapsto \mathbb{R}$ such that:

1. There exists a constant $c > 0$ such that $\int_{\mathbb{T}^n} |k_n(\mathbf{x})|\, \mathrm{d}\mathbf{x} \leq c$ for every k_n in the sequence;
2. For every k_n in sequence we have $\int_{\mathbb{T}^n} k_n(\mathbf{x})\, \mathrm{d}\mathbf{x} = 1$;
3. For every $0 < \delta < 1/2$, let $\mathcal{B}(\delta)$ be the open ball

$$\mathcal{B}(\delta) = \{\mathbf{x} \in \mathbb{T}^n \mid ||\mathbf{x}|| < \delta\},$$

and let $\mathcal{C}(\delta) = \mathbb{T}^n \setminus \mathcal{B}(\delta)$, then $\int_{\mathcal{C}(\delta)} |k_n(\mathbf{x})|\, \mathrm{d}\mathbf{x} \to 0$ as $n \to \infty$.

Definition 19. Let $f(\mathbf{x}), k(\mathbf{x})$ be continuous functions $\mathbb{T}^n \mapsto \mathbb{C}$. Then the *convolution* of f and k, denoted by $(f * k)$, is defined as

$$(f * k)(\mathbf{x}) = \int_{\mathbb{T}^n} f(\mathbf{x}-\mathbf{y}) k(\mathbf{y})\, \mathrm{d}\mathbf{y}.$$

Definition 20. The *Fejér kernel* $F_N^n(x_1, \cdots, x_n)$ is defined as

$$F_N^n(x_1, \cdots, x_n) = \sum_{\substack{\mathbf{m}\in\mathbb{Z}^n \\ |m_j \leq N|}} \left(1 - \frac{|m_1|}{N+1}\right) \cdots \left(1 - \frac{|m_1|}{N+1}\right) e^{2\pi i\langle \mathbf{m},\mathbf{x}\rangle}.$$

Lemma 21. The Fejér kernel F_N^n satisfies

$$F_N^n(x_1, \cdots, x_n) = \frac{1}{(N+1)^n} \prod_{j=1}^{n} \left(\frac{\sin(\pi(N+1)x_j)}{\sin(\pi x_j)} \right)^2.$$

Proof: Notice that

$$F_N^n(x_1, \cdots, x_n) = \prod_{j=1}^{n} F_N^1(x_j).$$

So it is sufficient to prove the lemma for the case $n = 1$. This sum is easy to do and I shall omit the details. □

Lemma 22. The Fejér kernels $F_1^n, F_2^n, \cdots$ constitute an approximate identity.

Proof: By Lemma 21 we have $F_N^n(\mathbf{x}) \geq 0$. Thus

$$\int_{\mathbb{T}^n} |F_N^n(\mathbf{x})| \, \mathrm{d}\mathbf{x} = \int_{\mathbb{T}^n} F_N^n(\mathbf{x}) \, \mathrm{d}\mathbf{x}.$$

Now notice that

$$\int_{\mathbb{T}^n} e^{2\pi i \langle \mathbf{m}, \mathbf{x} \rangle} = \prod_{j=1}^{n} \int_{-1/2}^{1/2} e^{2\pi i m_j x_j} \, \mathrm{d}x_j.$$

If $m_j \in \mathbb{Z} \setminus \{0\}$ then $e^{2\pi i m_j x_j}$ integrates to 0. Hence the only term that does not vanish in $\int_{\mathbb{T}^n} F_N^n(\mathbf{x}) \, \mathrm{d}\mathbf{x}$ is the term with $\mathbf{m} = \mathbf{0}$. We have

$$\int_{\mathbb{T}^n} |F_N^n(\mathbf{x})| \, \mathrm{d}\mathbf{x} = \int_{\mathbb{T}^n} F_N^n(\mathbf{x}) \, \mathrm{d}\mathbf{x} = 1.$$

Therefore properties 1 and 2 in Definition 18 are satisfied.

To prove property 3, first notice that $|\sin(\pi(N+1)x)| \leq \pi(N+1)|x|$ and $|\sin(\pi(N+1)x)| \leq 1$. Therefore

$$F_N^1(x) \leq \frac{1}{N+1} \min \left(\frac{\pi(N+1)|x|}{|\sin(\pi x)|}, \frac{1}{|\sin(\pi x)|} \right)^2.$$

Now we claim that, when $|x| \leq 1/2$,

$$1 \leq \pi|x|/|\sin(\pi x)| \leq \pi/2.$$

By symmetry it is sufficient to prove the case $x \geq 0$. The first inequality is obvious. To prove the second inequality, define $g(x) = \pi x / \sin(\pi x)$, then

$$g'(x) = \frac{\sin(x) - x\cos(x)}{\sin^2(x)}.$$

Define $h(x) = \sin(x) - x\cos(x)$, then

$$h'(x) = x\sin(x) \geq 0 \quad (|x| \leq \pi).$$

Therefore $h(x) \geq 0$ and $g'(x) \geq 0$ when $0 \leq x \leq 1/2$. Hence $g(x) \leq g(1/2) = \pi/2$ when $0 \leq x \leq 1/2$.

Now we have

$$F_N^1(x) \leq \frac{1}{N+1} \left(\frac{\pi|x|}{|\sin(\pi x)|} \right)^2 \min \left(N+1, \frac{1}{\pi|x|} \right)^2 \leq \frac{1}{N+1} \cdot \frac{\pi^2}{4} \min \left(N+1, \frac{1}{\pi|x|} \right)^2.$$

Thus for $\delta > 0$ we have

$$\int_{\delta \le |x| \le 1/2} F_N^1(x)\,\mathrm{d}x \le \frac{1}{N+1} \cdot \frac{\pi^2}{4} \int_{\delta \le |x| \le 1/2} \frac{\mathrm{d}\mathbf{x}}{|\pi x|^2} \le \frac{1}{4\delta^2(N+1)}.$$

As N increases, the integral approaches 0.

For general F_N^n, given any $\mathbf{x} \in [-1/2, 1/2]^n$ with $||\mathbf{x}|| \ge \delta$, at least one x_j satisfies $|x_j| \ge \delta/\sqrt{n}$. Therefore

$$\int_{\mathcal{C}(\delta)} F_N^n(\mathbf{x})\,\mathrm{d}\mathbf{x} \le \sum_{j=1}^{n} \left[\left(\int_{\frac{\delta}{\sqrt{n}} \le |x_j| \le 1/2} F_N^1(x_j)\,\mathrm{d}x_j \right) \prod_{k \ne j} \int_{-1/2}^{1/2} F_N^1(x_k)\,\mathrm{d}x_k \right] \le \frac{n}{4(\delta/\sqrt{n})^2(N+1)}.$$

The integral also approaches 0 as N increases. □

Lemma 23. Let $k_1, k_2, \cdots$ be an approximate identity. Then for every continuous function $f(\mathbf{x}) : \mathbb{T}^n \mapsto \mathbb{C}$, we have

$$\int_{\mathbb{T}^n} |(k_N * f)(\mathbf{x}) - f(\mathbf{x})|\,\mathrm{d}\mathbf{x} \to 0 \text{ as } N \to \infty.$$

Proof: Since $[-1/2, 1/2]^n$ is a compact set, $f(\mathbf{x})$ is uniformly continuous. For every $\varepsilon > 0$, we can find $\delta < 1/2$ such that

$$\forall \mathbf{x}, \mathbf{x}' \in \mathbb{T}^n,\ ||\mathbf{x} - \mathbf{x}'|| < \delta \Rightarrow |f(\mathbf{x}) - f(\mathbf{x}')| < \varepsilon.$$

Also, $|f(\mathbf{x})|$ has a maximum value M. Choose a sufficiently large N, such that

$$\int_{\mathcal{C}(\delta)} |k_N(\mathbf{x})|\,\mathrm{d}\mathbf{x} < \min(\varepsilon/M, \varepsilon).$$

Then we have

$$\begin{aligned}
\int_{\mathbb{T}^n} |(k_N * f)(\mathbf{x}) - f(\mathbf{x})|\,\mathrm{d}\mathbf{x} &= \int_{\mathbb{T}^n} \left| \left(\int_{\mathbb{T}^n} f(\mathbf{x}-\mathbf{y}) k_N(\mathbf{y})\,\mathrm{d}\mathbf{y} \right) - f(\mathbf{x}) \right| \mathrm{d}\mathbf{x} \\
&= \int_{\mathbb{T}^n} \left| \left(\int_{\mathcal{B}(\delta)} f(\mathbf{x}-\mathbf{y}) k_N(\mathbf{y})\,\mathrm{d}\mathbf{y} \right) + \left(\int_{\mathcal{C}(\delta)} f(\mathbf{x}-\mathbf{y}) k_N(\mathbf{y})\,\mathrm{d}\mathbf{y} \right) - f(\mathbf{x}) \right| \mathrm{d}\mathbf{x} \\
&\le \int_{\mathbb{T}^n} \left| \left(\int_{\mathcal{B}(\delta)} f(\mathbf{x}-\mathbf{y}) k_N(\mathbf{y})\,\mathrm{d}\mathbf{y} \right) - f(\mathbf{x}) \right| + \left| \int_{\mathcal{C}(\delta)} f(\mathbf{x}-\mathbf{y}) k_N(\mathbf{y})\,\mathrm{d}\mathbf{y} \right| \mathrm{d}\mathbf{x} \\
&\le \varepsilon + \int_{\mathbb{T}^n} \left| \left(\int_{\mathcal{B}(\delta)} f(\mathbf{x}-\mathbf{y}) k_N(\mathbf{y})\,\mathrm{d}\mathbf{y} \right) - f(\mathbf{x}) \right| \mathrm{d}\mathbf{x} \\
&= \varepsilon + \int_{\mathbb{T}^n} \left| \left(\int_{\mathcal{B}(\delta)} f(\mathbf{x}) k_N(\mathbf{y})\,\mathrm{d}\mathbf{y} \right) - f(\mathbf{x}) + \int_{\mathcal{B}(\delta)} (f(\mathbf{x}-\mathbf{y}) - f(\mathbf{x})) k_N(\mathbf{y})\,\mathrm{d}\mathbf{y} \right| \mathrm{d}\mathbf{x} \\
&\le \varepsilon + \int_{\mathbb{T}^n} \left| \left(\int_{\mathcal{B}(\delta)} f(\mathbf{x}) k_N(\mathbf{y})\,\mathrm{d}\mathbf{y} \right) - f(\mathbf{x}) \right| + \left| \int_{\mathcal{B}(\delta)} (f(\mathbf{x}-\mathbf{y}) - f(\mathbf{x})) k_N(\mathbf{y})\,\mathrm{d}\mathbf{y} \right| \mathrm{d}\mathbf{x}.
\end{aligned}$$

Because $\int_{\mathbb{T}^n} k_N(\mathbf{y})\,\mathrm{d}\mathbf{y} = 1$, we have

$$\int_{\mathcal{B}(\delta)} k_N(\mathbf{y})\,\mathrm{d}\mathbf{y} = 1 - \int_{\mathcal{C}(\delta)} k_N(\mathbf{y})\,\mathrm{d}\mathbf{y}.$$

But $|\int_{\mathcal{C}(\delta)} k_N(\mathbf{y})\,d\mathbf{y}| \le \int_{\mathcal{C}(\delta)} |k_N(\mathbf{y})|\,d\mathbf{y} \le \varepsilon$, so

$$1-\varepsilon \le \int_{\mathcal{B}(\delta)} k_N(\mathbf{y})\,d\mathbf{y} \le 1+\varepsilon.$$

We have

$$\left|\left(\int_{\mathcal{B}(\delta)} f(\mathbf{x})k_N(\mathbf{y})\,d\mathbf{y}\right) - f(\mathbf{x})\right| \le \varepsilon\cdot|f(\mathbf{x})| \le \varepsilon M.$$

Similarly

$$\left|\int_{\mathcal{B}(\delta)} (f(\mathbf{x}-\mathbf{y})-f(\mathbf{x}))k_N(\mathbf{y})\,d\mathbf{y}\right| \le \varepsilon\cdot(c-\varepsilon),$$

where c is the constant in requirement 1 of Definition 18. Together we get

$$\int_{\mathbb{T}^n} |(k_N * f)(\mathbf{x}) - f(\mathbf{x})|\,d\mathbf{x} \le \varepsilon\cdot(1+c+M-\varepsilon).$$

As $\varepsilon\to 0$, we see that $\int_{\mathbb{T}^n} |(k_N * f)(\mathbf{x}) - f(\mathbf{x})|\,d\mathbf{x} \to 0$. □

Lemma 24. Let $F(\mathbf{x})$ be a continuous function on $\mathbb{T}^n$. For each $\mathbf{x}\in\mathbb{T}^n$ we have

$$(F_N^n * F)(\mathbf{x}) = \sum_{\substack{\mathbf{m}\in\mathbb{Z}^n\\ |m_j|\le N}} \left(1-\frac{|m_1|}{N+1}\right)\cdots\left(1-\frac{|m_1|}{N+1}\right)\widehat{F}(-\mathbf{m})\,e^{2\pi i\langle\mathbf{m},\mathbf{x}\rangle}.$$

Proof:

$$\begin{aligned}
(F_N^n * F)(\mathbf{x}) &= \int_{\mathbb{T}^n} \sum_{\substack{\mathbf{m}\in\mathbb{Z}^n\\ |m_j|\le N}} \left(1-\frac{|m_1|}{N+1}\right)\cdots\left(1-\frac{|m_1|}{N+1}\right) e^{2\pi i\langle\mathbf{m},\mathbf{y}\rangle} f(\mathbf{x}-\mathbf{y})\,d\mathbf{y}\\
&= \sum_{\substack{\mathbf{m}\in\mathbb{Z}^n\\ |m_j|\le N}} \left(1-\frac{|m_1|}{N+1}\right)\cdots\left(1-\frac{|m_1|}{N+1}\right) \int_{\mathbb{T}^n} e^{2\pi i\langle\mathbf{m},\mathbf{y}\rangle} f(\mathbf{x}-\mathbf{y})\,d\mathbf{y}\\
&= \sum_{\substack{\mathbf{m}\in\mathbb{Z}^n\\ |m_j|\le N}} \left(1-\frac{|m_1|}{N+1}\right)\cdots\left(1-\frac{|m_1|}{N+1}\right) \int_{\mathbb{T}^n} e^{2\pi i\langle\mathbf{m},\mathbf{x}-\mathbf{y}\rangle} f(\mathbf{y})\,d\mathbf{y}\\
&= \sum_{\substack{\mathbf{m}\in\mathbb{Z}^n\\ |m_j|\le N}} \left(1-\frac{|m_1|}{N+1}\right)\cdots\left(1-\frac{|m_1|}{N+1}\right) e^{2\pi i\langle\mathbf{m},\mathbf{x}\rangle} \int_{\mathbb{T}^n} e^{2\pi i\langle-\mathbf{m},\mathbf{y}\rangle} f(\mathbf{y})\,d\mathbf{y}\\
&= \sum_{\substack{\mathbf{m}\in\mathbb{Z}^n\\ |m_j|\le N}} \left(1-\frac{|m_1|}{N+1}\right)\cdots\left(1-\frac{|m_1|}{N+1}\right) \widehat{F}(-\mathbf{m})\,e^{2\pi i\langle\mathbf{m},\mathbf{x}\rangle}.
\end{aligned}$$

□

Proof of Theorem 17: Define $H(\mathbf{x}) = F(\mathbf{x}) - G(\mathbf{x})$. Then $\widehat{H}(\mathbf{m}) = \widehat{F}(\mathbf{m}) - \widehat{G}(\mathbf{m}) = 0$. Hence $(F_N^n * H)(\mathbf{x}) = 0$ for

every $\mathbf{x} \in \mathbb{T}^n$. By Lemma 23 we have

$$\int_{\mathbb{T}^n} |(F_N^n * H)(\mathbf{x}) - H(\mathbf{x})| \, \mathrm{d}\mathbf{x} \to 0.$$

We conclude that $H(\mathbf{x}) = 0$, thus $F(\mathbf{x}) = G(\mathbf{x})$. □

We also get the following corollary:

Theorem 25 (Fourier inversion)**.** If $F(\mathbf{x})$ is a continuous function $\mathbb{T}^n \mapsto \mathbb{C}$, and

$$\sum_{\mathbf{m}\in\mathbb{Z}^n} |\widehat{F}(\mathbf{m})| < \infty,$$

then we have

$$F(\mathbf{x}) = \sum_{\mathbf{m}\in\mathbb{Z}^n} \widehat{F}(\mathbf{m}) \, e^{-2\pi i\langle \mathbf{m},\mathbf{x}\rangle}.$$

Proof: We use $G(\mathbf{x})$ to denote the RHS. We have

$$\begin{aligned}\widehat{G}(\mathbf{m}) &= \int_{\mathbb{T}^n} \sum_{\mathbf{m}'\in\mathbb{Z}^n} \widehat{F}(\mathbf{m}') \, e^{2\pi i\langle \mathbf{m}-\mathbf{m}',\mathbf{x}\rangle} \, \mathrm{d}\mathbf{x} \\ &= \sum_{\mathbf{m}'\in\mathbb{Z}^n} \left[\widehat{F}(\mathbf{m}') \int_{\mathbb{T}^n} e^{2\pi i\langle \mathbf{m}-\mathbf{m}',\mathbf{x}\rangle} \, \mathrm{d}\mathbf{x} \right].\end{aligned}$$

If $\mathbf{m} \neq \mathbf{m}'$ then $\int_{\mathbb{T}^n} e^{2\pi i\langle \mathbf{m}-\mathbf{m}',\mathbf{x}\rangle} \, \mathrm{d}\mathbf{x} = 0$. Hence $\widehat{G}(\mathbf{m}) = \widehat{F}(\mathbf{m})$, and we may conclude $F(\mathbf{x}) = G(\mathbf{x})$. □

2.2 Poisson's Summation Formula

We now return to the function $F(\mathbf{u}) = \sum_{\mathbf{x}\in\mathbb{Z}^n+\mathbf{u}} e^{2\pi i\langle \mathbf{x},\mathbf{y}\rangle} f(\mathbf{x})$.

