Lattices Part II — Dual Lattices, Fourier Transform, Smoothing Parameter, Public Key Encryption

Boaz Barak

May 2, 2008

The first two sections are based on Oded Regev's lecture notes, and the third one on his paper "New Lattice Based Cryptographic Constructions " (JACM 2004, preliminary version STOC 2003) and some personal communication with him.

1 Dual Lattices and Fourier Transform

Dual lattice If L is a lattice then $L^* = \{\mathbf{u} : \forall \mathbf{v} \in L \langle \mathbf{u}, \mathbf{v} \rangle \in \mathbb{Z}\}$. If B is a basis for L then $(B^{-1})^T$ is a basis for L^*. Indeed, for any $\mathbf{b} \in \mathbb{Z}^n$, $(B^{-1})^T \mathbf{b} \in L^*$ since for any $\mathbf{a} \in \mathbb{Z}^n$,

$$\left((B^{-1})^T \mathbf{b}\right)^T B\mathbf{a} = \mathbf{b}^T B^{-1} B\mathbf{a} = \langle \mathbf{a}, \mathbf{b} \rangle \in \mathbb{Z}$$

Similarly it can be shown that any vector in L^* can be obtained by integer combinations of the columns of $(B^{-1})^T$. As a corollary we obtain that $det(L^*) = 1/det(L)$ and $(L^*)^* = L$.

Fourier transform Consider the interval $[0, \lambda]$ and suppose that we identify the point λ with 0 (i.e., think of it as a Torus and work modulo λ). Another way to think about this is as the basic cell of the lattice $\lambda\mathbb{Z}$, whose dual is the lattice $(1/\lambda)\mathbb{Z}$. A periodic function on this torus has to period length of the form λ/n for an integer n. Thus, the Fourier transform of a function on this torus involves representing it as a sum of functions of the form $x \mapsto e^{-2\pi i n x/\lambda}$.

More generally, the Fourier transform of a function f on $\mathcal{P}(L)$ represents f as the sum of functions of the form $\mathbf{x} \mapsto e^{2\pi i \langle \mathbf{x}, \mathbf{y} \rangle}$ where $\mathbf{y}$ is an element in L^*.

That is, we have the following theorem:

Theorem 1.1. *Let $f : \mathcal{P}(B) \to \mathbb{R}$ be a nice (continuous, differentiable, integrable etc..) function. Then for every $\mathbf{x} \in \mathcal{P}(B)$,*

$$f(\mathbf{x}) = \sum_{\mathbf{y} \in L^*} \hat{f}(\mathbf{y}) e^{2\pi i \langle \mathbf{x}, \mathbf{y} \rangle} \tag{1}$$

where

$$\hat{f}(\mathbf{y}) = \frac{1}{det(L)} \int_{\mathcal{P}(B)} f(x) e^{-2\pi i \langle \mathbf{x}, \mathbf{y} \rangle} d\mathbf{x}$$

More generally, if f is an *L-periodic function* (i.e., $f(\mathbf{x}) = f(\mathbf{x} + \mathbf{z})$ for every $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{z} \in L$) then (1) holds for every $\mathbf{x} \in \mathbb{R}^n$.

Gaussian Let $\rho(\mathbf{x}) = e^{-\pi \|\mathbf{x}\|_2^2}$ (density function of Gaussian with standard deviation $1/\sqrt{2\pi}$). Then $\hat{\rho} = \rho$. If we let $\rho_s(\mathbf{x}) = e^{-\pi \|\mathbf{x}/s\|_2^2}$ then $\hat{\rho_s} = s^n \rho_{1/s}$. We have the following claim:

Claim 1.2. *For any lattice L,*

$$\sum_{\mathbf{z}\in L}\rho(\mathbf{z}/s) = s^n \sum_{\mathbf{z}\in L}\rho(\mathbf{z})$$

Proof. Consider the periodic function $f(\mathbf{x}) = \sum_{\mathbf{z}\in L}\rho(\mathbf{x}+\mathbf{z})$. Then

$$\sum_{\mathbf{z}\in L} g(\mathbf{z}) = f(\mathbf{0}) = \sum_{\mathbf{y}\in L^*}\hat{f}(\mathbf{y})$$

