Worst-case to Average-case Reductions

In this lecture, we will see a few basic theorems on lattices and a property called smoothing. We will then use smoothing as a tool to come up with worst-case to average-case reductions for SIS and LWE. In a nutshell, the worst-case to average-case reductions show how to transform any algorithm that solves SIS/LWE *on the average* into an algorithm that solves "approximate short vector problems" on lattices *in the worst case.*

1 Lattice Smoothing

1.1 Lattice Duality

For a rank-n lattice $\mathcal{L}$, its dual denoted $\mathcal{L}^*$ is defined as

$$\mathcal{L}^* = \{\mathbf{x} \in \mathbb{R}^n : \forall \mathbf{y} \in \mathcal{L}, \langle \mathbf{x}, \mathbf{y} \rangle \in \mathbb{Z}\}$$

Indeed, each dual lattice vector $\mathbf{x}$ corresponds to a linear function $\phi_{\mathbf{x}} : \mathcal{L} \to \mathbb{Z}$ and the dual lattice corresponds to a basis of the space of such linear functions.

Let us start with examples and some properties.

- In one dimension, the only possible lattices are $k\mathbb{Z}$. Its dual is $(1/k) \cdot \mathbb{Z}$.
- The dual of $\mathbb{Z}^n$ is $\mathbb{Z}^n$ itself.
- If $\mathcal{L} = \mathcal{L}(\mathbf{B})$ for a basis matrix $\mathbf{B} \in \mathbb{R}^{n \times n}$, then $\mathcal{L}^*$ is generated by the columns of $\mathbf{B}^{-T}$, its transposed inverse. Indeed, the pairwise inner products of the basis vectors and the dual basis vectors is captured in the matrix $\mathbf{B} \cdot (\mathbf{B}^{-T})^T = \mathbf{I}$.

The determinant of the dual lattice is immediately seen to be the inverse of the determinant of the lattice. In an intuitive sense, as the lattice gets sparser (the determinant gets larger), the dual lattice gets denser (its determinant gets smaller). This leads us to the following lemma.

Lemma 1. *For any rank-n lattice $\mathcal{L}$, $\lambda_1(\mathcal{L}^*) \cdot \lambda_1(\mathcal{L}) \leq n$.*

Proof. We know from Minkowski that

$$\lambda_1(\mathcal{L}) \leq \sqrt{n}(\det(\mathcal{L}))^{1/n} \text{ and } \lambda_1(\mathcal{L}^*) \leq \sqrt{n}(\det(\mathcal{L}^*))^{1/n}$$

Multiplying the two, we get

$$\lambda_1(\mathcal{L}) \cdot \lambda_1(\mathcal{L}^*) \leq n$$

as desired. □

In fact, using far more advanced tools, we can show something stronger, namely that $\lambda_1(\mathcal{L}^*) \cdot \lambda_n(\mathcal{L}) \leq n$. The following lemma goes in the other direction, has an elementary proof, and we will find it useful later on.

Lemma 2. *For any rank-n lattice $\mathcal{L}$, $\lambda_n(\mathcal{L}^*) \cdot \lambda_1(\mathcal{L}) \geq 1$.*

Proof. Let $\mathbf{x} \in \mathcal{L}$ be the shortest non-zero vector. Let $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathcal{L}^*$ be linearly independent. At least one of the $\mathbf{v}_i$ has a non-zero inner product with $\mathbf{x}$, say $\langle \mathbf{v}_i, \mathbf{x} \rangle > 0$. Since the inner products of lattice vectors and dual vectors are integers, $\langle \mathbf{v}_i, \mathbf{x} \rangle \geq 1$. Therefore,

$$\lambda_1(\mathcal{L}) = \|\mathbf{x}\| \geq 1/\|\mathbf{v}_i\| \geq 1/\lambda_n(\mathcal{L}^*) .$$

□

1.2 Gaussians

The Gaussian function over $\mathbb{R}$ with (zero mean and) parameter s is defined as

$$\rho_s(x) = e^{-\pi x^2/s^2}$$

We note that

$$\int_{-\infty}^{\infty} \rho_s(x)dx = \int_{-\infty}^{\infty} e^{-\pi x^2/s^2} dx = \frac{s}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-z^2} dz = s$$

where the second equality is by a change of variables and the third by using value of the Gaussian integral ($= \sqrt{\pi}$). This fact can be used to turn the Gaussian function into a probability distribution over the reals by scaling ρ_s by $1/s$.