Proof of Theorem 15: We have

$$\begin{aligned}\widehat{F}(\mathbf{m}) &= \int_{\mathbb{T}^n} e^{2\pi i\langle \mathbf{m},\mathbf{u}\rangle} \left(\sum_{\mathbf{x}\in\mathbb{Z}^n+\mathbf{u}} e^{2\pi i\langle \mathbf{x},\mathbf{y}\rangle} f(\mathbf{x}) \right) \mathrm{d}\mathbf{u} \\ &= \int_{\mathbb{T}^n} e^{2\pi i\langle \mathbf{m},\mathbf{u}\rangle} \left(\sum_{\mathbf{x}\in\mathbb{Z}^n} e^{2\pi i\langle \mathbf{x}+\mathbf{u},\mathbf{y}\rangle} f(\mathbf{x}+\mathbf{u}) \right) \mathrm{d}\mathbf{u} \\ &= \sum_{\mathbf{x}\in\mathbb{Z}^n} \int_{\mathbb{T}^n} e^{2\pi i\langle \mathbf{m},\mathbf{u}\rangle} \, e^{2\pi i\langle \mathbf{x}+\mathbf{u},\mathbf{y}\rangle} \, f(\mathbf{x}+\mathbf{u}) \, \mathrm{d}\mathbf{u} && \text{(Weierstrass M-test)} \\ &= \sum_{\mathbf{x}\in\mathbb{Z}^n} \int_{\mathbb{T}^n} e^{2\pi i\langle \mathbf{m},(\mathbf{x}+\mathbf{u})-\mathbf{x}\rangle} \, e^{2\pi i\langle \mathbf{x}+\mathbf{u},\mathbf{y}\rangle} \, f(\mathbf{x}+\mathbf{u}) \, \mathrm{d}\mathbf{u} \\ &= \sum_{\mathbf{x}\in\mathbb{Z}^n} \int_{\mathbb{T}^n} e^{2\pi i\langle \mathbf{m},\mathbf{x}+\mathbf{u}\rangle} \, e^{2\pi i\langle \mathbf{x}+\mathbf{u},\mathbf{y}\rangle} \, f(\mathbf{x}+\mathbf{u}) \, \mathrm{d}\mathbf{u} && (e^{2\pi i\langle \mathbf{m},\mathbf{x}\rangle} = 1) \\ &= \sum_{\mathbf{x}\in\mathbb{Z}^n} \int_{\mathbb{T}^n} e^{2\pi i\langle \mathbf{m}+\mathbf{y},\mathbf{x}+\mathbf{u}\rangle} \, f(\mathbf{x}+\mathbf{u}) \, \mathrm{d}\mathbf{u} \\ &= \int_{\mathbb{R}^n} e^{2\pi i\langle \mathbf{m}+\mathbf{y},\mathbf{x}\rangle} \, f(\mathbf{x}) \, \mathrm{d}\mathbf{x}\end{aligned}$$

$$= \widehat{f}(\mathbf{y}+\mathbf{m}).$$

By assumption we have $\sum_{\mathbf{m}\in\mathbb{Z}^n} |\widehat{F}(\mathbf{m})| < \infty$. By Theorem 25, we get

$$F(\mathbf{u}) = \sum_{\mathbf{m}\in\mathbb{Z}^n} \widehat{F}(\mathbf{m})\, e^{2\pi i\langle \mathbf{m},\mathbf{u}\rangle} = \sum_{\mathbf{z}\in\mathbb{Z}^n} e^{-2\pi i\langle \mathbf{z},\mathbf{u}\rangle}\, \widehat{f}(\mathbf{y}+\mathbf{z}).$$

□

2.3 General Lattices

Suppose that we want to sum not over $\mathbb{Z}^n + \mathbf{u}$ but over $L + \mathbf{u}$ where L may be any lattice of rank n. Let $\mathbf{v}_1, \cdots, \mathbf{v}_n$ be a basis of L. Let $\mathbf{L} = \begin{bmatrix}\mathbf{v}_1 & \cdots & \mathbf{v}_n\end{bmatrix}$. Define

$$g(\mathbf{x}) = f(\mathbf{L}\mathbf{x}),$$

so that we want to compute $\sum_{\mathbf{x}\in\mathbb{Z}^n+\mathbf{L}^{-1}\mathbf{u}} g(\mathbf{x})$. The Fourier transform of $g(\mathbf{x})$ is

$$\begin{aligned}
\hat{g}(\mathbf{y}) &= \int_{\mathbb{R}^n} e^{2\pi i\langle \mathbf{x},\mathbf{y}\rangle}\, g(\mathbf{x})\, \mathrm{d}\mathbf{x} \\
&= \int_{\mathbb{R}^n} e^{2\pi i\langle \mathbf{x},\mathbf{y}\rangle}\, f(\mathbf{L}\mathbf{x})\, \mathrm{d}\mathbf{x} \\
&= \frac{1}{\det \mathbf{L}} \int_{\mathbb{R}^n} e^{2\pi i\langle \mathbf{L}^{-1}\mathbf{z},\mathbf{y}\rangle}\, f(\mathbf{z})\, \mathrm{d}\mathbf{z} && (\mathbf{x} = \mathbf{L}^{-1}\mathbf{z}) \\
&= \frac{1}{\det \mathbf{L}} \int_{\mathbb{R}^n} e^{2\pi i\langle \mathbf{z},\mathbf{L}^{-\mathsf{T}}\mathbf{y}\rangle}\, f(\mathbf{z})\, \mathrm{d}\mathbf{z} \\
&= \frac{\widehat{f}(\mathbf{L}^{-\mathsf{T}}\mathbf{y})}{\det \mathbf{L}}.
\end{aligned}$$

Hence, by Theorem 15 we get

$$\begin{aligned}
\sum_{\mathbf{x}\in L+\mathbf{u}} e^{2\pi i\langle \mathbf{x},\mathbf{y}\rangle}\, f(\mathbf{x}) &= \sum_{\mathbf{x}\in\mathbb{Z}^n+\mathbf{L}^{-1}\mathbf{u}} e^{2\pi i\langle \mathbf{L}\mathbf{x},\mathbf{y}\rangle}\, g(\mathbf{x}) \\
&= \sum_{\mathbf{x}\in\mathbb{Z}^n+\mathbf{L}^{-1}\mathbf{u}} e^{2\pi i\langle \mathbf{x},\mathbf{L}^{\mathsf{T}}\mathbf{y}\rangle}\, g(\mathbf{x}) \\
&= \sum_{\mathbf{z}\in\mathbb{Z}^n} e^{2\pi i\langle \mathbf{z},\mathbf{L}^{-1}\mathbf{u}\rangle}\, \hat{g}(\mathbf{L}^{\mathsf{T}}\mathbf{y}+\mathbf{z}) \\
&= \frac{1}{\det \mathbf{L}} \sum_{\mathbf{z}\in\mathbb{Z}^n} e^{2\pi i\langle \mathbf{L}^{-\mathsf{T}}\mathbf{z},\mathbf{u}\rangle}\, \widehat{f}(\mathbf{L}^{-\mathsf{T}}(\mathbf{L}^{\mathsf{T}}\mathbf{y}+\mathbf{z})) \\
&= \frac{1}{\det \mathbf{L}} \sum_{\mathbf{L}^{\mathsf{T}}\mathbf{z}\in\mathbb{Z}^n} e^{2\pi i\langle \mathbf{z},\mathbf{u}\rangle}\, \widehat{f}(\mathbf{y}+\mathbf{z})
\end{aligned}$$

It is not hard to see that the vectors $\mathbf{z} \in \mathbb{R}^n$ satisfying $\mathbf{L}^{\mathsf{T}}\mathbf{z} \in \mathbb{Z}^n$ are precisely the vectors in L^*. Therefore

$$\sum_{\mathbf{x}\in L+\mathbf{u}} e^{2\pi i\langle \mathbf{x},\mathbf{y}\rangle}\, f(\mathbf{x}) = \frac{1}{d(L)} \sum_{\mathbf{z}\in L^*} e^{2\pi i\langle \mathbf{z},\mathbf{u}\rangle}\, \widehat{f}(\mathbf{y}+\mathbf{z}).$$

2.4 Positive-definite Functions

We additionally need the following result:

Definition 26. A continuous function $f(\mathbf{x}) : \mathbb{R}^n \mapsto \mathbb{C}$ is called *positive-definite*, if $f(\mathbf{x}) = \overline{f(-\mathbf{x})}$, and for every $\zeta_1, \cdots, \zeta_k \in \mathbb{C}$ and every $\mathbf{x}_1, \cdots, \mathbf{x}_k \in \mathbb{R}^n$ we have

$$\sum_{1 \le i,j \le k} \zeta_i \overline{\zeta_j} \, f(\mathbf{x}_i - \mathbf{x}_j) \ge 0.$$

Theorem 27. If $f(\mathbf{x}) \in \mathbb{R}$ and $f(\mathbf{x}) \ge 0$ for every $\mathbf{x} \in \mathbb{R}^n$, then $\widehat{f}(\mathbf{y})$ is a positive-definite function.

Proof: The Fourier transform of any function $f(\mathbf{x})$ satisfies $\widehat{f}(-\mathbf{x}) = \overline{\widehat{f}(\mathbf{x})}$. For any $\mathbf{x}, \mathbf{y}_1, \cdots, \mathbf{y}_k \in \mathbb{R}^n$ and any $\zeta_1, \cdots, \zeta_k \in \mathbb{C}$, define $c_p = \zeta_p \, e^{2\pi i \langle \mathbf{x}, \mathbf{y}_p \rangle}$, and we have

$$\sum_{1 \le p,q \le k} \zeta_p \overline{\zeta_q} \, e^{2\pi i \langle \mathbf{x}, \mathbf{y}_p - \mathbf{y}_q \rangle} = \sum_{1 \le p,q \le k} c_p \overline{c_q} = \left(\sum_{1 \le p \le k} c_p \right) \overline{\left(\sum_{1 \le p \le k} c_p \right)} \ge 0.$$

Hence

$$\begin{aligned} \sum_{1 \le p,q \le k} \zeta_p \overline{\zeta_q} \, \widehat{f}(\mathbf{y}_i - \mathbf{y}_j) &= \sum_{1 \le p,q \le k} \zeta_p \overline{\zeta_q} \int_{\mathbb{R}^n} e^{2\pi i \langle \mathbf{x}, \mathbf{y}_p - \mathbf{y}_q \rangle} \, f(\mathbf{x}) \, \mathrm{d}\mathbf{x} \\ &= \int_{\mathbb{R}^n} f(\mathbf{x}) \cdot \left(\sum_{1 \le p,q \le k} \zeta_p \overline{\zeta_q} \, e^{2\pi i \langle \mathbf{x}, \mathbf{y}_p - \mathbf{y}_q \rangle} \right) \mathrm{d}\mathbf{x} \\ &\ge 0. \end{aligned}$$

□

3 A Lemma of Banaszczyk

Let $\mathbf{\Sigma}$ be a symmetric positive definite matrix in $\mathbb{R}^{n \times n}$. The Gaussian distribution $\mathcal{D}$ on $\mathbb{R}^n$ with covariance $\mathbf{\Sigma}$ is a continuous distribution with density

$$\mathcal{D}(\mathbf{x}) = \frac{\exp\left(-\frac{1}{2} \mathbf{x}^\top \mathbf{\Sigma}^{-1} \mathbf{x}\right)}{(2\pi)^{n/2} \cdot \det \mathbf{\Sigma}}.$$

If L is a lattice on $\mathbb{R}^n$ then L is a discrete set, so we cannot impose the continuous Gaussian distribution onto L. Instead, Micciancio and Regev (2004) suggest to use the discrete Gaussian measure

$$\rho_{\mathbf{\Sigma}}(\mathbf{x}) = \frac{\exp(-\mathbf{x}^\top \mathbf{\Sigma}^{-1} \mathbf{x})}{\sum_{\mathbf{x} \in L} \exp(-\mathbf{x}^\top \mathbf{\Sigma}^{-1} \mathbf{x})},$$

from which they derived elegant security properties. In this section we prove a lemma from Banaszczyk (1993), which sets the foundation for the analysis of discrete Gaussian distributions.

3.1 Fourier Transform of Gaussian Functions

Lemma 28. For any $a > 0$ we have

$$\int_{-\infty}^{\infty} e^{-ax^2} = \sqrt{\pi/a}.$$

Proof: Simon (2015, Theorem 4.11.11, p. 286). □

Lemma 29. For any $a > 0$, the Fourier transform of $f(x) = e^{-ax^2}$ is

$$\widehat{f}(y) = \sqrt{\pi/a}\, e^{-\pi^2 y^2/a}.$$

Proof:

$$\begin{aligned}
\widehat{f}(y) &= \int_{-\infty}^{\infty} e^{-ax^2}\, e^{2\pi i x y}\, \mathrm{d}x \\
&= \int_{-\infty}^{\infty} e^{-ax^2}\, [\cos(2\pi x y) + i \sin(2\pi x y)]\, \mathrm{d}x \\
&= \int_{-\infty}^{\infty} e^{-ax^2} \cos(2\pi x y)\, \mathrm{d}x \\
&= \int_{-\infty}^{\infty} \left[e^{-ax^2} \sum_{k=0}^{\infty} \frac{(-1)^k (2\pi x y)^{2k}}{(2k)!} \right] \mathrm{d}x \\
&= \sum_{k=0}^{\infty} \left[\frac{(-4\pi^2 y^2)^k}{(2k)!} \int_{-\infty}^{\infty} e^{-ax^2} x^{2k}\, \mathrm{d}x \right].
\end{aligned}$$

Define $f_k(x) = e^{-ax^2}\, x^{2k}$. Integrating by parts gives us

$$\begin{aligned}
\int f_n(x)\, \mathrm{d}x &= e^{-ax^2}\, x^{2k+1} - \int x \cdot e^{-ax^2}\, (2k\, x^{2k-1} - 2a\, x^{2k+1})\, \mathrm{d}x \\
&= e^{-ax^2}\, x^{2k+1} - 2n \int f_k(x)\, \mathrm{d}x + 2a \int f_{k+1}(x)\, \mathrm{d}x.
\end{aligned}$$

As $x \to \pm\infty$ we have $e^{-ax^2}\, x^{2k+1} \to 0$. Therefore

$$\int_{-\infty}^{\infty} f_{k+1}(x)\, \mathrm{d}x = \frac{2k+1}{2a} \int_{-\infty}^{\infty} f_k(x)\, \mathrm{d}x.$$

We conclude that

$$\int_{-\infty}^{\infty} f_k(x)\, \mathrm{d}x = \frac{(2k-1)!!}{(2a)^k} \sqrt{\pi/a}.$$

Now we have

$$\begin{aligned}
\widehat{f}(y) &= \sum_{k=0}^{\infty} \left[\frac{(-4\pi^2 y^2)^k}{(2k)!} \int_{-\infty}^{\infty} e^{-ax^2} x^{2k}\, \mathrm{d}x \right] \\
&= \sum_{k=0}^{\infty} \frac{(-4\pi^2 y^2)^k}{(2k)!} \cdot \frac{(2k-1)!!}{(2a)^k} \sqrt{\pi/a}
\end{aligned}$$

$$
\begin{aligned}
&= \sqrt{\pi/a} \sum_{k=0}^{\infty} \frac{(-2\pi^2 y^2/a)^n}{(2k)!!} \\
&= \sqrt{\pi/a} \sum_{k=0}^{\infty} \frac{(-2\pi^2 y^2/a)^n}{2^k \cdot k!} \\
&= \sqrt{\pi/a}\, e^{-\pi^2 y^2/a}.
\end{aligned}
$$

□

Theorem 30. For any symmetric positive definite matrix $\mathbf{\Sigma}$, the Fourier transform of $f(\mathbf{x}) = \exp(-\mathbf{x}^\top \mathbf{\Sigma}^{-1}\mathbf{x})$ is

$$\widehat{f}(\mathbf{y}) = \sqrt{\pi^n \cdot \det \mathbf{\Sigma}}\, \exp(-\pi^2 \mathbf{y}^\top \mathbf{\Sigma}\mathbf{y}).$$

Proof: Since $\mathbf{\Sigma}$ is symmetric and positive definite, there exists an orthonormal basis $\mathbf{v}_1, \cdots, \mathbf{v}_n$ and positive real numbers $\lambda_1, \cdots, \lambda_n$, such that

$$\mathbf{\Sigma} = \sum_{k=1}^{n} \lambda_k \mathbf{v}_k \mathbf{v}_k^\top, \quad \mathbf{\Sigma}^{-1} = \sum_{k=1}^{n} \frac{1}{\lambda_k} \mathbf{v}_k \mathbf{v}_k^\top, \quad \det \mathbf{\Sigma} = \prod_{k=1}^{n} \lambda_k.$$

We now express everything in the $\{\mathbf{v}_k\}$ basis. For example, $y_k = \langle \mathbf{y}, \mathbf{v}_k \rangle$. We have

$$
\begin{aligned}
\widehat{f}(\mathbf{y}) &= \int_{\mathbb{R}^n} e^{2\pi i \langle \mathbf{x}, \mathbf{y} \rangle}\, e^{-\mathbf{x}^\top \mathbf{\Sigma}^{-1} \mathbf{x}}\, \mathrm{d}\mathbf{x} \\
&= \prod_{k=1}^{n} \int_{-\infty}^{\infty} e^{2\pi i x_k y_k}\, e^{-x_k^2/\lambda_k}\, \mathrm{d}x_k \\
&= \prod_{k=1}^{n} \sqrt{\pi \cdot \lambda_k}\, e^{-\pi^2 \lambda_k y_k^2} \\
&= \sqrt{\pi^n \cdot \det \mathbf{\Sigma}}\, \exp(-\pi^2 \mathbf{y}^\top \mathbf{\Sigma}\mathbf{y}).
\end{aligned}
$$

