6.876 Advanced Topics in Cryptography: Lattices October 21, 2015

Lecture 12

Lecturer: Vinod Vaikuntanathan *Scribe: Srinivasan Raghuraman*

In this lecture, we will present a "proof" of the Smoothing Lemma used in the worst-case to average-case reduction from the last lecture and then move on to present constructions of cryptographic objects from the SIS problem. In later lectures, we will explore constructions of more cryptographic objects from the LWE problem (because we can construct more objects with it!). We assume throughout the lecture that lattices under consideration are of full rank.

We begin with the construction of one-way collision-resistant hash functions which naturally arise from the SIS problem.

1 One-Way CRHFs from SIS

We recap the SIS problem:

Definition 1 ($\text{SIS}_{n,m,q,\beta}$)**.** *Given a random matrix $A \in \mathbb{Z}^{n \times m}$, find a short non-zero vector $e \in \mathbb{Z}^m$ such that $Ae = 0 \bmod q$ and $\|e\| \leq \beta$.*

We will be using the paramters $m > n \log q$ for reasons discussed in the previous lecture, and the ℓ_2-norm. We next define collision-resistant hash functions. We first define the notion of "negligible functions" which are simply functions which decay faster than any inverse polynomial.

Definition 2 (Negligible Functions)**.** *For every $n \in \mathbb{N}$, we denote by $\mathsf{negl}(n)$ the class of functions said to be negligible in n, where if $\mathcal{P}$ denotes the set of all univariate polynomials with real coefficients,*

$$\mathsf{negl}(n) = \{f : \mathbb{N} \to \mathbb{R} \mid \forall p \in \mathcal{P}, \exists N_0(p), \textit{such that } \forall n \geq N_0(p), f(n) < 1/p(n)\}$$

Although $\mathsf{negl}(n)$ denotes a set of functions, by abusing notation, if a function is a negligible function, that is it belongs to $\mathsf{negl}(n)$, we may sometimes write it as being equal to $\mathsf{negl}(n)$.

Definition 3 (Collision-resistant Hash Functions)**.** *A family of collision-resistant hash functions, denoted by $\{\mathcal{H}_n\}_{n \in \mathbb{N}}$, satisfies the properties that for any $n \in \mathbb{N}$, every $h \in \mathcal{H}_n$ is a function $h : \{0,1\}^n \to \{0,1\}^{m(n)}$, where $m(n) < n$ (h is a "shrinking" function), and that for all families of polynomial-sized circuits $\{\mathcal{A}_n\}_{n \in \mathbb{N}}$, for every $n \in \mathbb{N}$,*

$$\Pr_{h \leftarrow \mathcal{H}_n}[\mathcal{A}_n(h) = (x, y) : x \neq y \wedge h(x) = h(y)] = \mathsf{negl}(n)$$

Remarks:

1. For one to make any use of a family of collision-resistant hash functions or CRHFs, we require that the process of sampling a hash function, that is, the process $h \leftarrow \mathcal{H}_n$, as well as computing a hash function on an input be efficient (run in polynomial time in the security parameter, say n).
2. The above definition is inherently non-uniform. The reason why we need to have a family of functions as opposed to one in this case is that a non-uniform circuit can simply hardcode collisions. While this is usual of definitions in cryptography, there is no need for the same. One could work with a uniform definition where the adversaries are uniform Turing machines, although in reduction, one has to ensure that the reduction to security is uniform as well.
3. If one is willing to stick to a non-uniform definition, then there are ways to convert a single hash function into a family of hash functions which would now be collision-resistant.

We next define one-way functions.

Definition 4 (One-Way Functions (OWFs))**.** *A family of one-way functions, denoted by* $\{f_n\}_{n\in\mathbb{N}}$*, satisfies the properties that for any* $n \in \mathbb{N}$*,* $f_n : \{0,1\}^n \to \{0,1\}^*$*, and that for all families of polynomial-sized circuits* $\{\mathcal{A}_n\}_{n\in\mathbb{N}}$*, for every* $n \in \mathbb{N}$*,*

$$\Pr_{x\leftarrow\{0,1\}^n}[\mathcal{A}_n(f_n(x)) = x' : x' \in \{0,1\}^n \wedge f_n(x) = f_n(x')] = \mathsf{negl}(n)$$

Collision-resistance implies one-wayness.[1] We can show this through a reduction. Assume that we sample $h \leftarrow \mathcal{H}_n$, $x \leftarrow \{0,1\}^n$ and give $h_n(x)$. If an adversary succeeds in breaking one-wayness with high probability, he returns an $x' \in \{0,1\}^n$ such that $h(x) = h(x')$. If $x \neq x'$, then the adversary has actually found a collision and hence breaks the collision-resistance property of the family of hash functions. One can show that with very high probability $x \neq x'$ since the hash function is a shrinking function and hence there is no way to tell whether x itself or an $x \neq x'$ was used to compute $h(x)$.