But for every $\mathbf{y}\in L^*$,

$$\hat{f}(\mathbf{y}) = \tfrac{1}{det(L)}\int_{\mathcal{P}(B)} f(\mathbf{x})e^{-2\pi i\langle \mathbf{x},\mathbf{y}\rangle}d\mathbf{x} =$$

$$\frac{1}{det(L)}\sum_{\mathbf{z}\in L}\int_{\mathcal{P}(B)}\rho((\mathbf{x}+\mathbf{z})/s)e^{-2\pi i\langle \mathbf{x},\mathbf{y}\rangle}d\mathbf{x} = \text{ (change of var } \mathbf{x}+\mathbf{z}/s\mapsto\mathbf{x}\text{)}$$

$$s^n\tfrac{1}{det(L)}\sum_{\mathbf{z}\in L}\int_{\mathcal{P}(B)+\mathbf{z}}\rho(\mathbf{x})e^{-2\pi i\langle \mathbf{x}-\mathbf{z},\mathbf{y}\rangle}d\mathbf{x} = (\langle \mathbf{z},\mathbf{y}\rangle\in\mathbb{Z})$$

$$s^n\tfrac{1}{det(L)}\sum_{\mathbf{z}\in L}\int_{\mathcal{P}(B)+\mathbf{z}}\rho(\mathbf{x})e^{-2\pi i\langle \mathbf{x},\mathbf{y}\rangle}d\mathbf{x} =$$

$$s^n\tfrac{1}{det(L)}\int_{\mathbb{R}^n}\rho(\mathbf{x})e^{-2\pi i\langle \mathbf{x},\mathbf{y}\rangle}d\mathbf{x} = s^n\tfrac{1}{det(L)}\rho(\mathbf{y})$$

Hence we get that

$$\sum_{\mathbf{x}\in L}\rho_s(\mathbf{x}) = s^n\tfrac{1}{det(L^*)}\sum_{\mathbf{y}\in L^*}\rho(\mathbf{y})$$

Applying this equality again we get the RHS is equal to

$$s^n\tfrac{1}{det(L^*)}\tfrac{1}{det(L)}\sum_{\mathbf{x}\in L}\rho(\mathbf{x})$$

□

Summation over lattices In fact, the a similar proof to the one above yields a more general statement:

Lemma 1.3. *For every function $g:\mathbb{R}^n\to\mathbb{R}$ (not necessarily periodic)*

$$\sum_{\mathbf{x}\in L} g(\mathbf{x}) = det(L^*)\sum_{\mathbf{y}\in L^*}\hat{g}(\mathbf{y})$$

where $\hat{g}(\mathbf{y}) = \int_{\mathbb{R}^n} g(\mathbf{x})e^{-2\pi i\langle \mathbf{x},\mathbf{y}\rangle}d\mathbf{x}$.

2 Smoothing parameter

Gaussian For any s let $\mathcal{G}_s$ be the Gaussian distribution of vectors with expected norm s, obtained by taking n independent Gaussians with mean zero and standard deviation $s/\sqrt{n}$. The density function of this distribution is $\rho_{\sqrt{2\pi}s/\sqrt{n}}(\mathbf{x})$ where $\rho_s(\mathbf{x}) = s\rho(\mathbf{x}/s)$.

Adding noise to a lattice point The distribution $\mathcal{D}_s$ over $\mathcal{P}(B)$ is defined to be $\mathcal{G}_s \pmod{\mathcal{P}(B)}$. That is, for every $\mathbf{x} \in \mathcal{P}(B)$, $\mathcal{D}_s(\mathbf{x}) = \sum_{\mathbf{z}\in L} \mathcal{G}_s(\mathbf{x}+\mathbf{z}) = \sum_{\mathbf{z}\in L} \mathcal{G}_s(\mathbf{x}-\mathbf{z})$. Note that we can extend the latter definition to any point $\mathbf{x} \in \mathbb{R}^n$ to obtain an L-periodic function, which we can think of as the "distribution" obtained by taking a random lattice point and adding to it Gaussian noise.

Smoothing parameter The *smoothing paremeter* of L is the smallest s such that $\mathcal{D}_s$ is $2^{-n/100}$ statistically-close to the uniform distribution on $\mathcal{P}(B)$.