□

3.2 Banaszczyk's Lemma

Let L be a lattice of rank n. If we impose a discrete Gaussian distribution $\mathcal{D}$ onto L, and sample points of L according to $\mathcal{D}$, then the points most likely to be chosen should be points that are close to the origin. In this subsection, we prove a lemma from Banaszczyk (1993) that analytically bounds the probability of getting points that are "far" from the origin. For the remainder of this blogpost, when S is a countable subset of $\mathbb{R}^n$ and s is a positive real number, we write

$$\rho(S) = \sum_{\mathbf{x} \in S} e^{-\pi ||\mathbf{x}||^2},$$

$$\rho_s(S) = \sum_{\mathbf{x} \in S} e^{-\pi ||\mathbf{x}||^2/s^2}.$$

Lemma 31. For any $a > 0$ and $\mathbf{u}, \mathbf{y} \in \mathbb{R}^n$ we have

$$\sum_{\mathbf{x}\in L+\mathbf{u}} e^{2\pi i\langle \mathbf{x},\mathbf{y}\rangle}\, e^{-a||\mathbf{x}||^2} = \frac{\sqrt{\pi^n/a^n}}{d(L)} \sum_{\mathbf{z}\in L^*} e^{-2\pi i\langle \mathbf{z},\mathbf{u}\rangle}\, e^{-\pi^2||\mathbf{y}+\mathbf{z}||^2/a}.$$

Proof: This is a straightforward consequence of Theorem 15. □

Remark 32. In Lemma 31, if we set $\mathbf{y} = \mathbf{0}$, we get

$$\sum_{\mathbf{x}\in L+\mathbf{u}} e^{-a||\mathbf{x}||^2} = \frac{\sqrt{\pi^n/a^n}}{d(L)} \sum_{\mathbf{z}\in L^*} e^{-2\pi i\langle \mathbf{z},\mathbf{u}\rangle}\, e^{-\pi^2||\mathbf{z}||^2/a}.$$

Setting also $\mathbf{u} = \mathbf{0}$ gives us

$$\sum_{\mathbf{x}\in L} e^{-a||\mathbf{x}||^2} = \frac{\sqrt{\pi^n/a^n}}{d(L)} \sum_{\mathbf{z}\in L^*} e^{-\pi^2||\mathbf{z}||^2/a}.$$

Thus for any $\mathbf{u} \in \mathbb{R}^n$ we have

$$\begin{aligned}
\sum_{\mathbf{x}\in L+\mathbf{u}} e^{-a||\mathbf{x}||^2} &= \left| \frac{\sqrt{\pi^n/a^n}}{d(L)} \sum_{\mathbf{z}\in L^*} e^{-2\pi i\langle \mathbf{z},\mathbf{u}\rangle}\, e^{-\pi^2||\mathbf{z}||^2/a} \right| \\
&\leq \frac{\sqrt{\pi^n/a^n}}{d(L)} \sum_{\mathbf{z}\in L^*} \left| e^{-2\pi i\langle \mathbf{z},\mathbf{u}\rangle}\, e^{-\pi^2||\mathbf{z}||^2/a} \right| \\
&= \frac{\sqrt{\pi^n/a^n}}{d(L)} \sum_{\mathbf{z}\in L^*} e^{-\pi^2||\mathbf{z}||^2/a} \\
&= \sum_{\mathbf{x}\in L} e^{-a||\mathbf{x}||^2}.
\end{aligned}$$

On the other hand, we also have

$$\begin{aligned}
\sum_{\mathbf{x}\in L+\mathbf{u}} e^{-a||\mathbf{x}||^2} &= \sum_{\mathbf{x}\in L} e^{-a||\mathbf{x}+\mathbf{u}||^2} \\
&= \frac{1}{2} \sum_{\mathbf{x}\in L} \left(e^{-a||\mathbf{u}+\mathbf{x}||^2} + e^{-a||\mathbf{u}-\mathbf{x}||^2} \right) && \text{(Because both } \mathbf{x}, -\mathbf{x} \in L\text{)} \\
&= e^{-a||\mathbf{u}||^2} \sum_{\mathbf{x}\in L} e^{-a||\mathbf{x}||^2}\, \cosh(2a\langle \mathbf{x}, \mathbf{u}\rangle) \\
&\geq e^{-a||\mathbf{u}||^2} \sum_{\mathbf{x}\in L} e^{-a||\mathbf{x}||^2}.
\end{aligned}$$

□

Lemma 33. For any $a > 0$, $\mathbf{u} \in \mathbb{R}^n$, $k \in \{1, \cdots, n\}$, we have

$$\frac{\sum_{\mathbf{x}\in L+\mathbf{u}} x_k^2\, e^{-a||\mathbf{x}||^2}}{\sum_{\mathbf{x}\in L} e^{-a||\mathbf{x}||^2}} \leq 1/a.$$

If $\mathbf{u} = \mathbf{0}$ then the bound can be improved to $1/2a$.

Proof: Let $f(\mathbf{y}) = \sum_{\mathbf{x}\in L+\mathbf{u}} e^{-a||\mathbf{x}||^2}\, e^{2\pi i\langle \mathbf{x},\mathbf{y}\rangle}$. Let $f_{kk} = \partial^2 f/\partial y_k^2$. Then we have

$$f_{kk}(\mathbf{y}) = -4\pi^2 \sum_{\mathbf{x}\in L+\mathbf{u}} x_k^2\, e^{-a||\mathbf{x}||^2}\, e^{2\pi i\langle \mathbf{x},\mathbf{y}\rangle}.$$

Therefore

$$\sum_{\mathbf{x}\in L+\mathbf{u}} x_k^2\, e^{-a||\mathbf{x}||^2} = -\frac{f_{kk}(\mathbf{0})}{4\pi^2}.$$

But by Lemma 31 we also have

$$f(\mathbf{y}) = \frac{\sqrt{\pi^n/a^n}}{d(L)} \sum_{\mathbf{z}\in L^*} e^{-2\pi i\langle \mathbf{z},\mathbf{u}\rangle}\, e^{-\pi^2||\mathbf{y}+\mathbf{z}||^2/a},$$

$$f_{kk}(\mathbf{y}) = \frac{\sqrt{\pi^n/a^n}}{d(L)} \sum_{\mathbf{z}\in L^*} e^{-2\pi i\langle \mathbf{z},\mathbf{u}\rangle}\, e^{-\pi^2||\mathbf{y}+\mathbf{z}||^2/a} \left[4\pi^4(y_k+z_k)^2/a^2 - 2\pi^2/a\right],$$

$$f_{kk}(\mathbf{0}) = \frac{\sqrt{\pi^n/a^n}}{d(L)} \sum_{\mathbf{z}\in L^*} e^{-2\pi i\langle \mathbf{z},\mathbf{u}\rangle}\, e^{-\pi^2||\mathbf{z}||^2/a} \left[4\pi^4 z_k^2/a^2 - 2\pi^2/a\right].$$

Let $g(\mathbf{u}) = \sum_{\mathbf{z}\in L^*} e^{-\pi^2||\mathbf{z}||^2/a}\, e^{2\pi i\langle \mathbf{z},\mathbf{u}\rangle}$. Let $g_{kk} = \partial^2 g/\partial u_k^2$. We have

$$g_{kk}(\mathbf{u}) = -4\pi^2 \sum_{\mathbf{z}\in L^*} z_k^2\, e^{-\pi^2||\mathbf{z}||^2/a}\, e^{2\pi i\langle \mathbf{u},\mathbf{z}\rangle},$$

$$f_{kk}(\mathbf{0}) = \frac{\sqrt{\pi^n/a^n}}{d(L)} \left(-2\pi^2/a \cdot g(-\mathbf{u}) - \pi^2/a^2 \cdot g_{kk}(-\mathbf{u})\right),$$

$$\sum_{\mathbf{x}\in L} e^{-a||\mathbf{x}||^2} = \frac{\sqrt{\pi^n/a^n}}{d(L)} \sum_{\mathbf{z}\in L^*} e^{-\pi||\mathbf{z}||^2} = \frac{\sqrt{\pi^n/a^n}}{d(L)} g(\mathbf{0}).$$

Combining these equations, we get

$$\frac{\sum_{\mathbf{x}\in L+\mathbf{u}} x_k^2\, e^{-a||\mathbf{x}||^2}}{\sum_{\mathbf{x}\in L} e^{-a||\mathbf{x}||^2}} = \frac{g(-\mathbf{u})}{2a\cdot g(\mathbf{0})} + \frac{g_{kk}(-\mathbf{u})}{4a^2\cdot g(\mathbf{0})}.$$

In Remark 32, we showed that $g(-\mathbf{u}) \le g(\mathbf{0})$. We also have $g_{kk}(\mathbf{0}) < 0$. When $\mathbf{u} = \mathbf{0}$, this is sufficient to show that

$$\frac{\sum_{\mathbf{x}\in L} x_k^2\, e^{-a||\mathbf{x}||^2}}{\sum_{\mathbf{x}\in L} e^{-a||\mathbf{x}||^2}} \le \frac{1}{2a}.$$

By Theorem 27, $-g_{kk}$ is a positive-definite function. We have

$$-2g_{kk}(-\mathbf{u}) = -g_{kk}(-\mathbf{u}) - g_{kk}(\mathbf{u}) \ge 2g_{kk}(\mathbf{0}).$$

It is equivalent to $g_{kk}(-\mathbf{u}) \le -g_{kk}(\mathbf{0})$.

In Remark 32 we showed that $g(\mathbf{u})/g(\mathbf{0}) \ge e^{-a||\mathbf{u}||^2}$. Let $g_k = \partial g/\partial u_k$ and notice that $g_k(\mathbf{0}) = 0$. It follows that $g_{kk}(\mathbf{0})/g(\mathbf{0}) \ge -2a$, because otherwise $g(\mathbf{u})$ would decrease too quickly in a neighborhood around $\mathbf{0}$ and $g(\mathbf{u})/g(\mathbf{0}) \ge e^{-a||\mathbf{u}||^2}$ cannot hold.

We conclude that

$$\frac{g_{kk}(-\mathbf{u})}{4a^2 \cdot g(\mathbf{0})} \leq -\frac{-g_{kk}(\mathbf{0})}{4a^2 \cdot g(\mathbf{0})} \leq \frac{2a}{4a^2} = \frac{1}{2a},$$

$$\frac{\sum_{\mathbf{x}\in L+\mathbf{u}} x_k^2\, e^{-a||\mathbf{x}||^2}}{\sum_{\mathbf{x}\in L} e^{-a||\mathbf{x}||^2}} \leq \frac{1}{2a} + \frac{1}{2a} = \frac{1}{a}.$$

□

Lemma 34. For any $s \geq 1$ we have

$$\rho_s(L) \leq s^n \cdot \rho(L).$$

Proof: Let $f(a) = \rho_{\sqrt{a}}(L) = \sum_{\mathbf{x}\in L} e^{-\pi||\mathbf{x}||^2/a}$. We have

$$\begin{aligned} f'(a) &= \frac{\pi}{a^2} \sum_{\mathbf{x}\in L} ||\mathbf{x}||^2 e^{-\pi||\mathbf{x}||^2/a} \\ &= \frac{\pi}{a^2} \sum_{\mathbf{x}\in L} \sum_{k=1}^{n} x_k^2\, e^{-\pi||\mathbf{x}||^2/a} \\ &\leq \frac{\pi}{a^2} \sum_{k=1}^{n} \frac{a}{2\pi} \sum_{\mathbf{x}\in L} e^{-\pi||\mathbf{x}||^2/a} \\ &= \frac{n}{2a} \sum_{\mathbf{x}\in L} e^{-\pi||\mathbf{x}||^2/a} \\ &= \frac{n}{2a} f(a). \end{aligned}$$

Therefore $[\log f(a)]' \leq n/2a$. Integrating along a gives us $\log[f(a)/f(1)] \leq \log(a) \cdot n/2$, so $f(a)/f(1) \leq a^{n/2}$. Now $\rho_s(L)/\rho(L) = f(s^2)/f(1) \leq s^n$. □

Lemma 35. For any $s \geq 1$ and any $\mathbf{u} \in \mathbb{R}^n$ we have

$$\rho_s(L + \mathbf{u}) \leq 2s^n \cdot \rho(L).$$

Proof: Let $f(a) = \rho_{\sqrt{a}}(L) = \sum_{\mathbf{x}\in L} e^{-\pi||\mathbf{x}||^2/a}$. Let $g(a) = \rho_{\sqrt{a}}(L + \mathbf{u})$. We have

$$\begin{aligned} g'(a) &= \frac{\pi}{a^2} \sum_{\mathbf{x}\in L+\mathbf{u}} ||\mathbf{x}||^2 e^{-\pi||\mathbf{x}||^2/a} \\ &= \frac{\pi}{a^2} \sum_{\mathbf{x}\in L+\mathbf{u}} \sum_{k=1}^{n} x_k^2\, e^{-\pi||\mathbf{x}||^2/a} \\ &\leq \frac{\pi}{a^2} \sum_{k=1}^{n} \frac{a}{\pi} \sum_{\mathbf{x}\in L} e^{-\pi||\mathbf{x}||^2/a} \\ &= \frac{n}{a} \sum_{\mathbf{x}\in L} e^{-\pi||\mathbf{x}||^2/a} \\ &\leq n a^{n/2-1} f(1). \end{aligned}$$

Therefore

$$g(a) - g(1) \leq n \cdot f(1) \int_1^a t^{n/2-1}\, \mathrm{d}t = 2(a^{n/2} - 1) f(1).$$

By Remark 32 we have $g(1) \leq f(1)$. This finishes the proof. □

Lemma 36 (Banaszczyk's Bound). For any $c \geq 1/\sqrt{2\pi}$, let $\mathcal{B}(c)$ be the open ball

$$\mathcal{B}(c) = \{\mathbf{x} \in L \mid ||\mathbf{x}|| < c\sqrt{n}\},$$

then we have

$$\rho(L \setminus \mathcal{B}(c))/\rho(L) < \left[c\sqrt{2\pi e}\, e^{-\pi c^2}\right]^n.$$

Proof: For any $t \in (0,1)$ we have

$$\begin{aligned}\sum_{\mathbf{x}\in L} e^{-\pi t||\mathbf{x}||^2} &= \sum_{\mathbf{x}\in L} e^{\pi(1-t)||\mathbf{x}||^2}\, e^{-\pi||\mathbf{x}||^2} \\ &\geq \sum_{\substack{\mathbf{x}\in L \\ ||\mathbf{x}||^2 \geq c^2 n}} e^{\pi(1-t)||\mathbf{x}||^2}\, e^{-\pi||\mathbf{x}||^2} \\ &> e^{\pi(1-t)c^2 n} \sum_{\substack{\mathbf{x}\in L \\ ||\mathbf{x}||^2 \geq c^2 n}} e^{-\pi||\mathbf{x}||^2}.\end{aligned}$$

By Lemma 34 we also have

$$\sum_{\mathbf{x}\in L} e^{-\pi t||\mathbf{x}||^2} \leq t^{-n/2} \sum_{\mathbf{x}\in L} e^{-\pi||\mathbf{x}||^2}.$$

Therefore

$$\sum_{\substack{\mathbf{x}\in L \\ ||\mathbf{x}||^2 \geq c^2 n}} e^{-\pi||\mathbf{x}||^2} < t^{-n/2}\, e^{-\pi(1-t)c^2 n} \sum_{\mathbf{x}\in L} e^{-\pi||\mathbf{x}||^2}.$$

This can be written as

$$\rho(L \setminus \mathcal{B}(c))/\rho(L) < \left[t^{-1/2}\, e^{-\pi(1-t)c^2}\right]^n.$$

Set $t = 1/2\pi c^2$ and we get

$$\rho(L \setminus \mathcal{B}(c))/\rho(L) < \left[c\sqrt{2\pi e} \cdot e^{-\pi c^2}\right]^n.$$

□

4 The Smoothing Parameter

For any lattice L, we have

$$\rho_s(L) = 1 + \rho_s(L \setminus \{\mathbf{0}\}).$$

For any $\mathbf{c} \in \mathbb{R}^n$, by Lemma 31 we have

$$\rho_s(L + \mathbf{c}) = \frac{s^n}{d(L)} \sum_{\mathbf{z}\in L^*} e^{-2\pi i\langle \mathbf{c},\mathbf{z}\rangle}\, e^{-\pi s^2||\mathbf{z}||^2}.$$

Now suppose that $\rho_{1/s}(L^* \setminus \{\mathbf{0}\})$ is a very small value. Then the coefficient $e^{-2\pi i\langle \mathbf{c},\mathbf{z}\rangle}$ will not have a significant effect on the value of $\rho_s(L + \mathbf{c})$. Formally:

$$\rho_{1/s}(L^* \setminus \{\mathbf{0}\}) \leq \varepsilon \Rightarrow (1-\varepsilon)\frac{s^n}{d(L)} \leq \rho_s(L + \mathbf{c}) \leq (1+\varepsilon)\frac{s^n}{d(L)}.$$

Notice that the range of $\rho_s(L + \mathbf{c})$ does not depend on $\mathbf{c}$.

The significance of this inequality is as follows. If L' is a sublattice of L, then we can define an equivalence relation $\mathbf{v} \leftrightarrow \mathbf{v}'$ on L, determined by

$$\mathbf{v} \leftrightarrow \mathbf{v}' \equiv \mathbf{v} - \mathbf{v}' \in L'.$$

Each equivalence class of this relation can be written as $L' + \mathbf{c}$ for some $\mathbf{c} \in L$. Now if $\rho_s(L' + \mathbf{c})$ is roughly equal to a constant for each of these equivalence classes, then we can sample points from L according to the distribution $\rho_s(L)$, and the probability of getting points from each of these equivalence classes is roughly uniform. This is the core idea behind the regularity lemma.