Hence, it suffices to prove that the hash family we are going to construct is collision-resistant, and one-wayness follows from above.

1.1 The construction from SIS

For every $n \in \mathbb{N}$, fix $q = q(n)$ and $m = m(n)$. Let $\mathcal{H}_n$ be the family of functions $\{f_A\}_{A\in\mathbb{Z}^{n\times m}}$ indexed by A such that $f_A : \{0,1\}^m \to \mathbb{Z}_q^n$ and for $e \in \{0,1\}^n$, $f_A(e) = Ae \bmod q$ (find the sum of a subset of columns mod q). Again, we will be working with paramaters $m > n \log q$ so that f_A is shrinking. Note that this construction arises naturally from the definition of the SIS problem.

Lemma 5. *If an adversary succeeds in breaking the collision resistance of the above scheme with probability* ϵ*, then there exists an adversary which can solve the* $SIS_{n,m,q,\beta=\sqrt{m}}$ *problem with probability* ϵ*.*

Proof. Fix $n \in \mathbb{N}$. Suppose an adversary $\mathcal{A}$ on input f_A for a random $A \leftarrow \mathbb{Z}^{n\times m}$ returns $e, e' \in \mathbb{Z}^m$ with $e \neq e'$ and $Ae = Ae' \bmod q$. Then, consider the vector $e - e' \in \{-1, 0, 1\}^m \subset \mathbb{Z}^m$. We have $A(e - e') = 0 \bmod q$ and $\|e - e'\| \leq \sqrt{m} = \beta$. Thus, we can construct an adversary for the $\mathrm{SIS}_{n,m,q,\beta=\sqrt{m}}$ problem which given the random matrix A runs $\mathcal{A}$ on the description of f_A and outputs the difference of the two vectors $\mathcal{A}$ outputs as its solution. Clearly this adversary succeeds in solving the $\mathrm{SIS}_{n,m,q,\beta=\sqrt{m}}$ problem with the same probability that $\mathcal{A}$ succeeds in breaking the collision resistance of the above scheme. □

Remark. The above scheme is secure under a reduction to SIS with $\beta = \sqrt{m}$. In an attempt to improve this, one could consider improving the worst-case to average-case reduction, since solving SIS with $\beta = \sqrt{m}$ enables solving $\mathrm{SIVP}_{\tilde{\mathcal{O}}(n)\approx\beta\sqrt{n}}$.

Great, so we now can construct CRHFs and hence OWFs from SIS. It is known that OWFs imply a bunch of cryptographic primitives (which we do not define here) such as digital signatures, pseudorandom generators, pseudorandom functions, symmetric-key encryption and so on. So we can in principal construct all these primitives from SIS. There are two more questions one can ask at this point.

1. Is it possible to construct public-key cryptographic primitives from SIS?
2. Although OWFs do imply the above primitives, the reductions from OWFs would be inefficient in terms of the parameters one would achieve. Is it possible to construct them from SIS directly with more efficient parameters?

[1] In fact, historically, the above scheme was initially proposed as being a one-way function [Ajt96].

2 A Pseudo-proof of the Smoothing Lemma

We first state the Lemma. For completeness, we define a few notions.