Something very similar can be done in n dimensions. That is, the n-dimensional Gaussian function over $\mathbb{R}^n$ is defined as

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

This can again be turned into a probability distribution after scaling by $1/s^n$.

1.3 Basic Fourier Analysis

We call a function $f : \mathbb{R}^n \to \mathbb{C}$ "nice" if it is absolutely integrable, that is, $\int_{\mathbb{R}^n} |f(\mathbf{x})| d\mathbf{x} < \infty$.

Definition 3 (Fourier Transform). *For a nice function $f : \mathbb{R}^n \to \mathbb{C}$, we define its Fourier transform $\hat{f} : \mathbb{R}^n \to \mathbb{C}$ as*

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

If $f, \hat{f}$ are nice and f is continuous, we can recover a function from its Fourier transform using the inverse formula:

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

Lemma 4 (Fourier Transform of the Gaussian function). *Let $\hat{\rho}_s$ denote the Fourier transform of the Gaussian function ρ_s. Then,*

$$\hat{\rho}_s(\mathbf{x}) = s^n \cdot \rho_{1/s}(\mathbf{x})$$

Proof. We provide a proof in one dimension.

$$\begin{aligned}
\hat{\rho}_s(\mathbf{y}) &= \int_{\mathbb{R}^n} \rho_s(\mathbf{x}) e^{-2\pi i \langle \mathbf{x}, \mathbf{y} \rangle} d\mathbf{x} \\
&= \int_{\mathbb{R}^n} e^{-\pi||\mathbf{x}||^2/s^2} e^{-2\pi i \langle \mathbf{x}, \mathbf{y} \rangle} d\mathbf{x} \\
&= e^{-\pi s^2 ||\mathbf{y}||^2} \int_{\mathbb{R}^n} e^{-\pi||(\mathbf{x}/s + is\mathbf{y})||^2} d\mathbf{x}
\end{aligned}$$

The latter integral, on a complex change of variables becomes $s^n \cdot \int_{\mathbb{R}^n} e^{-\pi||\mathbf{z}||^2} d\mathbf{z}$ which is simply s^n. So,

$$\hat{\rho}_s(\mathbf{y}) = s^n e^{-\pi s^2 ||\mathbf{y}||^2} = s^n \cdot \rho_{1/s}(\mathbf{y})$$

□

For periodic functions, we have the closely related notion of Fourier series.

Definition 5 (Fourier Series)**.** *We will define Fourier series for* periodic *functions. For a "nice enough" function $f : \mathbb{R}^n \to \mathbb{C}$ that is $\mathcal{L}$-periodic, that is, $f(\mathbf{x}+\mathbf{y}) = f(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in \mathcal{L}$, we have its Fourier series $\hat{f} : \mathcal{L}^* \to \mathbb{C}$ defined as*

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

We will state the Fourier inversion formula below without proof.

Lemma 6 (Fourier Inversion)**.** $f(\mathbf{x}) = \sum_{\mathbf{y}\in\mathcal{L}^*} \hat{f}(\mathbf{y}) e^{2\pi i \langle \mathbf{x},\mathbf{y}\rangle}$.

An important fact that connects a function f and its Fourier transform is the Poisson Summation formula. The proof of this formula goes via the Fourier series.

Lemma 7 (Poisson Summation)**.** *Given $f : \mathbb{R}^n \to \mathbb{C}$, and any full-rank lattice $\mathcal{L}$, we have*

$$\sum_{\mathbf{x}\in\mathcal{L}} f(\mathbf{x}) = \frac{1}{\det(\mathcal{L})} \cdot \sum_{\mathbf{y}\in\mathcal{L}^*} \hat{f}(\mathbf{y}) = \det(\mathcal{L}^*) \cdot \sum_{\mathbf{y}\in\mathcal{L}^*} \hat{f}(\mathbf{y})$$