Transferance Theorem The following theorem is very useful in relating lattice parameters to one another:

Theorem 2.1. *The following parameters are equivalent to one another up to a multiplicative factor of $O(n)$:*

1. *The covering radius of the lattice: smallest r such that $dist(\mathbf{x}, L) \leq r$ for every $\mathbf{x} \in \mathbb{R}^n$.*
2. *The smoothing parameter of the lattice.*
3. *The length of the shortest independent vector collection: $\lambda_n(L)$.*
4. *The inverse of the shortest dual vector $1/\lambda_1(L^*)$.*

(These parameters are also roughly equivalent to the length of shortest basis of L: minimum over all bases $\mathbf{b}_1, \ldots, \mathbf{b}_n$ of L of $\max_i \|\mathbf{b}_i\|_2$, though we won't show that.)

It's easy to see that the smoothing parameter is larger than the covering radius.

It's also not hard to see that $1/\lambda_1(L^*) \leq \lambda_n(L)$. Indeed, if $\mathbf{u}$ is the shortest vector in L^* and $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are n independent vectors in L with $\|\mathbf{v}_i\|_2 \leq \lambda_n(L)$ then there must exist i such that $\langle \mathbf{u}, \mathbf{v}_i \rangle \neq 0$ and hence (since this is an integer) $|\langle \mathbf{u}, \mathbf{v}_i \rangle| \geq 1$. But by Cauchy Schwartz this means that $1 \leq \|\mathbf{u}\|_2 \|\mathbf{v}_i\|_2 \leq \lambda_1(L^*)\lambda_n(L)$.

Moreover one can show that the covering radius is at least $\lambda_n(L)/2$. Indeed, let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be a collection of linearly independent vectors in L such that $\mathbf{v}_n = \lambda_n(L)$. Then we claim that the vector $\mathbf{u} = \mathbf{v}_n/2$ has distance at least $\lambda_n(L)/2$ from L. Indeed if there exist $\mathbf{w} \in L$ such that $\|\mathbf{w} - \mathbf{u}\|_2 < \lambda_n(L)/2$ then both $\mathbf{w}$ and $\mathbf{w} - \mathbf{v}_n$ are lattice vectors with norm less than $\lambda_n(L)$, but they cannot be both in the hyperplane spanned by $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$ since that would imply $\mathbf{v}_n$ is in this hyperplane.

Proof of transferance theorem This means that the transferance theorem follows from the following theorem:

Theorem 2.2. *The smoothing parameter of L is at most $100n/\lambda_1(L^*)$.*

Proof. By scaling we may assume that $\lambda_1(L^*) = 10\sqrt{n}$. This means that it suffices to show that the Guassian distribution with standard deviation $\sqrt{n}$ (density function $e^{-\pi\|\mathbf{x}\|_2^2}$) is close to the uniform distribution modulo $\mathcal{P}(B)$. Since the Fourier transform of the uniform distribution $\mathcal{U}$ satisfied $\hat{\mathcal{U}}(\mathbf{0}) = 1$ and $\hat{\mathcal{U}}(\mathbf{y}) = 0$ for all $\mathbf{y} \neq \mathbf{0}$, and since the Gaussian distribution is its own Fourier transform, it means that we need to bound

$$\frac{1}{det(L)} \int_{\mathcal{P}(B)} \sum_{\mathbf{y}\in L^*\setminus\{\mathbf{0}\}} \rho(\mathbf{y}) e^{-2\pi i\langle \mathbf{x},\mathbf{y}\rangle} d\mathbf{x} =$$

$$\sum_{\mathbf{y}\in L^*\setminus\{\mathbf{0}\}} \rho(\mathbf{y}) \frac{1}{det(L)} \int_{\mathcal{P}(B)} e^{-2\pi i\langle \mathbf{x},\mathbf{y}\rangle} d\mathbf{x} \leq \sum_{\mathbf{y}\in L^*\setminus\{\mathbf{0}\}} \rho(\mathbf{y}) \frac{1}{det(L)} \left| \int_{\mathcal{P}(B)} e^{-2\pi i\langle \mathbf{x},\mathbf{y}\rangle} d\mathbf{x} \right| \leq \sum_{\mathbf{y}\in L^*\setminus\{\mathbf{0}\}} \rho(\mathbf{y})$$