Definition 6 (Statistical Distance)**.** *Let $\mathcal{X}, \mathcal{Y}$ be two probability distributions over some set S. Their* statistical distance *is*

$$\Delta(\mathcal{X}, \mathcal{Y}) \triangleq \frac{1}{2} \sum_{s \in S} \left| \Pr_{\mathcal{X}}[s] - \Pr_{\mathcal{Y}}[s] \right|$$

Analogously, if $\mathcal{X}, \mathcal{Y}$ are two probability distributions over a continuous domain $\mathcal{D}$, then

$$\Delta(\mathcal{X}, \mathcal{Y}) \triangleq \frac{1}{2} \int_{\mathcal{D}} \left| \Pr_{\mathcal{X}}[s] - \Pr_{\mathcal{Y}}[s] \right| \, ds$$

Definition 7 (Gaussian Distribution)**.** *The n-dimensional Gaussian distribution with zero mean and standard deviation paramater s is defined as*

$$\rho_s(x) = \frac{1}{s^n} \mathbf{exp}\left(-\pi \frac{\|x\|^2}{s^2} \right)$$

where $x \in \mathbb{R}^n$ and the norm used is the ℓ_2-norm.

Definition 8 (Dual Lattice)**.** *For any lattice $\mathcal{L} \subseteq \mathbb{R}^n$, we define the dual lattice, $\mathcal{L}^*$ as*

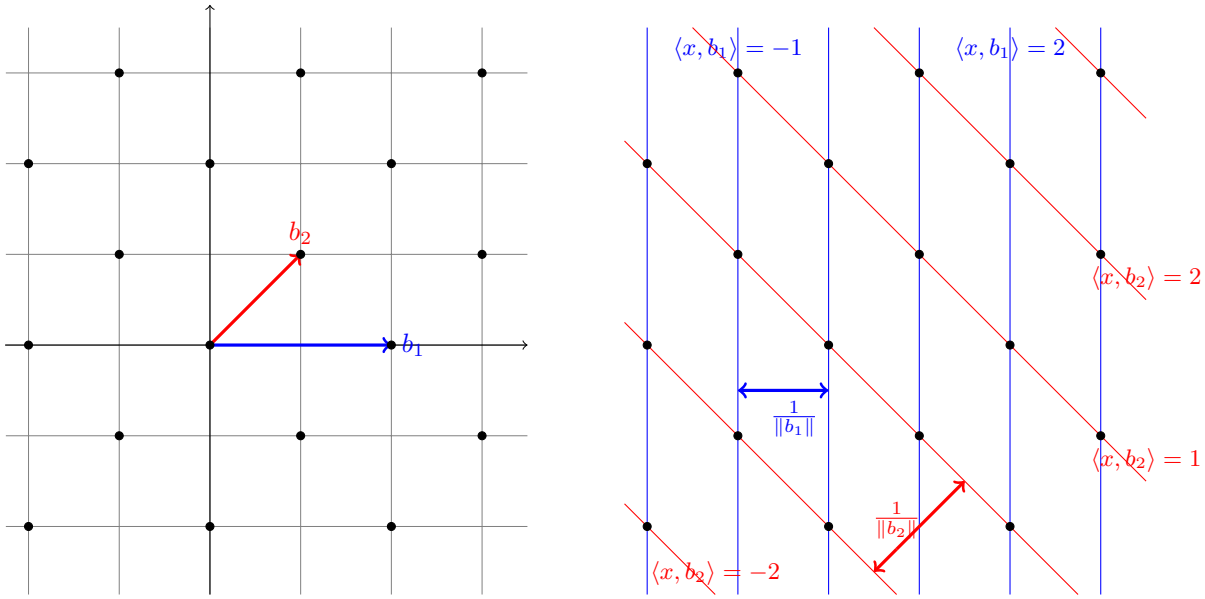
$$\mathcal{L}^* \triangleq \{x \in \mathbb{R}^n : \langle x, w \rangle \in \mathbb{Z} \ \forall w \in \mathcal{L}\}$$

Remarks:

1. The dual lattice is itself a lattice. Furthmore, the dual of the dual lattice is the lattice itself, that is $(\mathcal{L}^*)^* = \mathcal{L}$ for every lattice $\mathcal{L}$.
2. If $\mathbf{B}$ is a basis for $\mathcal{L}$, then $\mathbf{B}^* = (\mathbf{B}^{-1})^T$ is a basis for $\mathcal{L}^*$. One would expect that $(\mathbf{B}^*)^T\mathbf{B} \in \mathbb{Z}^{n \times n}$. It turns out that it is sufficient to ensure that $(\mathbf{B}^*)^T\mathbf{B} = \mathbf{I}_n$ where I_n denotes the $n \times n$ identity matrix and $\mathcal{L} \subseteq \mathbb{R}^n$.
3. For any lattice $\mathcal{L}$, $\det(\mathcal{L}^*) \cdot \det(\mathcal{L}) = 1$.
4. For any lattice $\mathcal{L}$ living in n dimensions, $\lambda_1(\mathcal{L}^*) \cdot \lambda_1(\mathcal{L}) \leq n$. This is very easy to prove using Minkowski's theorem and the previous property. In fact it is tight, that is, there are self-dual lattices $\mathcal{L}$ with $\lambda_1(\mathcal{L}) = c\sqrt{n}$ for some constant $c < 1$.
5. For any lattice $\mathcal{L}$ living in n dimensions, $1 \leq \lambda_1(\mathcal{L}^*) \cdot \lambda_n(\mathcal{L}) \leq n$. The lower bound is easy to see (we describe it ahead). The upper bound follows from a transference theorem due to Banaszczyk [Ban93]. It is also worth noting that both inequalities are tight.

We expose a little on the notion of a dual lattice. For instance, $\mathbb{Z}$ is a self-dual lattice and the dual lattice $k\mathbb{Z}$ is $(1/k)\mathbb{Z}$ for any $k \in \mathbb{Z}\backslash\{0\}$. As an example in two dimensions, consider the lattice $\mathcal{L}$ in Figure 1(a) described by the two basis vectors b_1 and b_2. The dual lattice is obtained by taking points of intersection of two sets of hyperplanes. One set of hyperplanes is the set of vectors with integer inner product with b_1 while the other is the set of vectors with integer inner product with b_2. Also note that the hyperplanes in each set are separated by a distance of $1/\|b_i\|$ for $i = 1, 2$. This gives an easy way to visualize $\mathcal{L}^*$ which suffices to prove the lower bound $1 \leq \lambda_1(\mathcal{L}^*) \cdot \lambda_n(\mathcal{L})$.

This is easy to see since the hyperplanes used for intersection live in $n-1$ dimensions while $\mathcal{L}$ lives in n dimensions. In order to obtain n linearly independent vectors, one needs to move out of a hyperplane and from the discrete structure, we see that this moves out by at least $\min_i \|b_i\|$. This means that $\lambda_1(\mathcal{L}^*) \geq 1/\lambda_n(\mathcal{L})$ which proves the above claim.



(a) A lattice $\mathcal{L}$

(b) The dual lattice $\mathcal{L}^*$

Figure 1: Visualizing the dual lattice: The images are not to scale.

Lemma 9 (Smoothing Lemma)**.** *For any lattice $\mathcal{L}$ described by a basis $\mathbf{B}$ and any parameter $s > 0$,*

$$\Delta\left(\mathcal{U}_{\mathcal{P}(\mathbf{B})}, \rho_s \bmod \mathcal{P}(\mathbf{B})\right) \leq \frac{1}{2}\rho_{1/s}(\mathcal{L}^* \backslash \{0\})$$

where

1. *$\mathcal{P}(\mathbf{B})$ denotes the fundamental parallelopiped of $\mathcal{L}$ as determined by the basis $\mathbf{B}$*
2. *$\mathcal{U}_{\mathcal{P}(\mathbf{B})}$ denotes the uniform distribution over $\mathcal{P}(\mathbf{B})$*
3. *$\rho_{1/s}(\mathcal{L}^* \backslash \{0\}) = \sum_{x \in \mathcal{L}^* \backslash \{0\}} \rho_{1/s}(x)$*

We now prove the Smoothing Lemma using Fourier Analysis.

2.1 Primer on Fourier Analysis

For any sufficiently well-behaved function[2] $f : \mathbb{R}^n \to \mathbb{C}$, one can define the Fourier transform of f to be the function $\hat{f} : \mathbb{R}^n \to \mathbb{C}$ defined by