Now how should we set the parameter s so that the value $\rho_{1/s}(L^* \setminus \{\mathbf{0}\})$ becomes negligible? To answer this question, Micciancio and Regev (2004) introduced the "smoothing parameter" and related it to other lattice properties.

Definition 37. For a given lattice L and $\varepsilon > 0$, the smallest positive real number s that satisfies

$$\rho_{1/s}(L^* \setminus \{\mathbf{0}\}) \leq \varepsilon$$

is called the *smoothing parameter* of L, denoted by $\eta_\varepsilon(L)$.

For a given lattice L, we use $\lambda_1(L)$ to denote the length of the shortest non-zero vector in L.

Lemma 38. If L is a lattice of rank n, and $\varepsilon = 2^{-2n}$, then $\eta_\varepsilon(L) \leq \sqrt{n}/\lambda_1(L^*)$.

Proof: If $s > \sqrt{n}/\lambda_1(L^*)$, then the only vector $\mathbf{v}$ in sL^* with $||\mathbf{v}|| < \sqrt{n}$ is $\mathbf{0}$. Then by Lemma 36 we have

$$\rho_{1/s}(L^* \setminus \{\mathbf{0}\}) = \rho(sL^* \setminus \{\mathbf{0}\}) < C^n \cdot \rho(sL^*) = C^n \cdot (1 + \rho(sL^* \setminus \{\mathbf{0}\}))$$

where $C = \sqrt{2\pi e} \cdot e^{-\pi} < 1/4$. Therefore

$$\rho_{1/s}(L^* \setminus \{\mathbf{0}\}) < \frac{C^n}{1 - C^n} < 2^{-2n}.$$

□

Lemma 39. For a given lattice L of rank n and any $s > 0, \varepsilon > 0$, we have

$$\rho_{1/s}(L) \leq \max\left(1, \left(\frac{\eta_\varepsilon(L^*)}{s}\right)^n\right)(1 + \varepsilon).$$

Proof: If $s \geq \eta_\varepsilon(L^*)$ then $\rho_{1/s}(L) \leq 1 + \varepsilon$ by the definition of $\eta_\varepsilon(L^*)$. If $s < \eta_\varepsilon(L^*)$ then let $\eta = \eta_\varepsilon(L^*)$ and

$$\rho_{1/s}(L) = \frac{s^{-n}}{d(L)}\rho_s(L^*) < \frac{s^{-n}}{d(L)}\rho_\eta(L^*) = (\eta/s)^n \cdot \rho_{1/\eta}(L) \leq (\eta/s)^n \cdot (1 + \varepsilon).$$

□

5 Algebraic Number Theory

In Ring-LWE the lattices under study are algebraically structured. They in fact correspond to ideals in rings of algebraic numbers. In this section we present the basic properties of these ideals. Jarvis (2014) is a good introduction to the theory, but only covers the basics. Neukirch (1999) presents the results in greater generality, but is more

difficult to read.

Let $P(X)$ be a polynomial with coefficients in $\mathbb{Q}$, such that $P(X)$ is irreducible over $\mathbb{Q}$. Suppose that

$$P(X) = c_0 + c_1X + \cdots + c_nX^n.$$

The *formal derivative* of $P(X)$ is defined to be

$$P'(X) = c_1 + 2c_2X + \cdots + nc_nX^{n-1}.$$

Since $P(X)$ is irreducible, it is coprime with its formal derivative. This is equivalent to $P(X)$ having no repeated roots over $\mathbb{C}$. Then $\mathbb{Q}[X]/P(X)$ is a field, and is called a finite and separable extension of $\mathbb{Q}$. Elements of $\mathbb{Q}[X]/P(X)$ can be represented as polynomials in X of degree at most $n-1$ and with rational coefficients. To avoid confusion, when representing elements of K we shall use γ to represent the placeholder variable of the polynomial, and leave $X, Y, \cdots$ for variables of other polynomials. We denote the field $\mathbb{Q}[X]/P(X)$ by K.

Definition 40. A non-zero rational polynomial $m(X)$ is an *annihilating polynomial* for an element $\alpha \in K$, if $m(\alpha) = 0$ in K.

Remark 41. If $m(X)$ is an annihilating polynomial for some $\alpha \in K$ then $\deg m \geq 1$. If $\deg m = 0$, then $m(X) = c$ for some $c \in \mathbb{Q}$, but $m(\alpha) = 0$ so $c = 0$. This contradicts the requirement that $m(X)$ is non-zero.

Definition 42. A non-zero rational polynomial $m(X)$ is *monic* if the leading coefficient of $m(X)$ is 1.

Lemma 43. For each $\alpha \in K$, there exists a unique monic annihilating polynomial $m(X)$ for α that has the lowest degree among all annihilating polynomials for α. We call $m(X)$ the *minimal polynomial* of α.

Proof: Jarvis (2014, Lemma 2.4, p. 20). □

Lemma 44. The minimal polynomial of any $\alpha \in K$ is irreducible.

Proof: Jarvis (2014, Lemma 2.6, p. 20). □

Lemma 45. If $m(X)$ is an annihilating polynomial of some $\alpha \in K$, then it is a multiple of the minimal polynomial of α.

Proof: Jarvis (2014, Lemma 2.7, p. 21). □

Definition 46. An element $\alpha \in K$ is called an *algebraic integer* if the minimal polynomial of α has only integer coefficients.

Lemma 47. An element $\alpha \in K$ is an algebraic integer iff it is the root of a monic polynomial with integer coefficients (not necessarily minimal).

Proof: Jarvis (2014, Lemma 2.22, p. 29). Suppose that $m'(X)$ is a monic polynomial with integer coefficients such that $m'(\alpha) = 0$. Then $m'(X)$ is a multiple of the minimal polynomial $m(X)$ of α, so we can write $m'(X) = m(X) \cdot g(X)$. Since $m(X)$ and $m'(X)$ are monic, so is $g(X)$. Suppose that some coefficients of $m(X)$ and $g(X)$ are not integers. Let a be the least common multiple (LCM) of the denominators of non-zero coefficients in $m(X)$. Let b be the LCM of the denominators of non-zero coefficients in $g(X)$. Then the polynomials $am(X)$ and $bg(X)$ have only integer coefficients.

The greatest common divisor (GCD) of the non-zero coefficients of $am(X)$ must be 1. Since $m(X)$ is monic, the leading coefficient of $am(X)$ is just a. If $a > 1$, let p be a prime factor of a, and let k be the highest integer such that p^k divides a. Then at least one non-zero coefficient c of $m(X)$ has a denominator that is a multiple of p^k, and the

numerator of that coefficient is not a multiple of p. We see that ac cannot be a multiple of p. Similarly, the GCD of the non-zero coefficients of $bg(X)$ must be 1.

The GCD of the non-zero coefficients of $am(X) \cdot bg(X)$ must be 1. If this is not the case, let p be a prime factor of the GCD. Suppose that

$$am(X) = m_0 + m_1 X + \cdots + m_r X^r, \quad bg(X) = g_0 + g_1 X + \cdots + g_s X^s,$$

$$am(X) \cdot bg(X) = \sum_{i=0}^{r+s} \sum_{j=0}^{i} m_j g_{i-j} X^i.$$

Since the GCD of the non-zero coefficients of $am(X)$ is 1, not all coefficients of $am(X)$ are multiples of p. Let $m_u X^u$ be the highest term of $am(X)$ that is not a multiple of p. Similarly, let $g_v X^v$ be the highest term of $bg(X)$ that is not a multiple of p. Consider the coefficient of X^{u+v} in $am(X) \cdot bg(X)$, which is equal to $\sum_{j=0}^{u+v} m_j g_{u+v-j}$. The terms having $j < u$ can be dropped since m_j is a multiple of p. The terms having $u + v - j < v$ can also be dropped since m_{u+v-j} is a multiple of p. The only remaining term is $m_u g_v$, which is not a multiple of p. We conclude that the coefficient of X^{u+v} cannot be a multiple of p. This contradicts the assumption that p is a factor of every coefficient of $am(X) \cdot bg(X)$.

Recall that we have assumed $m'(X) = m(X) \cdot g(X)$, so $am(X) \cdot bg(X) = ab \cdot m'(X)$. Since $m'(X)$ has only integer coefficients, ab is a factor of every coefficient of $am(X) \cdot bg(X)$. We must have $a = b = 1$. Hence both $m(X)$ and $g(X)$ have only integer coefficients. □

Neukirch (1999, p. 8) gives a different proof of this lemma, but Neukirch (1999) uses a different definition of algebraic integers, and proves that it is equivalent to Definition 46.

Lemma 48. For every $\alpha \in K$ there exists an integer k such that $k\alpha$ is an algebraic integer.

Proof: Neukirch (1999, p. 8). □

5.1 Algebraic Integers as a Ring

Let $\alpha_1, \cdots, \alpha_k$ be a finite number of elements in K. Then $\mathcal{R} = \mathbb{Z}[\alpha_1, \cdots, \alpha_k]$ is a subring of K. Each element in $\mathcal{R}$ can be written as a finite sum of monomials

$$z\alpha_1^{r_1} \cdots \alpha_k^{r_k}$$

where $z \in \mathbb{Z}, r_1, \cdots, r_k \in \mathbb{N}$. We say $\mathcal{R}$ is finitely generated, if there exists a finite number of elements $\omega_1, \cdots, \omega_l \in \mathcal{R}$, such that every element in $\mathcal{R}$ can be written as

$$r_1\omega_1 + \cdots + r_l\omega_l$$

where $r_1, \cdots, r_l \in \mathbb{Z}$.

Theorem 49. The following three propositions are equivalent:

1. $\mathbb{Z}[\alpha_1, \cdots, \alpha_k]$ is finitely generated;
2. The elements $\alpha_1, \cdots, \alpha_k$ are all algebraic integers;
3. All elements in $\mathbb{Z}[\alpha_1, \cdots, \alpha_k]$ are algebraic integers.

Proof: Jarvis (2014, Theorem 2.25, Corollary 2.26, Proposition 2.27, pp. 31–32). Let $\mathcal{R} = \mathbb{Z}[\alpha_1, \cdots, \alpha_k]$.

(3) ⇒ (2): Obvious.

(2) ⇒ (1): Let $m_i(X)$ be the minimal polynomial of α_i. Then $m_i(\alpha_i) = 0$, and $\alpha_i^{\deg m_i} = (X^{\deg m_i} - m_i(X))(\alpha_i)$. Notice that $\deg[X^{\deg m_i} - m_i(X)] < \deg m_i$. Therefore, each α_i^r with $r \geq \deg m_i$ can be expressed as an integral combination of $1, \alpha_i, \cdots, \alpha_i^{\deg m_i - 1}$. Hence each monomial in $\mathcal{R}$ can be expressed as an integral combination of monomials

$$\alpha_1^{r_1} \cdots \alpha_k^{r_k}$$

with $r_i \in \{0, 1, \cdots, \deg m_i - 1\}$. There are only a finite number of such monomials. Therefore, $\mathcal{R}$ is finitely generated.

(1) ⇒ (3): Let $\omega_1, \cdots, \omega_l$ be an integral basis of $\mathcal{R}$. For a given $\alpha \in \mathcal{R}$, the mapping $\beta \mapsto \alpha\beta$ is linear. Suppose that for each ω_i we have

$$\alpha\omega_i = \sum_{j=1}^{l} r_{ij}\omega_j.$$

Define the matrix

$$\mathbf{M}_\alpha = \begin{pmatrix} r_{11} & \cdots & r_{l1} \\ \vdots & \ddots & \vdots \\ 1l & \cdots & r_{ll} \end{pmatrix}.$$

Then for every $\beta \in \mathcal{R}$, if $\beta = s_1\omega_1 + \cdots + s_l\omega_l$, we have

$$\alpha\beta = \begin{pmatrix} \omega_1 & \cdots & \omega_l \end{pmatrix} \mathbf{M}_\alpha \begin{pmatrix} s_1 \\ \vdots \\ s_l \end{pmatrix}.$$

Furthermore, if $f(X)$ is a polynomial, then

$$f(\alpha) \cdot \beta = \begin{pmatrix} \omega_1 & \cdots & \omega_l \end{pmatrix} f(\mathbf{M}_\alpha) \begin{pmatrix} s_1 \\ \vdots \\ s_l \end{pmatrix}.$$

Let $F(X)$ be the characteristic polynomial of $\mathbf{M}_\alpha$, which is a monic polynomial of degree l with integer coefficients. It is known the characteristic polynomial of a matrix is an annihilating polynomial of that matrix (Caylay-Hamilton theorem). Hence $F(\alpha) = 0$, and α is an algebraic integer by Lemma 47. □

Theorem 50. The set of algebraic integers forms a subring of K.

Proof: If α, β are two algebraic integers in K, then $\alpha + \beta, \alpha\beta \in \mathbb{Z}[\alpha, \beta]$, and they are both algebraic integers by Theorem 49. □

5.2 Extension Field as a Vector Space

The field K can be seen as an n-dimensional vector space over $\mathbb{Q}$, and $\{1, \gamma, \gamma^2, \cdots, \gamma^{n-1}\}$ is a basis for this vector space. In this section we analyze the structure of bases for this vector space.

Lemma 51. The minimal polynomial of any $\alpha \in K$ has degree at most n.

Proof: The vectors corresponding to $1, \alpha, \alpha^2, \cdots, \alpha^n$ cannot be linearly independent, since the dimension of the vector space is n. Hence we can express 0 as a non-trivial linear combination of $1, \alpha, \cdots, \alpha^n$. □

For a given $\alpha \in K$, let $m(X)$ be the minimal polynomial of α and let $r = \deg m(X)$. If $r = 1$ then $\alpha \in \mathbb{Q}$ and we ignore this case. The vectors corresponding to $1, \alpha, \cdots, \alpha^{r-1}$ are linearly independent and span a subspace of K. They do not span the full space unless $r = n$. Let $f(X)$ be a polynomial of degree at most $r - 1$. As $m(X)$ is irreducible, we

can find polynomials $g(X), h(X)$ such that $f(X) \cdot g(X) = m(X) \cdot h(X) + 1$. Then $f(\alpha) \cdot g(\alpha) = 1$. As such, the inverse of any element spanned by $1, \alpha, \cdots, \alpha^{r-1}$ is also spanned by $1, \alpha, \cdots, \alpha^{r-1}$. The elements that can be written as a polynomial in α form a subfield of K. We denote this field by $\mathbb{Q}(\alpha)$.

Let β be an element of K not in $\mathbb{Q}(\alpha)$. We claim that $1, \alpha, \cdots, \alpha^{r-1}, \beta, \alpha\beta, \cdots, \alpha^{r-1}\beta$ are linearly independent. If this is not the case, we express 0 as a non-trivial linear combination of these elements. The combination must involve at least one of $\beta, \cdots, \alpha^{r-1}\beta$, since $1, \cdots, \alpha^{r-1}$ are linearly independent. Then we can factor out β and write the combination as

$$\beta \cdot F(\alpha) = G(\alpha) \Rightarrow \beta = \frac{G(\alpha)}{F(\alpha)},$$

where $F(X), G(X)$ are polynomials of degree at most $r-1$. This means β is already in $\mathbb{Q}(\alpha)$.

Now let k be the least positive integer such that the vectors

$$1, \cdots, \alpha^{r-1}, \beta, \cdots, \alpha^{r-1}\beta, \cdots, \beta^k, \cdots, \alpha^{r-1}\beta^k$$

are linearly dependent. We have just shown that $k \geq 2$. We claim that the set of linearly independent vectors

$$S = \{1, \cdots, \alpha^{r-1}, \beta, \cdots, \alpha^{r-1}\beta, \cdots, \beta^{k-1}, \cdots, \alpha^{r-1}\beta^{k-1}\}$$

spans a larger subfield of K. Since adding $\beta^k, \cdots, \alpha^{r-1}\beta^{k-1}$ causes the set to become linearly dependent, we can factor out β^k and write

$$\beta^k \cdot F(\alpha) = G(\alpha, \beta) = g_0 + g_1\beta + \cdots + g_{k-1}\beta^{k-1}$$

where $g_0, \cdots, g_{k-1} \in \mathbb{Q}(\alpha)$. Then we have

$$\beta^k = \frac{g_0}{F(\alpha)} + \frac{g_1}{F(\alpha)}\beta + \cdots + \frac{g_{k-1}}{F(\alpha)}\beta^{k-1},$$

where each of the fractions is in $\mathbb{Q}(\alpha)$ and the sum can be expressed as a linear combination of elements in S. We may call the polynomial

$$m_\beta(X) = X^k - \frac{g_{k-1}}{F(\alpha)}X^{k-1} - \cdots - \frac{g_0}{F(\alpha)}$$

the minimal polynomial of β **over** $\mathbb{Q}(\alpha)$. It is not the same thing as the minimal polynomial of β (**over** $\mathbb{Q}$), since we are allowing coefficients in $\mathbb{Q}(\alpha)$.

A simple corollary of the above result is that each β^q with $q \geq k$ can be expressed as a linear combination of elements in S. We now look at how to express the inverse elements. Let $m_\beta(X)$ be the minimal polynomial of β over $\mathbb{Q}(\alpha)$, as defined above, and let $k = \deg m_\beta(X)$. Then $m_\beta(X)$ is irreducible over $\mathbb{Q}(\alpha)$, otherwise $m_\beta(X)$ would not be minimal. Since $\mathbb{Q}(\alpha)$ is a field, we can do polynomial long division in $\mathbb{Q}(\alpha)$. Thus for each polynomial $F(X)$ of degree at most $k-1$ and with coefficients in $\mathbb{Q}(\alpha)$, we can find polynomials $G(X), H(X)$ (with coefficients in $\mathbb{Q}(\alpha)$) such that $F(X) \cdot G(X) = m_\beta(X) \cdot H(X) + 1$. This implies $G(\beta) = 1/F(\beta)$.