But

$$\sum_{\mathbf{y}\in L^*\setminus\{\mathbf{0}\}} \rho(\mathbf{y}) = \sum_{\mathbf{y}\in L^*\setminus\{\mathbf{0}\}} e^{-\pi\|\mathbf{y}\|_2^2} \leq \sum_{\mathbf{y}\in L^*\setminus\{\mathbf{0}\}} e^{-\pi\|\mathbf{y}/2\|_2^2-(3/4)(10\sqrt{n})^2} =$$
$$e^{-75n}\sum_{\mathbf{y}\in L^*\setminus\{\mathbf{0}\}} e^{-\pi\|\mathbf{y}/2\|_2^2} \leq e^{-75n}\sum_{\mathbf{y}\in L^*} e^{-\pi\|\mathbf{y}/2\|_2^2} = e^{-75n}2^n\sum_{\mathbf{y}\in L^*}\rho(\mathbf{y})$$

Hence letting $X = \sum_{\mathbf{y}\in L^*\setminus\{\mathbf{0}\}}\rho(\mathbf{y})$ we get that

$$X \leq 2^{-70n}\left(\rho(\mathbf{0}) + X\right)$$

In particular, $2X \leq \rho(\mathbf{0})2^{-70n}$ but since $\rho(\mathbf{0}) \leq 1$ we complete the proof. □

3 Regev's First Cryptosystem

Main Theorem Suppose that some lattice problem is worst-case hard, then for every $m = \text{poly}(n)$, no polynomial-time algorithm can distinguish whether m samples come from:

1. The uniform distribution on $[0,1)^n$.
2. The distribution $\mathcal{T}_{\mathbf{u}}$ which is obtained by taking the uniform distribution over all vectors $\mathbf{v}$ in $[0,1)^n$ such that $\langle\mathbf{u},\mathbf{v}\rangle \in \mathbb{Z}$ and adding to it the Gaussian distribution with standard deviation $1/(n^4\|\mathbf{u}\|_2)$.[1] The vector $\mathbf{u}$ has integer coordinates with a random direction and norm chosen uniformly from $[10^n, 100^n]$.

Corollary The following is a secure public key encryption: private key is $\mathbf{u}$, public key is vectors $\mathbf{v}_1,\ldots,\mathbf{v}_{n^2}$ chosen uniformly from $\mathcal{T}_{\mathbf{u}}$. To encrypt one, choose randomly $S \subseteq [n^2]$ and output $\sum_{i\in S}\mathbf{v}_i \pmod{[0,1)^n}$ and add to it some slight noise, namely a Gaussian with deviation $2^{-n^{1.5}}$. To encrypt zero, output a random vector in $[0,1)^n$. When decrypting $\mathbf{z}$, output zero if $\langle\mathbf{z},\mathbf{u}\rangle \pmod 1 \in (0.1, 0.9)$.

Proof of Main Theorem We'll reduce from the following promise problem on lattices— input is a lattice L.

Yes instance L^* has a nonzero vector $\mathbf{u}$ of length $\Delta = 1/n^{10}$ and all other vectors in L^* that are not proportional to $\mathbf{u}$ have length at least n.

No instance All nonzero vectors in L^* have length at least n.

The lattice is given by a basis with rational coordinates with common denominator 2^n and numerators in $[-2^{2n}, +2^{2n}]$. Regev showed that this problem is equivalent to the approximating the unique shortest vector problem up to a certain polynomial factor.[2]

Distribution We let $\mathcal{D}_L$ be the periodic function over $\mathbb{R}^n$ obtained by adding n^4-standard deviation Guassian noise to each lattice point. Using the relation between the smoothing parameter of L and the inverse of the shortest vector of L^* we prove:

1. In the **No** case, $\mathcal{D}_L$ is exponentially close to the uniform distribution U over $\mathbb{R}^n$ (i.e., the constant function 1).

[1] By Guassian distribution with std s we mean that we choose n independent gaussians so that the expected norm of the resulting vector is s.

[2] Actually Regev only showed this where Δ is not known exactly but up to a factor of two, we'll tackle this issue below.

2. In the **Yes** case, $\mathcal{D}_L$ is exponentially close to the distribution $\mathcal{T}_L$ that is obtained by adding n^4 standard deviation Gaussian noise to the uniform distribution over the the hyperplanes $\{H_k\}_{k\in\mathbb{Z}}$ where $H_k = \{\mathbf{v} : \langle \mathbf{v}, \mathbf{u}\rangle = k\}$. That is, $\mathcal{T}_L(\mathbf{v}) \sim e^{-(\langle \mathbf{v},\mathbf{u}\rangle \mod 1)^2/n^8}$

Therefore, taking any bounded nice shape C in R^n such that we can (approximately) sample a random lattice point in C, under our assumption it is hard to distinguish the distribution obtained by restricting $\mathcal{T}_L$ to C and the uniform distribution on C.