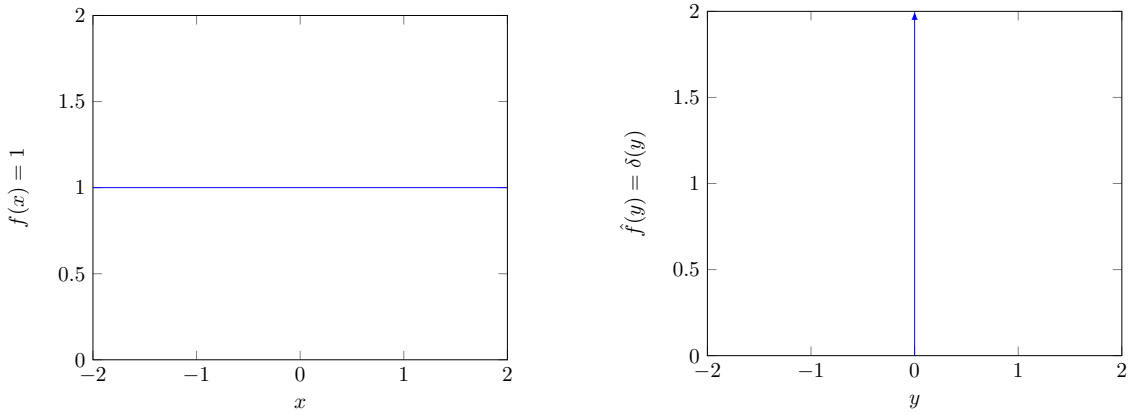
$$\hat{f}(y) = \int_{\mathbb{R}^n} f(x) e^{-2\pi i \langle x, y \rangle} \, \mathrm{d}x$$

Analogously, we can also define the inverse Fourier transform as

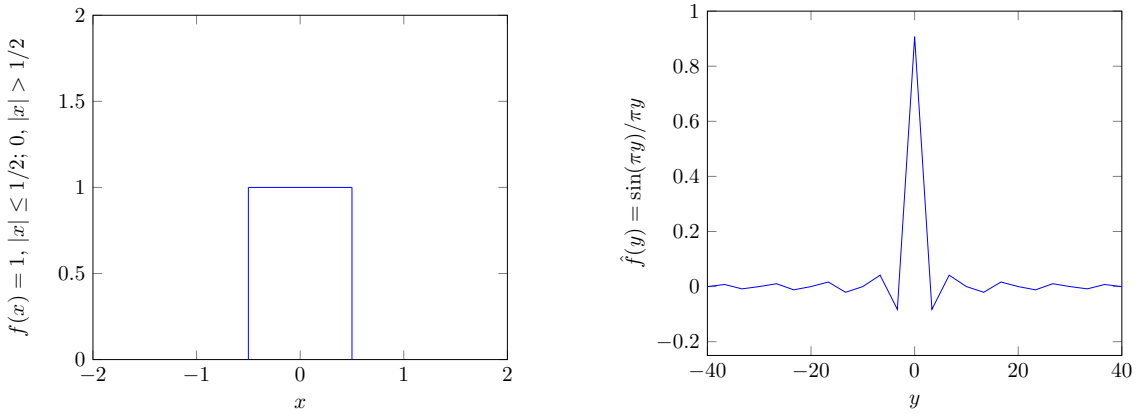
$$f(x) = \int_{\mathbb{R}^n} \hat{f}(y) e^{2\pi i \langle x, y \rangle} \, \mathrm{d}y$$

[2]For example, f could be piece-wise continuous with only finitely many discontinuities, of bounded total variation, etc. One can write less restrictive conditions but these suffice for our purposes.

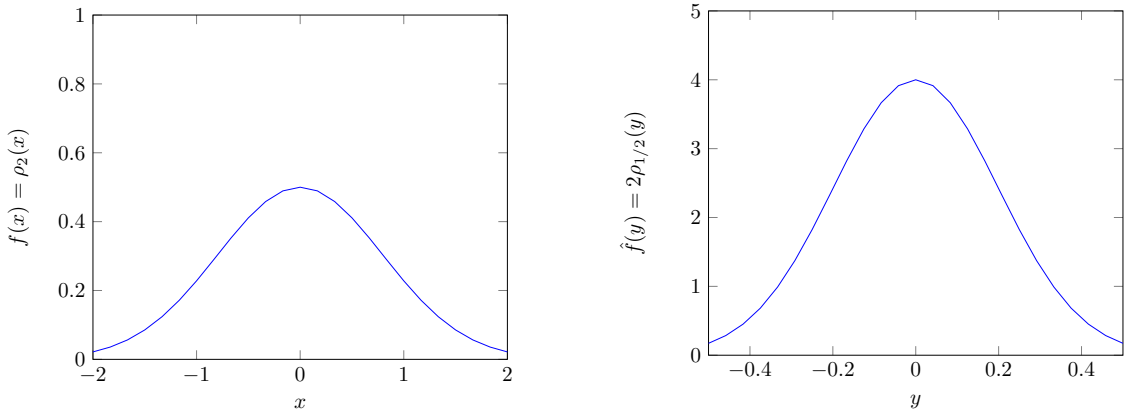
As an illustration, one can look at the functions in Figure 2 and their Fourier transforms. For instance, the constant function in Figure 2(a) has a Fourier transform of the Dirac-delta function, while the square wave in Figure 2(b) has a Fourier transform of the sinc function which is more spread out.