We have thus constructed a new subfield $\mathbb{Q}(\alpha, \beta)$ of K. It is strictly larger than $\mathbb{Q}(\alpha)$ because $\beta \notin \mathbb{Q}(\alpha)$. If $\mathbb{Q}(\alpha, \beta) \neq K$, then we can repeat the above construction to construct further larger subfields of K. Notice that the basis of $\mathbb{Q}(\alpha, \beta)$ (as a vector space over $\mathbb{Q}$) contains rk elements, which is a multiple of r. If one repeats the above construction to construct another subfield $\mathbb{Q}(\alpha, \beta, \delta)$, then the basis of $\mathbb{Q}(\alpha, \beta, \delta)$ (as a vector space over $\mathbb{Q}$) would contain rkt elements for some positive integer $t \geq 2$. This leads us to the following lemma:

Lemma 52. If $m(X)$ is the minimal polynomial of some $\alpha \in K$, then $\deg m$ is a factor of n.

Proof: Starting from the subfield $\mathbb{Q}(\alpha)$, repeat the subfield extension procedure described above until $\mathbb{Q}(\alpha, \beta, \cdots) =$

K. Then the basis of $\mathbb{Q}(\alpha, \beta, \cdots)$ (as a vector space over $\mathbb{Q}$) contains $d \cdot \deg m$ elements, where d is some positive integer. But $\{1, \gamma, \cdots, \gamma^{n-1}\}$ is also a basis of K, so $d \cdot \deg m = n$, and $\deg m$ is a factor of n. □

5.3 The Canonical Basis

Since $P(X)$ is irreducible over $\mathbb{Q}$, there exists n distinct roots $\gamma_1, \cdots, \gamma_n$ of $P(X)$ in $\mathbb{C}$. Each root γ_i induces an embedding σ_i of K into $\mathbb{C}$, by having $\sigma_i(\gamma) = \gamma_i$. It is easy to see that these are all the possible embeddings of K into $\mathbb{C}$, since γ must be mapped to a root of $P(X)$.

Consider the matrix

$$\mathbf{M} = \begin{pmatrix} 1 & \gamma_1 & \cdots & \gamma_1^{n-1} \\ 1 & \gamma_2 & \cdots & \gamma_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \gamma_n & \cdots & \gamma_n^{n-1} \end{pmatrix}.$$

We see that $\mathbf{M}$ is a Vandermonde matrix, and

$$\det \mathbf{M} = \prod_{1 \le i < j \le n} (\gamma_j - \gamma_i) \neq 0.$$

Thus the vectors $\mathbf{v}_i = (\gamma_1^i, \cdots, \gamma_n^i)$ form a basis of $\mathbb{C}^n$. We call $\{\mathbf{v}_i\}$ the canonical basis of K in $\mathbb{C}^n$.

Suppose that we embed K as a vector space into $\mathbb{C}^n$, by mapping γ^i to $\mathbf{v}_i$ in the canonical basis. Then the image of α is $\begin{bmatrix} \sigma_1(\alpha) & \cdots & \sigma_n(\alpha) \end{bmatrix}^\mathsf{T}$. If $F(X)$ is a rational polynomial, then the image of $F(\gamma)$ is $(F(\gamma_1), \cdots, F(\gamma_n))$. From this we see that:

Lemma 53. For any rational polynomial $F(X)$, $F(\gamma) = 0$ in K iff $F(\sigma(\gamma)) = 0$ for every embedding σ of K into $\mathbb{C}$. □

5.4 Norms and Traces

For each $\alpha \in K$, the mapping $x \mapsto \alpha x$ is a linear transformation on the vector space K, and can be represented by a matrix $\mathbf{M}_\alpha \in \mathbb{Q}^{n \times n}$. If $\alpha \neq 0$ then $\mathbf{M}_\alpha$ is invertible, since the mapping $x \mapsto \alpha x$ has an inverse $x \mapsto \alpha^{-1} x$. We denote the determinant of this mapping by $\mathrm{N}(\alpha)$, and its trace by $\mathrm{T}(\alpha)$. It is well-known that characteristic polynomials, determinants, and traces do not depend on the basis used for representation.

Let $m(X)$ be the minimal polynomial of α. We suppose that

$$m(X) = X^r + c_{r-1} X^{r-1} + \cdots + c_0.$$

In Lemma 52 we showed that $\deg m(X)$ is a factor of n. Let $r = \deg m(X)$ and $d = n/r$. Find a basis of K of the form

$$S = \{1, \alpha, \cdots, \alpha^{r-1}, \beta_1, \beta_1\alpha, \cdots, \beta_1\alpha^{r-1}, \cdots, \beta_{d-1}, \beta_{d-1}\alpha, \cdots, \beta_{d-1}\alpha^{r-1}\}.$$

The matrix $\mathbf{M}_\alpha$ under this basis consists of d blocks along the diagonal, each having the form

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & -c_0 \\ 1 & 0 & \cdots & 0 & -c_1 \\ 0 & 1 & \cdots & 0 & -c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -c_{r-1} \end{pmatrix}.$$

Matrices of this form are called *companion matrices*. The characteristic polynomial of each block is $m(X)$. Hence

the characteristic polynomial of $\mathbf{M}_\alpha$ is $m(X)^d$. From this we get:

Lemma 54. If $\alpha \in K$ is an algebraic integer, then both $\mathrm{N}(\alpha)$ and $\mathrm{T}(\alpha)$ are integers.

Proof: $\mathrm{N}(\alpha)$ is $(-1)^n$ times the constant term of the characteristic polynomial $f(\lambda)$ of $\mathbf{M}(\alpha)$, and $\mathrm{T}(\alpha)$ is -1 times the coefficient of the second-highest term of $f(\lambda)$. Since $f(\lambda)$ is a power of a polynomial with only integer coefficients, both values are integers. □

Now consider the two subfields $\mathcal{F}_1 = \mathbb{Q}(\alpha)$ and $\mathcal{F}_2 = \mathbb{Q}(\alpha, \gamma)$. It is obvious that $\mathcal{F}_2 = K$. From our previous discussion, γ has a minimal polynomial $G(X)$ over $\mathcal{F}_1$, and $\deg G(X) = n/r = d$. Suppose that

$$G(X) = X^d + g_{d-1}X^{d-1} + \cdots + g_0$$

where $g_0, \cdots, g_{d-1} \in \mathcal{F}_1$. Each g_i can be written as a polynomial in α. We assume that

$$g_i = t_{i,r-1}\alpha^{r-1} + t_{i,r-2}\alpha^{r-2} + \cdots + t_{i,0}.$$

Let $\alpha_1, \cdots, \alpha_r$ be the r distinct roots of $m(X)$ in $\mathbb{C}$. We make r copies $G(X)$ and denote them by $G_1(X), \cdots, G_r(X)$. In $G_i(X)$, we replace α with α_i. Then we define

$$H(X) = G_1(X) \cdots G_r(X).$$

It is easy to see that $H(X)$ is a monic polynomial of degree $rd = n$. For each $i < n$, the coefficient h_i of X^i is

$$h_i = \sum_{\substack{0 \le k_1, \cdots, k_r \le d \\ k_1 + \cdots + k_r = i}} \prod_{j=1}^{r} g_{k_j}(\alpha_j)$$

where $g_d = 1$. For given indices $s_1, \cdots, s_r$, the coefficient of $\alpha_1^{s_1} \cdots \alpha_r^{s_r}$ in h_i is

$$u(s_1, \cdots, s_r) = \sum_{\substack{0 \le k_1, \cdots, k_r \le d \\ k_1 + \cdots + k_r = i}} \prod_{j=1}^{r} t_{k_j, s_j}.$$

If $s'_1, \cdots, s'_r$ is a permutation of $s_1, \cdots, s_r$, we see that $u(s_1, \cdots, s_r) = u(s'_1, \cdots, s'_r)$. This is because we can apply the same permutation to the indices $k_1, \cdots, k_r$. Thus h_i is a symmetric polynomial in $\alpha_1, \cdots, \alpha_r$. By Vieta's formulas, we see that each h_i is a rational number, and $H(X)$ is a rational polynomial.

For each embedding σ of K into $\mathbb{C}$ we must have $H(\sigma(\gamma)) = 0$. This is because each embedding σ must satisfy $P(\sigma(\gamma)) = 0$ and $f(\sigma(\gamma)) = \sigma(\alpha)$ where $f(\gamma)$ is the polynomial representation of α. As $m(\sigma(\alpha)) = 0$ is a consequence of these two equations, $\sigma(\alpha)$ must be a one of $\alpha_1, \cdots, \alpha_r$, and the corresponding factor $G_i(X)$ becomes 0. Thus by Lemma 53, $H(X)$ is an annihilating polynomial of γ. Since $\deg H(X) = n$, it must be identical to $P(X)$.

What we have shown above is that, the n roots of $P(X)$ in $\mathbb{C}$ can be classified into r groups. Each group contains the d distinct roots of $G_i(X)$. For $j \neq i$, roots of $G_i(X)$ cannot be roots of $G_j(X)$, since the n roots of $P(X) = H(X)$ are distinct. Now if an embedding σ of K into $\mathbb{C}$ satisfies $G_i(\sigma(\gamma)) = 0$, we must also have $\sigma(\alpha) = \alpha_i$. This is because $\sigma(G(\gamma)) = 0$ must be true, but $\sigma(G(X))$ must be one of $G_1(X), \cdots, G_r(X)$ depending on $\sigma(\alpha)$. Thus each root of $m(X) = 0$ occurs with multiplicity d within $\sigma_1(\alpha), \cdots, \sigma_n(\alpha)$.

Lemma 55. Let $\sigma_1, \cdots, \sigma_n$ be the embeddings of K into $\mathbb{C}$. Then for each $\alpha \in K$ we have

$$\mathrm{N}(\alpha) = \prod_{k=1}^{n} \sigma_k(\alpha), \quad \mathrm{T}(\alpha) = \sum_{k=1}^{n} \sigma_k(\alpha).$$

Proof: Jarvis (2014, Proposition 3.16, p. 47). Simply check that $\sigma_1(\alpha), \cdots, \sigma_n(\alpha)$ are exactly the roots of the characteristic polynomial of $\mathbf{M}_\alpha$. □

5.5 Algebraic Integers as a Lattice

Via the canonical basis we can embed K as a vector space into $\mathbb{C}^n$. However, such a representation is not yet compatible with our notion of lattice, because we need a representation in $\mathbb{R}^n$. This issue can be resolved as follows. Let $\gamma_1, \cdots, \gamma_n$ be the n roots of $P(X)$. Suppose that there are s real roots and t complex roots. Since the coefficients of $P(X)$ are real, the complex roots always appear in conjugate pairs. Therefore we may assume $\gamma_1, \cdots, \gamma_s \in \mathbb{R}$, and $\gamma_{s+k} = \overline{\gamma_{n+1-k}}$. Then for every $\alpha \in K$ we have $\sigma_{s+k}(\alpha) = \overline{\sigma_{n+1-k}(\alpha)}$. We can thus represent α by the vector

$$\boldsymbol{\alpha} = \begin{pmatrix} \sigma_1(\alpha) \\ \vdots \\ \sigma_s(\alpha) \\ \mathrm{Re}(\sigma_{s+1}(\alpha)) + \mathrm{Im}(\sigma_{s+1}(\alpha)) \\ \mathrm{Re}(\sigma_{s+1}(\alpha)) - \mathrm{Im}(\sigma_{s+1}(\alpha)) \\ \vdots \\ \mathrm{Re}(\sigma_{s+t/2}(\alpha)) + \mathrm{Im}(\sigma_{s+t/2}(\alpha)) \\ \mathrm{Re}(\sigma_{s+t/2}(\alpha)) - \mathrm{Im}(\sigma_{s+t/2}(\alpha)) \end{pmatrix}.$$

Notice that we are only taking the first $t/2$ complex embeddings. However, each complex embedding is split into two components, so in total we still have n components. This embedding of K into $\mathbb{R}^n$ is equivalent as an inner product space to the embedding into $\mathbb{C}^n$. Recall that the standard inner product on $\mathbb{C}^n$ is

$$\langle \mathbf{z}, \mathbf{z}' \rangle = \sum_{k=1}^{n} z_i \overline{z_i'}.$$

Since the embeddings are in conjugate pairs, we have

$$\begin{aligned} \sigma_{s+k}(\alpha)\overline{\sigma_{s+k}(\beta)} + \sigma_{n+1-k}(\alpha)\overline{\sigma_{n+1-k}(\beta)} &= \sigma_{s+k}(\alpha)\overline{\sigma_{s+k}(\beta)} + \overline{\sigma_{s+k}(\alpha)}\sigma_{s+k}(\beta) \\ &= 2[\mathrm{Re}(\sigma_{s+k}(\alpha))\mathrm{Re}(\sigma_{s+k}(\beta)) + \mathrm{Im}(\sigma_{s+k}(\alpha))\mathrm{Im}(\sigma_{s+k}(\beta))]. \end{aligned}$$

So in $\mathbb{C}^n$ the inner product is

$$\langle \alpha, \beta \rangle = \sum_{k=1}^{n} \sigma_k(\alpha)\overline{\sigma_k(\beta)} = \sum_{k=1}^{s} \sigma_k(\alpha)\sigma_k(\beta) + 2\sum_{k=1}^{t/2} \mathrm{Re}(\sigma_{s+k}(\alpha))\mathrm{Re}(\sigma_{s+k}(\beta)) + \mathrm{Im}(\sigma_{s+k}(\alpha))\mathrm{Im}(\sigma_{s+k}(\beta)).$$

The same is true for the $\mathbb{R}^n$ embedding. Furthermore, notice that

$$\begin{pmatrix} \frac{1-i}{2} & \frac{1+i}{2} \\ \frac{1+i}{2} & \frac{1-i}{2} \end{pmatrix} \begin{pmatrix} \sigma_{s+k}(\alpha) \\ \sigma_{n+1-k}(\alpha) \end{pmatrix} = \begin{pmatrix} \mathrm{Re}(\sigma_{s+k}(\alpha)) + \mathrm{Im}(\sigma_{s+k}(\alpha)) \\ \mathrm{Re}(\sigma_{s+k}(\alpha)) - \mathrm{Im}(\sigma_{s+k}(\alpha)) \end{pmatrix}, \quad \det \begin{pmatrix} \frac{1-i}{2} & \frac{1+i}{2} \\ \frac{1+i}{2} & \frac{1-i}{2} \end{pmatrix} = -i.$$

So for any $\alpha_1, \cdots, \alpha_n$, if $\boldsymbol{\alpha}_i$ is the embedding of α in $\mathbb{C}^n$ and $\boldsymbol{\alpha}'_i$ is its embedding in $\mathbb{R}^n$ then

$$\left|\det\begin{bmatrix}\boldsymbol{\alpha}_1 & \cdots & \boldsymbol{\alpha}_n\end{bmatrix}\right| = \left|\det\begin{bmatrix}\boldsymbol{\alpha}'_1 & \cdots & \boldsymbol{\alpha}'_n\end{bmatrix}\right|.$$

Now let $\omega_1, \cdots, \omega_n$ be n elements in K. We define

$$\mathbf{M} = \begin{pmatrix}\sigma_1(\omega_1) & \cdots & \sigma_1(\omega_n)\\ \vdots & \ddots & \vdots\\ \sigma_n(\omega_1) & \cdots & \sigma_n(\omega_n)\end{pmatrix}, \quad \Delta(\omega_1, \cdots, \omega_n) = (\det \mathbf{M})^2.$$

Notice that

$$(\det \mathbf{M})^2 = \det \mathbf{M}\mathbf{M}^{\mathsf{T}} = \det \mathbf{T}$$

where

$$t_{ij} = \sum_{k=1}^{n} \sigma_k(\omega_i)\sigma_k(\omega_j) = \sum_{k=1}^{n} \sigma_k(\omega_i\omega_j) = \mathrm{T}(\omega_i\omega_j) \in \mathbb{Q}.$$

Thus $\Delta(\omega_1, \cdots, \omega_n)$ is a real number, and $|\Delta(\omega_1, \cdots, \omega_n)|$ is the square of the volume of the fundamental parallelepiped of the lattice with basis $\omega_1, \cdots \omega_n$, when they are embedded into $\mathbb{R}^n$ as explained above.

Lemma 56. If $\omega_1, \cdots, \omega_n$ are algebraic integers, then $\Delta(\omega_1, \cdots, \omega_n)$ is an integer.

Proof: Jarvis (2014, Corollary 3.20, p. 48). Notice that each t_{ij} is an integer, by Lemma 54. □

Let $\alpha_1, \cdots, \alpha_n$ be a basis of the vector space K. By Lemma 48, we may scale each α_i by an integer k_i, so that each $\beta_i = k_i\alpha_i$ is an algebraic integer. Since sums of algebraic integers are still algebraic integers, each expression of the form

$$q_1\beta_1 + \cdots + q_n\beta_n$$

where $q_1, \cdots, q_n \in \mathbb{Z}$ is an algebraic integer. However, it is not guaranteed that all algebraic integers can be expressed this way. Notice that if α is an algebraic integer, then $\alpha\beta_i$ is also an algebraic integer, and so $\mathrm{T}(\alpha\beta_i) \in \mathbb{Z}$. Each $\mathrm{T}(\alpha\beta_i)$ is a linear function of α. Therefore, we can find $\delta_1, \cdots, \delta_n \in K$ such that

$$\mathrm{T}(\delta_i\beta_j) = \begin{cases}1 & (i = j)\\ 0 & (i \neq j)\end{cases}.$$

Then it is evident that every algebraic integer can be expressed as $q_1\delta_1 + \cdots + q_n\delta_n$, but it is not guaranteed that all such expressions are algebraic integers. Thus $\mathbb{Z}_K$ is "sandwiched" between two lattices L, L', where L is generated by $\beta_1, \cdots, \beta_n$ and L' is generated by $\delta_1, \cdots, \delta_n$.