Completing the proof We now complete the reduction. Apply a random rotation to the lattice— this makes the shortest vector $\mathbf{u}$ have a random direction.

Let N be an number chosen at random in $[10^n/\Delta, 100^n/\Delta]$ and let $\mathbf{e}_1$ be a lattice vector that is 2^n-close to $(N, 0, \ldots, 0)$ (such a vector can be found using LLL). Define $\mathbf{e}_2, \ldots, \mathbf{e}_n$ similarly. Let C be the parallelepiped of $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Note that (1) these are nearly orthogonal vectors and so in particular linearly independent and (2) we can sample a random lattice vector in C by taking a random linear combination of the basis vectors with coefficients from a large enough range and reducing the resulting vector modulo C (because the $\mathbf{e}_i$'s are lattice vectors, the resulting vector will stay in the lattice).

We map this parallelepiped C to $[0,1)^n$ using the linear transformation T that maps $\mathbf{e}_i$ to the i^{th} standard basis. Let $\mathbf{u}'$ be the vector $(\beta_1, \ldots, \beta_n)$ where $\beta_i = \langle \mathbf{e}_i, \mathbf{u}\rangle$. Note that β_i is an integer and is equal to $N\mathbf{u}_i$ (where $\mathbf{u}_i$ denotes the i^{th} coordinate of $\mathbf{u}$) up to a multiplicative factor of (1 ± 2^{-n}).

Letting $\mathcal{T}_{L,C}$ denoting the restriction of $\mathcal{T}_L$ to C, we claim that $T(\mathcal{T}_{L,C})$ is equal to the distribution $\mathcal{T}_{\mathbf{u}'}$ on $[0,1]^n$. Indeed, for every point $\mathbf{v} \in C$, writing $\mathbf{v} = \sum \alpha_i \mathbf{e}_i$, we see that $\langle \mathbf{v}, \mathbf{u}\rangle = \sum \alpha_i \langle \mathbf{e}_i, \mathbf{u}\rangle = \sum \alpha_i \beta_i = \langle T(\mathbf{v}), \mathbf{u}'\rangle)$ and hence $\mathcal{T}_L(\mathbf{v})$ is proportional to $\mathcal{T}_{\mathbf{u}'}(T(\mathbf{v})$. Since the transformation T is a linear function (with fixed derivative), this means that the two distributions are proportional and hence equal.

Now up to a multiplicative 1 ± 2^{-n} factor, $\mathbf{u}'$ is equal to $N\mathbf{u}$ in every coordinate and hence up to this factor $\mathbf{u}'$ has uniform Guassian direction, and a random length in range $[10^n, 100^n]$ completing the proof. □

Dealing with unknown Δ The above assumed that we know Δ while in reality we'll only know that it's in the interval $[1/n^{10}, 2/n^{10}]$. Still if we have an distinguisher D we can determine the probability that it outputs 1 in the uniform distribution, and then run it many times with many guesses for Δ. If in even one of these times it deviates from this probability then we know that we are in the Yes case. If the adversary has ϵ advantage in guessing then we'll not need more than $O(1/\epsilon)$ repetitions.

Reducing to one dimension Regev's actual cryptosystem was one dimensional. We can show that the one-dimensional problem is also hard by projecting $[0,1)^n$ to $[0,1)$. The idea is to partition $[0,1)^{n-1}$ to very small equal sized sets $S_1, \ldots, S_M$ for some (exponentially large M) such that S_i is extremely close to S_{i+1} and then we project the set $S_i \times [0,1)$ to the i^{th} interval $[(i-1)/M, i/M)$ of length $1/M$ in $[0,1)$. That is, if we get a sample $\mathbf{x}$ from a distribution on $[0,1)^n$, then we map $\mathbf{x}$ to $(i-1)/M + \mathbf{x}_n$ where i is the index of S_i such that $(\mathbf{x}_1, \ldots, \mathbf{x}_{n-1}) \in S_i$. (We choose a nice enough partition so that i is easy to compute.)