(a) $f(x) = 1$, $\hat{f}(y) = \delta(y)$



(b) $f(x) = 1$ if $|x| \leq 1/2$ and 0 if $|x| > 1/2$, $\hat{f}(y) = \sin(\pi y)/\pi y$



(c) $f(x) = \rho_2(x)$, $\hat{f}(y) = 2\rho_{1/2}(y)$

Figure 2: Examples of Fourier transforms

Finally, Figure 2(c) tells us what the big fuss about the Gaussian function is. In fact, one can show that for $f = \rho_s$, $\hat{(f)} = s^n \rho_{1/s}$, that is, the Fourier transform of a Gaussian function is another (scaled) Gaussian function.

We describe two more properties (both of which can be easily proved) of the Fourier transform which will enable us to prove the Smoothing Lemma.

1. A translation of the variable for a function implies a phase change of its Fourier transform. Formally, for all $a \in \mathbb{R}^n$, if $f_a(x) = f(x + a)$, then
$$\widehat{f_a}(y) = \hat{f}(y) e^{2\pi i \langle y, a \rangle}$$
2. **Poisson Summation.** For any sufficiently well-behaved function f,
$$\sum_{x \in \mathbb{Z}^n} f(x) = \sum_{y \in \mathbb{Z}^n} \hat{f}(y)$$
One could generalize this to a summation over any lattice domain $\mathcal{L}$. Then, we obtain that for any sufficiently well-behaved function f and lattice $\mathcal{L}$,
$$\sum_{x \in \mathcal{L}} f(x) = \det(\mathcal{L}^*) \sum_{y \in \mathcal{L}^*} \hat{f}(y)$$

We are now ready to prove the Smoothing Lemma (Lemma 9).

Proof. (Lemma 9) We have for all $y \in \mathcal{P}(\mathbf{B})$, $\Pr_{\mathcal{U}_{\mathcal{P}(\mathbf{B})}}[y] = 1/\det(\mathcal{L}) = \det(\mathcal{L}^*)$. Also, for all $y \in \mathcal{P}(\mathbf{B})$,

$$\Pr_{\rho_s \bmod \mathcal{P}(\mathbf{B})}[y] = \frac{1}{s^n} \sum_{z \in y + \mathcal{L}} \rho_s(z)$$

This is because all elements of $y + \mathcal{L}$ land up at y once we go mod $\mathcal{P}(\mathbf{B})$. The $1/s^n$ is a normalizing factor. We have

$$\Pr_{\rho_s \bmod \mathcal{P}(\mathbf{B})}[y] = \frac{1}{s^n} \det(\mathcal{L}^*) s^n \sum_{w \in \mathcal{L}^*} \rho_{1/s}(w) \cdot e^{2\pi i \langle y, w \rangle}$$

using the shift property and Poisson summation over the lattice. Hence, we have

$$\Pr_{\rho_s \bmod \mathcal{P}(\mathbf{B})}[y] = \det(\mathcal{L}^*) \left(1 + \sum_{w \in \mathcal{L}^* \setminus \{0\}} \rho_{1/s}(w) \cdot e^{2\pi i \langle y, w \rangle} \right)$$