We can check whether L contains all algebraic integers by checking each equivalence class of L'/L. Suppose that we have found some $\alpha \in \mathbb{Z}_K$ such that $\alpha \notin L$. We can still express α as a linear combination

$$\alpha = r_1\beta_1 + \cdots + r_n\beta_n$$

but some r_i will be in $\mathbb{Q} \setminus \mathbb{Z}$. Without loss of generality, suppose that $r_1 \in \mathbb{Q} \setminus \mathbb{Z}$. Let $k = \lfloor r_1 \rfloor$. Replace β_1 with $\alpha - k\beta_1$. Then

$$|\Delta(\beta'_1, \beta_2, \cdots, \beta_n)| = (r_1 - k)^2|\Delta(\beta_1, \cdots, \beta_n)| < |\Delta(\beta_1, \cdots, \beta_n)|.$$

But $|\Delta(\beta_1, \cdots, \beta_n)|$ is always a positive integer and cannot decrease infinitely. Thus after a finite number of steps we obtain $\beta_1, \cdots, \beta_n \in \mathbb{Z}_K$ such that all algebraic integers can be expressed as integral combinations of these elements. We say $\beta_1, \cdots, \beta_n$ is an integral basis of $\mathbb{Z}_K$.

If $\beta_1, \cdots, \beta_n$ and $\beta'_1, \cdots, \beta'_n$ are two integral bases of $\mathbb{Z}_K$, then we must have $|\Delta(\beta_1, \cdots, \beta_n)| = |\Delta(\beta'_1, \cdots, \beta'_n)|$. The determinant is called the *discriminant* of $\mathbb{Z}_K$.

5.6 Ideals

Definition 57. An *ideal* $\mathcal{I}$ of $\mathbb{Z}_K$ is a subset of $\mathbb{Z}_K$ such that:

1. $0 \in \mathcal{I}$;
2. If $\alpha, \beta \in \mathcal{I}$ then $\alpha + \beta \in \mathcal{I}$;
3. If $\alpha \in \mathcal{I}, \beta \in \mathbb{Z}_K$ then $\alpha\beta \in \mathcal{I}$.

Remark 58. The singleton set $\mathcal{I} = \{0\}$ is an ideal of $\mathbb{Z}_K$. Also, $\mathbb{Z}_K$ is an ideal of itself. A *non-zero* ideal is an ideal that contains at least one non-zero element. A *proper* ideal is an ideal that is a proper subset of $\mathbb{Z}_K$.

Definition 59. For a given $\alpha \in \mathbb{Z}_K$, the set $\mathcal{I} = \{\alpha\beta \mid \beta \in \mathbb{Z}_K\}$ is an ideal. We call it the *principal ideal* generated by α and denote it by $\langle\alpha\rangle$.

Remark 60. For a given non-zero ideal $\mathcal{I}$, let α be any non-zero element of $\mathcal{I}$. Then $\alpha, \alpha\gamma, \cdots, \alpha\gamma^{n-1}$ are linearly independent elements of $\mathcal{I}$. Therefore, $\mathcal{I}$ is a sublattice of $\mathbb{Z}_K$. It is called an *ideal lattice*. This does not mean every element in $\mathcal{I}$ can be written as $\alpha \cdot \beta$ for some $\beta \in \mathbb{Z}_K$. We still need to follow the procedure described earlier to find an integral basis for $\mathcal{I}$.

Definition 61. Let $\mathcal{I}, \mathcal{J}$ be two ideals of $\mathbb{Z}_K$.

1. $\mathcal{I} + \mathcal{J} = \{\alpha + \beta \mid \alpha \in \mathcal{I}, \beta \in \mathcal{J}\}$;
2. $\mathcal{I}\mathcal{J} = \{\alpha_1\beta_1 + \alpha_2\beta_2 + \cdots + \alpha_k\beta_k \mid k \in \mathbb{N}, \alpha_i \in \mathcal{I}, \beta_i \in \mathcal{J}\}$.

Remark 62. If $\mathcal{I}, \mathcal{J}, \mathcal{K}$ are three ideals such that $\mathcal{I} \subseteq \mathcal{J}$, then $\mathcal{I}\mathcal{K} \subseteq \mathcal{J}\mathcal{K}$.

Remark 63. It is easy to see that $\mathcal{I}\mathcal{J} \subseteq \mathcal{I} \cap \mathcal{J}$. In general it is not true that $\mathcal{I}\mathcal{J} = \mathcal{I} \cap \mathcal{J}$. However, if $\mathcal{I}, \mathcal{J}$ are *coprime*, i.e. $\mathcal{I} + \mathcal{J} = \mathbb{Z}_K$, then $\mathcal{I}\mathcal{J} = \mathcal{I} \cap \mathcal{J}$.

If $\mathcal{I} + \mathcal{J} = \mathbb{Z}_K$, then there exists $x \in \mathcal{I}, y \in \mathcal{J}$ such that $x + y = 1$. Then for every $c \in \mathcal{I} \cap \mathcal{J}$ we have $c = c(x + y) = cx + cy \in \mathcal{I}\mathcal{J}$.

Definition 64. An proper ideal $\mathcal{I}$ is *prime* if for every $\alpha, \beta \in \mathbb{Z}_K$, if $\alpha\beta \in \mathcal{I}$ then $\alpha \in \mathcal{I} \vee \beta \in \mathcal{I}$.

Remark 65. An alternative characterization of prime ideals (Jarvis, 2014, Lemma 5.13, p. 96) is as follows. An ideal $\mathcal{K}$ is a prime ideal iff whenever $\mathcal{I}\mathcal{J} \subseteq \mathcal{K}$, then $\mathcal{I} \subseteq \mathcal{K} \vee \mathcal{J} \subseteq \mathcal{K}$.

Suppose that $\mathcal{K}$ is prime and $\mathcal{I}\mathcal{J} \subseteq \mathcal{K}$. If neither $\mathcal{I} \subseteq \mathcal{K}$ nor $\mathcal{J} \subseteq \mathcal{K}$ holds, then there exists $\alpha \in \mathcal{I}, \beta \in \mathcal{J}$ such that $\alpha, \beta \notin \mathcal{K}$. However $\alpha\beta \in \mathcal{I}\mathcal{J}$ and so $\alpha\beta \in \mathcal{K}$. Because $\mathcal{K}$ is prime, either $\alpha \in \mathcal{K}$ or $\beta \in \mathcal{K}$, which is a contradiction.

Suppose that $\mathcal{K}$ is not prime. Then there exists $\alpha, \beta \notin \mathcal{K}$ such that $\alpha\beta \in \mathcal{K}$. Then we have neither $\langle\alpha\rangle \subseteq \mathcal{K}$ nor $\langle\beta\rangle \subseteq \mathcal{K}$. However we have $\langle\alpha\rangle\langle\beta\rangle = \langle\alpha\beta\rangle \subseteq \mathcal{K}$.

Definition 66. A proper ideal $\mathcal{I}$ is *maximal* if $\mathcal{I}$ is not a subset of any other proper ideal.

Remark 67. If $\mathcal{I}$ is an ideal and $\beta \in \mathbb{Z}_K$ then $\mathcal{I}' = \mathcal{I} + \langle\beta\rangle$ is an ideal. If $\beta \notin \mathcal{I}$ then $\mathcal{I}'$ is strictly larger than $\mathcal{I}$. If $\mathcal{I}$ is maximal, then $\mathcal{I}'$ cannot be a proper ideal, so $\mathcal{I}' = \mathbb{Z}_K$ and $1 \in \mathcal{I}'$.

As such, if $\mathcal{I}$ is maximal, then for each $\beta \notin \mathcal{I}$ we can find $\delta \in \mathbb{Z}_K$ such that $\beta\delta + \alpha = 1$ for some $\alpha \in \mathcal{I}$. Note that this implies $\delta \notin \mathcal{I}$. Otherwise, $1 \in \mathcal{I}$ and $\mathcal{I} = \mathbb{Z}_K$. This is usually formulated as: if $\mathcal{I}$ is maximal then $\mathbb{Z}_K/\mathcal{I}$ is a field.

Conversely, if $\mathbb{Z}_K/\mathcal{I}$ is a field, i.e. for each $\beta \notin \mathcal{I}$ there exists $\delta \in \mathbb{Z}_K$ such that $\beta\delta + \alpha = 1$ for some $\alpha \in \mathcal{I}$, then any ideal $\mathcal{I}'$ that is strictly larger than $\mathcal{I}$ must contain 1 and so $\mathcal{I}' = \mathbb{Z}_K$. As such $\mathcal{I}$ is maximal.

Remark 68. A maximal ideal is always prime. Suppose that $\mathcal{I}$ is maximal but not prime. Then there exists $\beta, \delta \in \mathbb{Z}_K$ such that $\beta, \delta \notin \mathcal{I}$ but $\beta\delta \in \mathcal{I}$. Find β' such that $\beta\beta' + \alpha = 1$ for some $\alpha \in \mathcal{I}$. Then $\beta\beta'\delta = \delta - \alpha\delta \in \mathcal{I}$, but then $\delta \in \mathcal{I}$ which is a contradiction.

The converse is not true in general. However, in algebraic number rings it is true.

Lemma 69. Every non-zero prime ideal of $\mathbb{Z}_K$ is maximal.

Proof: Jarvis (2014, Proposition 5.21, p. 98). Notice that every ideal $\mathcal{I}$ of $\mathbb{Z}_K$ is a sublattice of $\mathbb{Z}_K$. By Lemma 10, $\mathbb{Z}_K/\mathcal{I}$ has only a finite number of equivalence classes. For a given prime ideal $\mathcal{I}$ and $\alpha \notin \mathcal{I}$, consider the sequence $\alpha, \alpha^2, \cdots$. By induction we see that each α^k is not in $\mathcal{I}$. Eventually we can find $j < k$ such that α^j and α^k that belong to the same equivalence class in $\mathbb{Z}_K/\mathcal{I}$. Then we have $\alpha^j(1 - \alpha^{k-j}) \in \mathcal{I}$. Since $\alpha^j \notin \mathcal{I}$, we must have $1 - \alpha^{k-j} \in \mathcal{I}$. Thus $\alpha \cdot \alpha^{k-j-1} = 1$ in $\mathbb{Z}_K/\mathcal{I}$. Hence $\mathbb{Z}_K/\mathcal{I}$ is a field, and $\mathcal{I}$ is maximal. □

Definition 70. A *fractional ideal* $\mathcal{I}$ of $\mathbb{Z}_K$ is a subset of K such that $\mathcal{J} = \alpha\mathcal{I}$ is an ideal of $\mathbb{Z}_K$ for some non-zero $\alpha \in \mathbb{Z}_K$. We may write $\mathcal{I} = \mathcal{J}/\alpha$.

Remark 71. Fractional ideals may contain elements of K that are not in $\mathbb{Z}_K$. Therefore, in general they are not ideals of $\mathbb{Z}_K$. Fractional ideals can be seen as lattices in the $\mathbb{R}^n$ space, but they are not sublattices of $\mathbb{Z}_K$.

Remark 72. Addition and multipliation of ideals can be extended to fractional ideals in the following natural way:

$$\mathcal{I}/\alpha + \mathcal{J}/\beta = (\beta\mathcal{I} + \alpha\mathcal{J})/(\alpha\beta), \quad \mathcal{I}/\alpha \cdot \mathcal{J}/\beta = \mathcal{I}\mathcal{J}/(\alpha\beta).$$

Lemma 73. If $\mathcal{I}, \mathcal{J}$ are two ideals with $\mathcal{I} \subseteq \mathcal{J}$, and $\mathcal{K}$ is a fractional ideal, then $\mathcal{I}\mathcal{K} \subseteq \mathcal{J}\mathcal{K}$.

Proof: Suppose that $\mathcal{K} = \mathcal{K}'/\alpha$, then $\mathcal{I}\mathcal{K} = \mathcal{I}\mathcal{K}'/\alpha$ and $\mathcal{J}\mathcal{K} = \mathcal{J}\mathcal{K}'/\alpha$. Then it is sufficient to notice that $\mathcal{I}\mathcal{K}' \subseteq \mathcal{J}\mathcal{K}'$. □

Lemma 74. If a fractional ideal $\mathcal{I}$ is a subset of an ideal $\mathcal{J}$, then $\mathcal{I}$ is an ideal.

Proof: We have $\mathcal{I} \subseteq \mathcal{J} \subseteq \mathbb{Z}_K$. It is straightforward to check that $\mathcal{I}$ satisfies other requirements of an ideal. □

Lemma 75. If $\mathcal{I}$ is a non-zero ideal of $\mathbb{Z}_K$, then

$$\mathcal{I}^{-1} = \{\alpha \in K \mid \alpha\mathcal{I} \subseteq \mathbb{Z}_K\}$$

is a fractional ideal, and $\mathcal{I}\mathcal{I}^{-1} = \mathbb{Z}_K$.

Proof: Jarvis (2014, Lemma 5.25, Lemma 5.28, Lemma 5.29, pp. 100–101). □

Theorem 76 (Unique Factorization of Ideals)**.** Every non-zero proper ideal of $\mathbb{Z}_K$ can uniquely written as a product of prime ideals of $\mathbb{Z}_K$.

Proof: Jarvis (2014, Lemma 5.31, Theorem 5.32, p. 102). □

Lemma 77. Let $\mathcal{I}, \mathcal{J}$ be two non-zero ideals of $\mathbb{Z}_K$. Then $\mathcal{I} \subseteq \mathcal{J}$ iff there exists an ideal $\mathcal{K}$ such that $\mathcal{I} = \mathcal{J}\mathcal{K}$.

Proof: If $\mathcal{J} = \mathbb{Z}_K$ then the proposition is trivial, because every $\mathcal{I}$ satisfies $\mathcal{I} \subseteq \mathbb{Z}_K$, and also satisfies $\mathcal{I} = \mathbb{Z}_K\mathcal{I}$. Also, if $\mathcal{I} = \mathcal{J}\mathcal{K}$ then it is easy to see that $\mathcal{I} \subseteq \mathcal{J}$.

Hence assume $\mathcal{J}$ is a proper ideal and $\mathcal{I} \subseteq \mathcal{J}$. By Theorem 76, $\mathcal{J}$ can be written as

$$\mathcal{J} = \mathfrak{p}_1^{a_1} \cdots \mathfrak{p}_k^{a_k}$$

where $\mathfrak{p}_i$ are prime ideals and a_i are positive integers. We shall do induction on $a_1 + \cdots + a_k$.

The base case is that $\mathcal{J}$ is a prime ideal. Consider the factorization of $\mathcal{I}$:

$$\mathcal{I} = \mathfrak{q}_1^{b_1} \cdots \mathfrak{q}_k^{b_l}$$

By Remark 65, at least one factor $\mathfrak{q}_i$ is a subset of $\mathcal{J}$. However, by Lemma 69, $\mathfrak{q}_i$ is maximal, so $\mathfrak{q}_i = \mathcal{J}$ and we take $\mathcal{K}$ to be the other factors of $\mathcal{I}$.

In the inductive case, we have $\mathcal{J} = \mathcal{J}'\mathfrak{p}$ such that p is a prime ideal. The reasoning is similar to the base case. The ideal $\mathcal{I}$ must have a prime factor equal to $\mathfrak{p}$. By Lemma 73, $\mathcal{I} \subseteq \mathcal{J}$ implies $\mathcal{I}\mathfrak{p}^{-1} \subseteq \mathcal{J}\mathfrak{p}^{-1} = \mathcal{J}'$. By Lemma 74, $\mathcal{I}\mathfrak{p}^{-1}$ is an ideal. Now $\mathcal{J}'$ has one less prime factor than $\mathcal{J}$, and the induction hypothesis applies. □

By Lemma 77, $\mathcal{I}$ is a multiple of $\mathcal{J}$, or $\mathcal{J}$ divides $\mathcal{I}$, iff $\mathcal{I} \subseteq \mathcal{J}$. Thus ideal divisibility is equivalent to reverse containment.

Remark 78. If $\mathfrak{p}$ is a prime ideal, then by Lemma 69 we see that $\mathbb{Z}_K/\mathfrak{p}$ is a finite field. A finite field always has a positive characteristic p, and we must have $p \in \mathfrak{p}$. Then we have $\langle p \rangle \subseteq \mathfrak{p}$, so $\mathfrak{p}$ divides $\langle p \rangle$. Conversely, if $\mathfrak{p}$ is a factor of $\langle p \rangle$, then $\langle p \rangle \subseteq \mathfrak{p}$ and $p \in \mathfrak{p}$. A proper ideal cannot contain two prime numbers, otherwise $1 \in \mathfrak{p}$ and so $\mathfrak{p} = \mathbb{Z}_K$. Hence, all factors of $\langle p \rangle$ have characteristic p. This shows that, to understand the prime ideals in $\mathbb{Z}_K$, it is sufficient to understand how each principal ideal $\langle p \rangle$ factors in $\mathbb{Z}_K$.

Lemma 79. Let $\mathcal{I}, \mathcal{J}$ be two non-zero ideals. Suppose that

$$\mathcal{I} = \mathfrak{p}_1^{a_1} \cdots \mathfrak{p}_k^{a_k}, \quad \mathcal{J} = \mathfrak{p}_1^{b_1} \cdots \mathfrak{p}_k^{b_k}$$

where $\mathfrak{p}_i$ are prime ideals and $a_i, b_i \in \mathbb{N}$. Note that we allow a_i, b_i to be 0, so $\mathcal{I}, \mathcal{J}$ can still have distinct prime ideal factors. The case where all exponents are 0 corresponds to $\mathbb{Z}_K$. Then we have

$$\mathcal{I} + \mathcal{J} = \mathfrak{p}_1^{\min(a_1,b_1)} \cdots \mathfrak{p}_k^{\min(a_k,b_k)}.$$

In other words, $\mathcal{I} + \mathcal{J}$ is the "greatest common denominator" of $\mathcal{I}$ and $\mathcal{J}$.