Proof. Although f is not periodic, the proof of Poisson summation goes through the Fourier series of a "periodized" f. In particular, consider the function

$$\phi(\mathbf{x}) = \sum_{\mathbf{z}\in\mathcal{L}} f(\mathbf{x}+\mathbf{z})$$

Clearly ϕ is periodic over $\mathcal{L}$, therefore $\hat{\phi}$ is defined over $\mathcal{L}^*$. For any $\mathbf{y} \in \mathcal{L}^*$, we have

$$\begin{aligned}
\hat{\phi}(\mathbf{y}) &= \det(\mathcal{L}^*) \int_{\mathbf{x}\in\mathcal{P}(\mathcal{L})} \phi(\mathbf{x}) e^{-2\pi i \langle \mathbf{x},\mathbf{y}\rangle} d\mathbf{x} \\
&= \det(\mathcal{L}^*) \int_{\mathbf{x}\in\mathcal{P}(\mathcal{L})} \left(\sum_{\mathbf{z}\in\mathcal{L}} f(\mathbf{x}+\mathbf{z}) \right) e^{-2\pi i \langle \mathbf{x},\mathbf{y}\rangle} d\mathbf{x} \\
&= \det(\mathcal{L}^*) \sum_{\mathbf{z}\in\mathcal{L}} \int_{\mathbf{x}\in\mathcal{P}(\mathcal{L})} f(\mathbf{x}+\mathbf{z}) e^{-2\pi i \langle \mathbf{x},\mathbf{y}\rangle} d\mathbf{x} \\
&= \det(\mathcal{L}^*) \sum_{\mathbf{z}\in\mathcal{L}} \int_{\mathbf{x}\in\mathcal{P}(\mathcal{L})} f(\mathbf{x}+\mathbf{z}) e^{-2\pi i \langle \mathbf{x}+\mathbf{z},\mathbf{y}\rangle} d\mathbf{x} \\
&= \det(\mathcal{L}^*) \int_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}) e^{-2\pi i \langle \mathbf{x},\mathbf{y}\rangle} d\mathbf{x} \\
&= \det(\mathcal{L}^*) \hat{f}(\mathbf{y})
\end{aligned}$$

where the first equality used the definition of the Fourier series for ϕ, the second used the definition of ϕ, the third used the "niceness" of f to switch the integral and summation, the fourth used the fact that $\langle \mathbf{y},\mathbf{z}\rangle \in \mathbb{Z}$, and the final one used the definition of the Fourier transform of f.

Now use Fourier inversion for ϕ to show that

$$\sum_{\mathbf{x}\in\mathcal{L}} f(\mathbf{x}) = \phi(\mathbf{0}) = \sum_{\mathbf{y}\in\mathcal{L}^*} \hat{\phi}(\mathbf{y}) = \det(\mathcal{L}^*) \sum_{\mathbf{y}\in\mathcal{L}^*} \hat{f}(\mathbf{y})$$

□

1.4 Smoothing Lemma and Proof

Let ϕ_s denote the distribution obtained by picking a vector from the (continuous) Gaussian distribution defined by ρ_s and reducing it modulo the parallelepiped $\mathcal{P}(\mathbf{B})$. Thus,

$$\phi_s(\mathbf{x}) = 1/s^n \cdot \sum_{\mathbf{y} \in \mathcal{L}(\mathbf{B})} \rho_s(\mathbf{x} + \mathbf{y}) := 1/s^n \cdot \rho_s(\mathbf{x} + \mathcal{L}(\mathbf{B}))$$

Now, since ϕ_s is clearly a periodic function over the lattice $\mathcal{L}(\mathbf{B})$, we can compute it alternatively using the Poisson summation formula. For any $\mathbf{x} \in \mathcal{P}(\mathbf{B})$, we have

$$\begin{aligned}
\phi_s(\mathbf{x}) &= \sum_{\mathbf{y} \in \mathcal{L}^*} \hat{\phi}_s(\mathbf{y}) e^{2\pi i \langle \mathbf{x}, \mathbf{y} \rangle} \\
&= \det(\mathcal{L}^*) \cdot (1/s^n) \sum_{\mathbf{y} \in \mathcal{L}^*} \hat{\rho}_s(\mathbf{y}) e^{2\pi i \langle \mathbf{x}, \mathbf{y} \rangle} \\
&= s^n \cdot \det(\mathcal{L}^*) \cdot (1/s^n) \sum_{\mathbf{y} \in \mathcal{L}^*} \rho_{1/s}(\mathbf{y}) e^{2\pi i \langle \mathbf{x}, \mathbf{y} \rangle} \\
&= \det(\mathcal{L}^*) \cdot \left(1 + \sum_{\mathbf{y} \in \mathcal{L}^* \setminus \{\mathbf{0}\}} \rho_{1/s}(\mathbf{y}) \cdot e^{2\pi i \langle \mathbf{y}, \mathbf{x} \rangle} \right)
\end{aligned}$$

where the first equality is by the definition of Fourier inversion, the second by the definition of ϕ_s and by the linearity of the Fourier transform, the third by the Fourier transform of the Gaussian function (Lemma 4), and the final one just by grouping terms together.