This gives

$$\begin{aligned}
\Delta\left(\mathcal{U}_{\mathcal{P}(\mathbf{B})}, \rho_s \bmod \mathcal{P}(\mathbf{B})\right) &= \frac{1}{2} \int_{\mathcal{P}(\mathbf{B})} \left| \Pr_{\mathcal{U}_{\mathcal{P}(\mathbf{B})}}[y] - \Pr_{\rho_s \bmod \mathcal{P}(\mathbf{B})}[y] \right| \, \mathrm{d}y \\
&= \frac{1}{2} \int_{\mathcal{P}(\mathbf{B})} \det(\mathcal{L}^*) \sum_{w \in \mathcal{L}^* \setminus \{0\}} \rho_{1/s}(w) \cdot e^{2\pi i \langle y, w \rangle} \, \mathrm{d}y \\
&\leq \frac{1}{2} \det(\mathcal{L}^*) \det(\mathcal{L}) \max_y \left\{ \sum_{w \in \mathcal{L}^* \setminus \{0\}} \rho_{1/s}(w) \cdot e^{2\pi i \langle y, w \rangle} \right\} \\
&= \frac{1}{2} \sum_{w \in \mathcal{L}^* \setminus \{0\}} \rho_{1/s}(w) = \frac{1}{2} \rho_{1/s}(\mathcal{L}^* \setminus \{0\})
\end{aligned}$$

□

Remarks:

1. The fact that the Gaussian does not change (upto scaling) makes the above proof that much simpler. Although, it begs the question as what one can show using other distributions, probably with some hard-work.
2. A very reasonable question to ask: Can one prove the Smoothing Lemma without the machinery of Fourier Analysis? This could suggest an alternate proof of the Lemma due to Banaszczyk (Lemma 10) which is presented ahead.

3 Using the Smoothing Lemma

We show what we can do with the Smoothing Lemma (what was needed to do the worst-case to average-case reduction). We state the following lemma due to Banaszczyk without proof.

Lemma 10. *For all lattices $\mathcal{L}$ living in n dimensions with $\lambda_1(\mathcal{L}) > \sqrt{n}$,*

$$\rho_1(\mathcal{L}\backslash\{0\}) \leq 2^{-n+1}$$

Remark. We know that the Gaussian is concentrated on vectors of length $\leq \sqrt{n}$. So, if $\lambda_1(\mathcal{L}) > \sqrt{n}$, then clearly lattice vectors have a very low mass. Why the lemma is not immediately obvious is because there are still many many points in $\mathcal{L}\backslash\{0\}$. But the numbers still work out, roughly because there aren't many lattice vectors of the same length to add up and produce a high enough mass.

Theorem 11. *For all lattices $\mathcal{L}$ living in n dimensions with basis $\mathbf{B}$, if $s > \sqrt{n}\lambda_n(\mathcal{L})$,*

$$\Delta\left(\mathcal{U}_{\mathcal{P}(\mathbf{B})}, \rho_s \bmod \mathcal{P}(\mathbf{B})\right) \leq 2^{-n}$$

Proof. We have

$$s > \sqrt{n}\lambda_n(\mathcal{L}) > \frac{\sqrt{n}}{\lambda_1(\mathcal{L}^*)} \Rightarrow s\lambda_1(\mathcal{L}^*) > \sqrt{n}$$

Consider the lattice $\mathcal{L}^\dagger = s\mathcal{L}^*$. Clearly, $\lambda_1(\mathcal{L}^\dagger) = s\lambda_1(\mathcal{L}^*) > \sqrt{n}$. From Lemma 10,

$$\rho_1(\mathcal{L}^\dagger\backslash\{0\}) \leq 2^{-n+1}$$

But we know that

$$\rho_1(\mathcal{L}^\dagger\backslash\{0\}) = \rho_1(s\mathcal{L}^*\backslash\{0\}) = \rho_{1/s}(\mathcal{L}^*\backslash\{0\})$$

Hence

$$\rho_{1/s}(\mathcal{L}^*\backslash\{0\}) \leq 2^{-n+1}$$

Now, from the Smoothing Lemma (Lemma 9), we have

$$\Delta\left(\mathcal{U}_{\mathcal{P}(\mathbf{B})}, \rho_s \bmod \mathcal{P}(\mathbf{B})\right) \leq 2^{-n}$$

□

We note that Theorem 11 is exactly what defines the smoothing parameter used in the worst-case to average-case reduction in the previous lecture.

References

[Ajt96] Miklós Ajtai. Generating hard instances of lattice problems (extended abstract). In *Proceedings of the Twenty-Eighth Annual ACM Symposium on the Theory of Computing, Philadelphia, Pennsylvania, USA, May 22-24, 1996*, pages 99–108, 1996.

[Ban93] Wojciech Banaszczyk. New bounds in some transference theorems in the geometry of numbers. In *Mathematische Annalenn, 296(4)*, pages 625–635, 1993.