Proof: Let $\mathcal{K} = \mathfrak{p}_1^{\min(a_1,b_1)} \cdots \mathfrak{p}_k^{\min(a_k,b_k)}$ and write $\mathcal{I} = \mathcal{I}'\mathcal{K}, \mathcal{J} = \mathcal{J}'\mathcal{K}$, so that $\mathcal{I}', \mathcal{J}'$ have no common prime ideal factors. It is sufficient to prove that $\mathcal{I}' + \mathcal{J}' = \mathbb{Z}_K$. If not, then $\mathcal{I}' + \mathcal{J}'$ has a prime ideal factor $\mathfrak{q}$. But then $\mathcal{I}' \subseteq \mathcal{I}' + \mathcal{J}' \subseteq \mathfrak{q}$. Similarly, $\mathcal{J}' \subseteq \mathfrak{q}$. This contradicts that $\mathcal{I}', \mathcal{J}'$ have no common prime ideal factors. □

Lemma 80. Similar to Lemma 79, we also have the "least common multiple"

$$\mathcal{I} \cap \mathcal{J} = \mathfrak{p}_1^{\max(a_1,b_1)} \cdots \mathfrak{p}_k^{\max(a_k,b_k)}.$$

Proof: Analogous to Lemma 79. See Remark 63. □

Definition 81. Let L be the lattice of algebraic integers. Let $\mathcal{I}$ be a non-zero ideal or fractional ideal of $\mathbb{Z}_K$. Let L' be the lattice associated to $\mathcal{I}$. The *norm* of $\mathcal{I}$, denoted by $\mathrm{N}(\mathcal{I})$, is defined to be $|L/L'| = |d(L')/d(L)|$.

Lemma 82. For any $\alpha \in \mathbb{Z}_K$ with $\alpha \neq 0$ we have $\mathrm{N}(\langle \alpha \rangle) = |\mathrm{N}(\alpha)|$.

Proof: Jarvis (2014, Lemma 5.35, p. 104). □

Lemma 83. For any non-zero ideal or fractional ideal $\mathcal{I}$ of $\mathbb{Z}_K$, let L be the lattice associated to $\mathcal{I}$, then $|d(L)| = \mathrm{N}(\mathcal{I}) \cdot \sqrt{|\Delta|}$ where Δ is the discriminant of $\mathbb{Z}_K$.

Proof: Immediate from the definition of $\mathrm{N}(\mathcal{I})$ and Δ. □

Lemma 84. For any two non-zero ideals or fractional ideals $\mathcal{I}, \mathcal{J}$ of $\mathbb{Z}_K$ we have $\mathrm{N}(\mathcal{I}\mathcal{J}) = \mathrm{N}(\mathcal{I})\mathrm{N}(\mathcal{J})$.

Proof: Jarvis (2014, Lemma 5.36, Lemma 5.37, pp. 104–105). The extension to fractional ideals is straightforward. □

Remark 85. If p is a prime number and $\mathfrak{p}$ is a prime ideal with $p \in \mathfrak{p}$, then $\mathbb{Z}_K/\mathfrak{p}$ is a finite field with characteristic p. In this case we have $\mathrm{N}(\mathfrak{p}) = p^f$ for some positive integer f, and f is called the *inertial degree* of $\mathfrak{p}$.

We are now ready to state how each principal ideal $\langle p \rangle$ factors in $\mathbb{Z}_K$. The theorem below applies only when $\mathbb{Z}_K = \mathbb{Z}[\gamma]$. This special case is sufficient for understanding Lyubashevsky et al. (2013). The more general case requires the theory of the "conductor" ideal and is detailed in Conrad ([n. d.]).

Theorem 86. Recall that $P(X)$ is the minimal polynomial of γ. Since we assume $\mathbb{Z}_K = \mathbb{Z}[\gamma]$, γ is an algebraic integer so $P(X)$ contains only integer coefficients. Let p be a prime, and let $\overline{P}(X)$ be the factorization of $P(X)$ in $\mathbb{F}_p$:

$$\overline{P}(X) = \overline{P}_1(X)^{e_1} \cdots \overline{P}_k(X)^{e_k}.$$

Then there exists distinct prime ideals $\mathfrak{p}_1, \cdots, \mathfrak{p}_k$ of $\mathbb{Z}_K$ such that

$$\langle p \rangle = \mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_k^{e_k},$$

and the inertial degree of $\mathfrak{p}_i$ is $\deg \overline{P}_i(X)$.

Proof: Jarvis (2014, Proposition 5.42, p. 109). □

Remark 87. From the factorization we see that

$$n = \deg P(X) = \deg \overline{P}(X) = \sum_{i=1}^{k} e_i \cdot \deg \overline{P}_i(X).$$

This is true in general. See Jarvis (2014, Theorem 5.41, p. 108).

5.7 The Cyclotomic Fields

Definition 88. Let m be a positive integer. An m-th *root of unity* is a solution to the equation $\zeta^m = 1$ in $\mathbb{C}$. In fact, $\zeta = e^{2\pi i k/m}$ for $k = 0, 1, 2, \ldots$

Definition 89. An m-th root of unity is *primitive* if $\zeta^k \neq 1$ for any $1 \leq k < m$. This corresponds to $\zeta = e^{2\pi i k/m}$ for k coprime to m.

Definition 90. The m-th cyclotomic polynomial $\Phi_m(X)$ is the irreducible polynomial whose roots are the m-th primitive roots of unity:

$$\Phi_m(X) = \prod_{k \text{ coprime to } m} (X - e^{2\pi i k/m}).$$

The degree of $\Phi_m(X)$ is $n = \varphi(m)$.

Let $K_m = \mathbb{Q}[X]/\Phi_m(X)$. The results we need about K_m and $\mathbb{Z}_{K_m}$ are:

Lemma 91. $\mathbb{Z}_{K_m} = \mathbb{Z}[\zeta]$ where ζ is any m-th primitive root of unity.

Proof: Neukirch (1999, Proposition 10.2, p. 60). □

Lemma 92. The discriminant of $\mathbb{Z}_{K_m}$ is

$$\Delta_{K_m} = \left(\frac{m}{\prod_{\text{prime } p|m} p^{1/(p-1)}} \right)^n \leq n^n.$$

Proof: Shurman ([n. d.]). □

Lemma 93. Let p be a prime number. The factoring of principal ideal $\langle p \rangle$ in $\mathbb{Z}_{K_m}$ is as follows. Let r_p be the largest integer such that m is a multiple of p^{r_p}. Let f_p be the smallest positive integer such that $p^{f_p} \equiv 1 \bmod m/p^{r_p}$. Let $d = n/[f_p \cdot \varphi(p^{r_p})]$. Then we have

$$\langle p \rangle = (\mathfrak{p}_1 \cdots \mathfrak{p}_d)^{\varphi(p^{r_p})}$$

for some distinct prime ideals $\mathfrak{p}_1, \cdots, \mathfrak{p}_d$, each having norm p^{f_p}.

Proof: Neukirch (1999, Proposition 10.3, p. 61). □

6 The Regularity Lemma

We now have all necessary background information to state and prove the regularity lemma. Let m be an integer with $m \geq 3$ and $n = \varphi(m)$. Let K be the m-th cyclotomic field, and let $R = \mathbb{Z}[\zeta]$ be its ring of algebraic integers. Let q be an integer. Let R_q be the subset of R with the coefficient of each ζ^i within $[0, q-1]$. If a, b are positive integers, let $R^{[a]}$ be the set of vectors of length a with entries in R, and $R^{[a]\times[b]}$ be the set of matrices of size $a \times b$ with entries in R. Similarly, $R_q^{[a]}$ is the set of vectors of length a with entries in R_q, and $R_q^{[a]\times[b]}$ is the set of matrices of size $a \times b$ with entries in R_q. Let $\mathcal{D}_s$ be a discrete Gaussian distribution on R with density

$$\mathcal{D}_s(x) = \frac{\exp(-\pi ||x||^2/s^2)}{\sum_{x \in R} \exp(-\pi ||x||^2/s^2)}$$

where $||x||$ means the length of x under the canonical embedding into $\mathbb{R}^n$.

Let k, l be two positive integers with $k \leq l < 2^n$. The regularity lemma states: let $\overline{\mathbf{A}}$ is a matrix uniformly random in $R_q^{[k]\times[l-k]}$, let $\mathbf{A} = [\mathbf{I}_k | \overline{\mathbf{A}}]$, let $\mathbf{x}$ be a vector in $R_q^{[l]}$ with each component sampled from R according to $\mathcal{D}_s(x)$ with $s > 2n \cdot q^{k/l+2/(nl)}$ and reduced by qR, then $\mathbf{Ax} \bmod qR$ is almost uniformly random over $R_q^{[k]}$.

Proof of regularity lemma: For any matrix $\overline{\mathbf{A}} \in R_q^{[k]\times[l-k]}$, let $\mathbf{A} = [\mathbf{I}_k | \overline{\mathbf{A}}]$. Define the set

$$\Lambda^\perp(\mathbf{A}) = \{\mathbf{z} \in R^{[l]} \mid \mathbf{Az} = \mathbf{0} \bmod qR\}.$$

For any $\mathbf{z} \in R^{[l]}$ we have $\mathbf{A}(q\mathbf{z}) = q\mathbf{Az} = \mathbf{0} \bmod qR$. Therefore $\Lambda^\perp(\mathbf{A})$ is a lattice of rank nl. The idea of the proof is to estimate the smoothing factor of $\Lambda^\perp(\mathbf{A})$. If $s > \eta_\varepsilon(\Lambda^\perp(\mathbf{A}))$, then as the components of $\mathbf{z}$ are sampled according to $\mathcal{D}_s$, the probability of $\mathbf{z}$ falling into each equivalence class of $R^{[l]}/\Lambda^\perp(\mathbf{A})$ should be roughly uniform. Also, since $\mathbf{A}$ has full rank, the number of equivalence classes of $\mathbf{z}$ satisfying $\mathbf{Az} = \mathbf{t} \bmod qR$ for each $\mathbf{t} \in R_q^{[k]}$ should be equal. Together, we get an almost uniform distribution of $\mathbf{Az} \bmod qR$.

Suppose that $\mathbf{z} \in \Lambda^\perp(\mathbf{A})$. Let $\mathbf{z}^\top = [\mathbf{z}_1^\top \mid \mathbf{z}_2^\top]$ where $\mathbf{z}_1 \in R^{[k]}$ and $\mathbf{z}_2 \in R^{[l-k]}$. Then $\mathbf{Az} = \mathbf{z}_1 + \overline{\mathbf{A}}\mathbf{z}_2 \in qR^{[k]}$. Suppose that $\mathbf{Az} = q\mathbf{z}_3$. Then $\mathbf{z}_1 = q\mathbf{z}_3 - \overline{\mathbf{A}}\mathbf{z}_2$. As such, we have

$$\Lambda^\perp(\mathbf{A}) = \left\{ \begin{bmatrix} q\mathbf{z}_3 - \overline{\mathbf{A}}\mathbf{z}_2 \\ \mathbf{z}_2 \end{bmatrix} \;\middle|\; \mathbf{z}_2 \in R^{[l-k]}, \mathbf{z}_3 \in R^{[k]} \right\}.$$

Suppose that $\mathbf{y} \in \Lambda^\perp(\mathbf{A})^*$. Let $\mathbf{y}^\top = [\mathbf{y}_1^\top \mid \mathbf{y}_2^\top]$ where $\mathbf{y}_1 \in R^{[k]}, \mathbf{y}_2 \in R^{[l-k]}$. By definition, for each $\mathbf{z} \in \Lambda^\top(\mathbf{A})$ we have

$$\langle \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{y}_1, q\mathbf{z}_3 - \overline{\mathbf{A}}\mathbf{z}_2 \rangle + \langle \mathbf{y}_2, \mathbf{z}_2 \rangle = q \cdot \langle \mathbf{y}_1, \mathbf{z}_3 \rangle - \left\langle \overline{\mathbf{A}}^\top \mathbf{y}_1, \mathbf{z}_2 \right\rangle + \langle \mathbf{y}_2, \mathbf{z}_2 \rangle \in \mathbb{Z}.$$

First, suppose that $\mathbf{z}_2 = \mathbf{0}$. This means for any $\mathbf{z}_3 \in R^{[k]}$ we have $\langle q\mathbf{y}_1, \mathbf{z}_3 \rangle \in \mathbb{Z}$. Hence $\mathbf{y}_1 \in \frac{1}{q}\left(R^{[k]}\right)^* = \frac{1}{q}(R^*)^{[k]}$, where R^* is the dual lattice of R and is a fractional ideal. Next, suppose that $\mathbf{z}_3 = \mathbf{0}$, then for any $\mathbf{z}_2 \in R^{[l-k]}$ we have $\left\langle \mathbf{y}_2 - \overline{\mathbf{A}}^\top \mathbf{y}_1, \mathbf{z}_2 \right\rangle \in \mathbb{Z}$. Therefore $\mathbf{y}_2 - \overline{\mathbf{A}}^\top \mathbf{y}_1 \in (R^*)^{[l-k]}$. Hence, $\Lambda^\perp(\mathbf{A})^*$ can be represented as

$$\Lambda^\perp(\mathbf{A})^* = \left\{ \begin{bmatrix} \mathbf{y}_3/q \\ \mathbf{y}_4 + \overline{\mathbf{A}}^\top \mathbf{y}_3/q \end{bmatrix} \;\middle|\; \mathbf{y}_3 \in (R^*)^{[k]}, \mathbf{y}_4 \in (R^*)^{[l-k]} \right\}.$$

Now each $\mathbf{y}_3 \in (R^*)^{[k]}$ can be uniquely written as $\mathbf{y}_3 = \mathbf{y}_5 + \mathbf{y}_6$ with $\mathbf{y}_5 \in (qR^*)^{[k]}$ and $\mathbf{y}_6 \in (R_q^*)^{[k]}$. Now $[\mathbf{y}_5^\top/q \mid \mathbf{y}_4^\top + \mathbf{y}_5^\top\overline{\mathbf{A}}/q] \in (R^*)^{[l]}$. Therefore

$$\Lambda^\perp(\mathbf{A})^* = (R^*)^{[l]} + \frac{1}{q}\mathbf{A}^\top (R_q^*)^{[k]}.$$

We now estimate the value of $\rho_{1/s}\left(\Lambda^\perp(\mathbf{A})^*\right)$ for a uniformly random $\overline{\mathbf{A}}$.

$$\begin{aligned}
\mathbb{E}_{\overline{A}}\left[\rho_{1/s}\left(\Lambda^\perp(\mathbf{A})^*\right)\right] &= \frac{1}{q^{nk(l-k)}} \sum_{\overline{A}} \sum_{\mathbf{v} \in (R_q^*)^{[k]}} \rho_{1/s}\left((R^*)^{[l]} + \frac{1}{q}\mathbf{A}^\top \mathbf{v}\right) \\
&= \frac{1}{q^{nk(l-k)}} \sum_{\mathbf{u} \in (R^*)^{[l]}} \sum_{\mathbf{v} \in (R_q^*)^{[k]}} \sum_{\overline{A}} e^{-\pi s^2 ||\mathbf{u} + \mathbf{A}^\top \mathbf{v}/q||^2} \\
&= \frac{1}{q^{nk(l-k)}} \sum_{\mathbf{u}_1 \in (R^*)^{[k]}} \sum_{\mathbf{u}_2 \in (R^*)^{[l-k]}} \sum_{\mathbf{v} \in (R_q^*)^{[k]}} \sum_{\overline{A}} e^{-\pi s^2 ||\mathbf{u}_1 + \mathbf{v}/q||^2} e^{-\pi s^2 ||\mathbf{u}_2 + \overline{\mathbf{A}}^\top \mathbf{v}/q||^2} \\
&= \frac{1}{q^{nk(l-k)}} \sum_{\mathbf{u}_1 \in (R^*)^{[k]}} \sum_{\mathbf{v} \in (R_q^*)^{[k]}} \left(e^{-\pi s^2 ||\mathbf{u}_1 + \mathbf{v}/q||^2} \sum_{\mathbf{u}_2 \in (R^*)^{[l-k]}} \sum_{\overline{A}} e^{-\pi s^2 ||\mathbf{u}_2 + \overline{\mathbf{A}}^\top \mathbf{v}/q||^2} \right).
\end{aligned}$$

Because each component of $\mathbf{u}_2 + \overline{\mathbf{A}}^\top \mathbf{v}/q$ varies independently during summation, we have

$$\sum_{\mathbf{u}_2 \in (R^*)^{[l-k]}} \sum_{\overline{A}} e^{-\pi s^2 ||\mathbf{u}_2 + \overline{\mathbf{A}}^\top \mathbf{v}/q||^2} = \left(\sum_{u \in R^*} \sum_{\mathbf{a} \in R_q^{[k]}} e^{-\pi s^2 ||u + \langle \mathbf{a}, \mathbf{v} \rangle/q||^2} \right)^{l-k}.$$

Suppose that $\mathbf{v} = (v_1, \cdots, v_k)$. Define $\mathcal{I}_\mathbf{v} = v_1 R + \cdots + v_k R + qR^*$. It is a fractional ideal, the "greatest common denominator" of $v_1 R, \cdots, v_k R$ and qR^*. We have $\langle \mathbf{a}, \mathbf{v} \rangle + qu \in \mathcal{I}_\mathbf{v}$.

Since $v_1, \cdots, v_k \in R^*$, we have $qR^* \subseteq \mathcal{I}_\mathbf{v} \subseteq R^*$. Consider the quotient $\mathcal{I}_\mathbf{v}/qR^*$. By the definition of $\mathcal{I}_\mathbf{v}$, for each

equivalence class in $\mathcal{I}_{\mathbf{v}}/qR^*$ there exists at least one $\mathbf{a} \in R_q^{[k]}$ such that $\langle \mathbf{a}, \mathbf{v} \rangle$ belongs to this class. Also, suppose that for a given equivalence class, there exists two vectors $\mathbf{a}_1, \mathbf{a}_2 \in R_q^{[k]}$ such that $\langle \mathbf{a}, \mathbf{v} \rangle$ falls into this class, then $\langle \mathbf{a}_1 - \mathbf{a}_2, \mathbf{v} \rangle \in qR^*$. Hence, for any other vector $\mathbf{a}_3 \in R_q^{[k]}$, $\langle \mathbf{a}_3, \mathbf{v} \rangle$ and $\langle \mathbf{a}_3 + (\mathbf{a}_1 - \mathbf{a}_2) \bmod qR, \mathbf{v} \rangle$ must also belong to the same equivalence class. We conclude that as $\mathbf{a}$ varies over $R_q^{[k]}$ uniformly, $\langle \mathbf{a}, \mathbf{v} \rangle$ also varies over $\mathcal{I}_{\mathbf{v}}/qR^*$ uniformly. Hence we have

$$\sum_{u \in R^*} \sum_{\mathbf{a} \in R_q^{[k]}} e^{-\pi s^2 ||u + \langle \mathbf{a}, \mathbf{v} \rangle / q||^2} = \frac{q^{nk}}{|\mathcal{I}_{\mathbf{v}}/qR^*|} \cdot \rho_{1/s}(\mathcal{I}_{\mathbf{v}}/q).$$

We see that

$$\frac{1}{q^{nk(l-k)}} \left(\sum_{u \in R^*} \sum_{\mathbf{a} \in R_q^{[k]}} e^{-\pi s^2 ||u + \langle \mathbf{a}, \mathbf{v} \rangle / q||^2} \right)^{l-k} = \left(\frac{\rho_{1/s}(\mathcal{I}_{\mathbf{v}}/q)}{|\mathcal{I}_{\mathbf{v}}/qR^*|} \right)^{l-k}.$$

We get

$$\begin{aligned} \mathbb{E}_{\overline{A}} \left[\rho_{1/s} \left(\Lambda^{\perp}(\mathbf{A})^* \right) \right] &= \sum_{\mathbf{u}_1 \in (R^*)^{[k]}} \sum_{\mathbf{v} \in (R_q^*)^{[k]}} e^{-\pi s^2 ||\mathbf{u}_1 + \mathbf{v}/q||^2} \cdot \left(\frac{\rho_{1/s}(\mathcal{I}_{\mathbf{v}}/q)}{|\mathcal{I}_{\mathbf{v}}/qR^*|} \right)^{l-k} \\ &= \sum_{\mathbf{v} \in (R_q^*)^{[k]}} \rho_{1/s} \left((R^*)^{[k]} + \mathbf{v}/q \right) \cdot \left(\frac{\rho_{1/s}(\mathcal{I}_{\mathbf{v}}/q)}{|\mathcal{I}_{\mathbf{v}}/qR^*|} \right)^{l-k}. \end{aligned}$$

Let T be the set of all fractional ideal $\mathcal{I}$ of $\mathbb{Z}_K$ such that $qR^* \subseteq \mathcal{I} \subseteq R^*$. This is a finite set. Let d be an integer such that dR^* is an ideal. Then $qdR^* \subseteq d\mathcal{I}$, so $d\mathcal{I}$ is factor of qdR^*. Simply iterate through the factors of qdR^* and divide each by d. We have

$$\mathbb{E}_{\overline{A}} \left[\rho_{1/s} \left(\Lambda^{\perp}(\mathbf{A})^* \right) \right] = \sum_{\mathcal{I} \in T} \left[\left(\frac{\rho_{1/s}(\mathcal{I}/q)}{|\mathcal{I}/qR^*|} \right)^{l-k} \sum_{\substack{\mathbf{v} \in (R_q^*)^{[k]} \\ \mathcal{I}_{\mathbf{v}} = \mathcal{I}}} \rho_{1/s} \left((R^*)^{[k]} + \mathbf{v}/q \right) \right].$$

If $\mathbf{v} = \mathbf{0}$ then it is clear that $\mathcal{I}_{\mathbf{v}} = qR^*$. Conversely, suppose that $\mathcal{I}_{\mathbf{v}} = qR^*$. Then each component v_i of $\mathbf{v}$ satisfies $v_i \in qR^*$. However, we have $v_i \in R_q^*$, so $v_i = 0$ and $\mathbf{v} = \mathbf{0}$. Thus we can write $T = \{qR^*\} \cup T'$ where T' is the set of fractional ideals $\mathcal{I}$ satisfying $qR^* \subsetneq \mathcal{I} \subseteq R^*$. In the special case of $\mathcal{I} = qR^*$ we have

$$\left(\frac{\rho_{1/s}(\mathcal{I}/q)}{|\mathcal{I}/qR^*|} \right)^{l-k} \sum_{\substack{\mathbf{v} \in (R_q^*)^{[k]} \\ \mathcal{I}_{\mathbf{v}} = \mathcal{I}}} \rho_{1/s} \left((R^*)^{[k]} + \mathbf{v}/q \right) = \rho_{1/s}(R^*)^{l-k} \cdot \rho_{1/s}((R^*)^{[k]}) = \rho_{1/s}(R^*)^l.$$

So we have

$$\mathbb{E}_{\overline{A}} \left[\rho_{1/s} \left(\Lambda^{\perp}(\mathbf{A})^* \right) \right] = \rho_{1/s}(R^*)^l + \sum_{\mathcal{I} \in T'} \left[\left(\frac{\rho_{1/s}(\mathcal{I}/q)}{|\mathcal{I}/qR^*|} \right)^{l-k} \sum_{\substack{\mathbf{v} \in (R_q^*)^{[k]} \\ \mathcal{I}_{\mathbf{v}} = \mathcal{I}}} \rho_{1/s} \left((R^*)^{[k]} + \mathbf{v}/q \right) \right].$$

We can estimate the inner summation by assuming every $\mathbf{v} \neq \mathbf{0}$ with components in $\mathcal{I}$ satisfy $\mathcal{I}_{\mathbf{v}} = \mathcal{I}$. We have

$$\begin{aligned}
\mathbb{E}_{\overline{A}}\left[\rho_{1/s}\left(\Lambda^{\perp}(\mathbf{A})^*\right)\right] &\leq \rho_{1/s}(R^*)^l + \sum_{\mathcal{J}\in T'}\left[\left(\frac{\rho_{1/s}(\mathcal{I}/q)}{|\mathcal{I}/qR^*|}\right)^{l-k}\left(\rho_{1/s}(\mathcal{I}/q)^k - 1\right)\right] \\
&= \rho_{1/s}(R^*)^l + \sum_{\mathcal{J}\in T'}\frac{\rho_{1/s}(\mathcal{I}/q)^l - \rho_{1/s}(\mathcal{I}/q)^{l-k}}{|\mathcal{I}/qR^*|^{l-k}} \\
&\leq \rho_{1/s}(R^*)^l + \sum_{\mathcal{J}\in T'}\frac{\rho_{1/s}(\mathcal{I}/q)^l - 1}{|\mathcal{I}/qR^*|^{l-k}} \\
&= 1 + \sum_{\mathcal{J}\in T}\frac{\rho_{1/s}(\mathcal{I}/q)^l - 1}{|\mathcal{I}/qR^*|^{l-k}}.
\end{aligned}$$

The value of $\rho_{1/s}(\mathcal{I}/q)$ can be estimated as follows. By the arithmetic-geometric inequality, for any element $x \in \mathcal{I}$ we have

$$||x||^2 = \sum_{i=1}^{n}|\sigma_i(x)|^2 \geq n\left(\prod_{i=1}^{n}|\sigma_i(x)|\right)^{2/n} = n\cdot|\mathrm{N}^{2/n}(x)|.$$

We also have $|\mathrm{N}(x)| = \mathrm{N}(\langle x\rangle) \geq \mathrm{N}(\mathcal{I})$ since $\langle x\rangle \subseteq \mathcal{I}$. Therefore

$$\lambda_1(\mathcal{I}) \geq \sqrt{n}\cdot\mathrm{N}^{1/n}(\mathcal{I}).$$

Scaling by $1/q$ gives us

$$\lambda_1(\mathcal{I}/q) \geq \frac{\sqrt{n}\cdot\mathrm{N}^{1/n}(\mathcal{I})}{q}.$$

The dual lattice of $\mathcal{I}/q$ is $q\mathcal{I}^*$. By Lemma 38 we have

$$\eta_{2^{-2n}}(q\mathcal{I}^*) \leq q\cdot\mathrm{N}^{-1/n}(\mathcal{I}).$$

By Lemma 39 we have

$$\rho_{1/s}(\mathcal{I}/q)^l \leq \max\left(1, (\mathrm{N}(\mathcal{I})^{-1}\cdot q^n\cdot s^{-n})^l\right)(1+2^{-2n})^l.$$

By the binomial theorem,

$$(1+2^{-2n})^l = \sum_{k=0}^{l}\binom{l}{k}2^{-2nk}.$$

By the inequality

$$\binom{l}{k} \leq 2\left(\frac{l}{2}\right)^k$$

we get

$$(1+2^{-2n})^l \leq 1 + 2\sum_{k=1}^{l}(l\cdot 2^{-2n-1})^k = 1 + 2l\cdot\frac{1-(2^{-2n-1}\cdot l)^l}{2^{2n+1}-l}.$$

Since $l < 2^n$ we have $(1+2^{-2n})^l \leq 1 + l\cdot 2^{-2n+1} < 2$, and

$$\rho_{1/s}(\mathcal{I}/q)^l \leq 1 + l\cdot 2^{-2n+1} + 2(\mathrm{N}(\mathcal{I})^{-1}\cdot q^n\cdot s^{-n})^l.$$

We have

$$\mathbb{E}_{\overline{A}}\left[\rho_{1/s}\left(\Lambda^{\perp}(\mathbf{A})^{*}\right)\right] \leq 1 + l \cdot 2^{-2n+1}\left(\sum_{\mathcal{I}\in T} |\mathcal{I}/qR^{*}|^{k-l}\right) + 2s^{-nl}\sum_{\mathcal{I}\in T}\frac{(q^{n}\cdot \mathrm{N}(\mathcal{I})^{-1})^{l}}{|\mathcal{I}/qR^{*}|^{l-k}}.$$

Notice that

$$|\mathcal{I}/qR^{*}| = \frac{d(qR^{*})}{d(\mathcal{I})} = \frac{q^{n}}{\sqrt{|\Delta_{K_m}|}\cdot d(\mathcal{I})} = \frac{q^{n}\cdot d(\mathcal{I})^{-1}}{\sqrt{|\Delta_{K_m}|}} = q^{n}\cdot \mathrm{N}(\mathcal{I}^{*}) = \mathrm{N}(q\mathcal{I}^{*}).$$

Since $qR^{*} \subseteq \mathcal{I} \subseteq R^{*}$, by Lemma 11, $R \subseteq \mathcal{I}^{*} \subseteq R/q$, hence $qR \subseteq q\mathcal{I}^{*} \subseteq R$. We see that $q\mathcal{I}^{*}$ is an ideal and is a factor of $\langle q\rangle$. We also have

$$q^{n}\cdot \mathrm{N}(\mathcal{I})^{-1} = q^{n}\cdot\frac{\sqrt{|\Delta_{K_m}|}}{d(\mathcal{I})} = d(q\mathcal{I}^{*})\cdot\sqrt{|\Delta_{K_m}|} = \mathrm{N}(q\mathcal{I}^{*})\cdot|\Delta_{K_m}|.$$

Therefore

$$\frac{(q^{n}\cdot \mathrm{N}(\mathcal{I})^{-1})^{l}}{|\mathcal{I}/qR^{*}|^{l-k}} = |\Delta_{K_m}|^{l}\cdot \mathrm{N}(q\mathcal{I}^{*})^{k} \leq n^{nl}\cdot \mathrm{N}(q\mathcal{I}^{*})^{k}.$$

Let S be the set of ideals that are factors of $\langle q\rangle$. We first claim that

$$\sum_{\mathcal{I}\in T}|\mathcal{I}/qR^{*}|^{k-l} \leq 1 + \frac{|S|-1}{2^{l-k}} \leq 1 + \frac{q^{2n}}{2^{l-k}},$$

so the first sum becomes negligible as l increases. Notice that $k = l$ then the sum is equal to $|S|$. Thus we begin by estimating the value of $|S|$.

If $q = q_1q_2$ where q_1, q_2 are two coprime integers, then $\langle q_1\rangle, \langle q_2\rangle$ are also coprime. Let S_1, S_2 be the set of factors of $\langle q_1\rangle, \langle q_2\rangle$ respectively. We have $|S| = |S_1|\cdot|S_2|$. In fact we more generally have

$$\sum_{\mathcal{I}|\langle q\rangle}\mathrm{N}(\mathcal{I})^{k} = \left(\sum_{\mathcal{I}|\langle q_1\rangle}\mathrm{N}(\mathcal{I})^{k}\right)\left(\sum_{\mathcal{I}|\langle q_2\rangle}\mathrm{N}(\mathcal{I})^{k}\right).$$

Therefore it sufficies to consider the case where q is a prime power. Let $q = p^{t}$. By Lemma 93, $\langle q\rangle$ factors as

$$\langle p\rangle = \mathfrak{p}_1^{th}\cdots\mathfrak{p}_d^{th}$$

where r_p is the largest integer such that p^{r_p} divides m, f_p is the multiplicative order of p modulo m/p^{r_p}, $h = \varphi(p^{r_p})$, and $d = n/hf$. Hence $|S| = (th+1)^{d} \leq (2th)^{d}$. Notice that $hd \leq n$. Let $f(x) = (2tx)^{n/x}$. This function reaches its maximum when $2tx = e$. Therefore $|S| \leq \exp(2tn/e) \leq q^{2n}$. The second inequality is because $\exp(1/e) \approx 1.44 < p$, and $q = p^{k}$.

Among the elements of S there is one special element, namely R itself. We have $\mathrm{N}(R) = 1$, and $1^{k-l} = 1$ regardless of k and l. Except this special element, all other elements have $\mathrm{N}(\mathcal{I}) \geq 2$, and so $\mathrm{N}(\mathcal{I})^{k-l} \leq 2^{k-l}$. Together we get

$$\sum_{\mathcal{I}\in T}|\mathcal{I}/qR^{*}|^{k-l} \leq 1 + \frac{q^{2n}}{2^{l-k}}.$$

To bound the second sum, notice that

$$2s^{-nl}\sum_{\mathcal{I}\in T}\frac{(q^{n}\cdot \mathrm{N}(\mathcal{I})^{-1})^{l}}{|\mathcal{I}/qR^{*}|^{l-k}} \leq 2(s/n)^{-nl}\sum_{\mathcal{I}\in S}\mathrm{N}(\mathcal{I})^{k}.$$

And we have

$$\begin{aligned}\sum_{\mathcal{J}\in S}\mathrm{N}(\mathcal{J})^k &= \prod_{i=1}^{d}(1+\mathrm{N}(\mathfrak{p}_i)^k+\cdots+\mathrm{N}(\mathfrak{p}_i)^{thk})\\ &= \left(1+p^{f_p k}+\cdots+p^{f_p thk}\right)^d\\ &\le p^{f_p thkd}(1-p^{-f_p k})^{-d}\\ &\le q^{nk}\exp(d\cdot p^{-f_p k}). \quad (f_p hd = n, p^t = q)\end{aligned}$$

Finally, we have $p^{f_p} \ge m/p^{r_p}$ and $g \le n/\varphi(p^{r_p}) = \varphi(m/p^{r_p})$, so $g\cdot p^{-f_p k} \le 1$, and the sum is bounded.

To conclude, what we have proven is that, with a uniformly random $\overline{\mathbf{A}}$, we have

$$\mathbb{E}_{\overline{A}}\left[\rho_{1/s}\left(\Lambda^{\perp}(\mathbf{A})^*\right)\right] \le 1 + l\cdot 2^{-2n+1}(1+q^{2n}/2^{l-k}) + 2(s/n)^{-nl}q^{kn+2}.$$

And so with suitably large l and s, $\rho_{1/s}\left(\Lambda^{\perp}(\mathbf{A})^* \setminus \{\mathbf{0}\}\right)$ becomes negligible. Thus the probabilitiy distribution of

$$\overline{\mathbf{A}}\mathbf{s} + \mathbf{e} \bmod Rq$$

where $\mathbf{s} \in R_q^{[l-k]}, \mathbf{e} \in R_q^{[k]}$, is close to uniformly random.

References

W. Banaszczyk. 1993. New bounds in some transference theorems in the geometry of numbers. *Math. Ann.* 296 (1993), 625–635. https://doi.org/10.1007/BF01445125

Keith Conrad. [n. d.]. The Conductor Ideal of an Order. https://kconrad.math.uconn.edu/blurbs/gradnumthy/conductor.pdf.

Loukas Grafakos. 2014. *Classical Fourier Analysis*. Springer, New York, NY. https://doi.org/10.1007/978-1-4939-1194-3

Frazer Jarvis. 2014. *Algebraic Number Theory*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-319-07545-7

Vadim Lyubashevsky, Chris Peikert, and Oded Regev. 2013. A Toolkit for Ring-LWE Cryptography. In *Advances in Cryptology – EUROCRYPT 2013*, Thomas Johansson and Phong Q. Nguyen (Eds.). Springer, Berlin, Heidelberg, 35–54. https://doi.org/10.1007/978-3-642-38348-9_3

D. Micciancio and O. Regev. 2004. Worst-case to average-case reductions based on Gaussian measures. In *45th Annual IEEE Symposium on Foundations of Computer Science*. IEEE, Piscataway, NJ, 372–381. https://doi.org/10.1109/FOCS.2004.72

Jürgen Neukirch. 1999. *Algebraic Number Theory*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-03983-0

Jerry Shurman. [n. d.]. Cyclotomic Integer Rings, In General. https://people.reed.edu/~jerry/361/lectures/discresults.pdf.

Barry Simon. 2015. *A Comprehensive Course in Analysis, Part I: Real Analysis*. American Mathematical Society, Providence, RI.