We will use this formulation to compute the statistical distance of ϕ_s from the uniform distribution over the paralellepiped whose density function is $U_{\mathcal{P}(\mathbf{B})}(\mathbf{x}) = 1/\det(\mathcal{L}) = \det(\mathcal{L}^*)$.

$$\begin{aligned}
\Delta(\phi_s, U_{\mathcal{P}(\mathbf{B})}) &= \int_{\mathcal{P}(\mathbf{B})} |\phi_s(\mathbf{x}) - U_{\mathcal{P}(\mathbf{B})}(\mathbf{x})| \, d\mathbf{x} \\
&= \det(\mathcal{L}^*) \int_{\mathcal{P}(\mathbf{B})} \left| \sum_{\mathbf{y} \in \mathcal{L}^* \setminus \{\mathbf{0}\}} \rho_{1/s}(\mathbf{y}) \cdot e^{2\pi i \langle \mathbf{y}, \mathbf{x} \rangle} \right| d\mathbf{x} \\
&= \det(\mathcal{L}^*) \cdot \det(\mathcal{L}) \cdot \max_{\mathbf{x} \in \mathcal{P}(\mathbf{B})} \left| \sum_{\mathbf{y} \in \mathcal{L}^* \setminus \{\mathbf{0}\}} \rho_{1/s}(\mathbf{y}) \cdot e^{2\pi i \langle \mathbf{y}, \mathbf{x} \rangle} \right| \\
&\le \sum_{\mathbf{y} \in \mathcal{L}^* \setminus \{\mathbf{0}\}} \rho_{1/s}(\mathbf{y}) := \rho_{1/s}(\mathcal{L}^* \setminus \{\mathbf{0}\}) \qquad (1)
\end{aligned}$$

In other words, we established $\rho_{1/s}(\mathcal{L}^* \setminus \{\mathbf{0}\})$ as the quantity that governs the variation (or statistical) distance between the continuous Gaussian reduced modulo $\mathcal{P}(\mathbf{B})$ and the uniform distribution over $\mathcal{P}(\mathbf{B})$. We will now bound this quantity.

Bounding the Gaussian Weight of Non-Zero (Dual) Lattice Vectors. Let us first try to build some intuition for why we should expect to bound the Gaussian weight $\rho_{1/s}(\mathcal{L}^*)$ by something close to 1. First of all, the heaviest vector is the zero vector that gets a weight of 1. Secondly, if $\lambda_1(\mathcal{L}^*) \gtrapprox (1/s\sqrt{2\pi}) \cdot \omega(\sqrt{\log n})$, then the next heaviest vector has weight $e^{-\omega(\log n)}$ which is negligible in n. However, there could be exponentially many vectors of that length which could make the collective contribution much larger. We have to balance these two effects: the fact that a large λ_1 results in the Gaussian weight of each individual non-zero lattice vector to be tiny, versus the fact that there may be exponentially many lattice vectors of a given length.

First, let us come up with a simple upper bound on the number of lattice vectors of a given length using a packing argument.

Lemma 8. *Let $\mathcal{L}$ be a rank-n lattice. The number of lattice vectors of length at most r is at most* $\left(1+\frac{2r}{\lambda_1(\mathcal{L})}\right)^n$.

Proof. Draw balls of radius $\lambda_1/2$ around each lattice point. These balls do not intersect. As long as the length of each such lattice point is at most r, these balls are all contained in the ball of radius $r + \lambda_1/2$ around the origin. By a volume argument, we have

$$\mathsf{vol}_n(r + \frac{\lambda_1}{2}) \geq N_r \cdot \mathsf{vol}_n(\frac{\lambda_1}{2})$$

where N_r is the number of lattice vectors of length at most r. Put together, we get

$$N_r \leq \frac{\mathsf{vol}_n(r + \frac{\lambda_1}{2})}{\mathsf{vol}_n(\frac{\lambda_1}{2})} = \left(\frac{r + \frac{\lambda_1}{2}}{\frac{\lambda_1}{2}}\right)^n = \left(1 + \frac{2r}{\lambda_1(\mathcal{L})}\right)^n$$

□

We now use this to bound the sum $\sum_{\mathbf{y}\in\mathcal{L}} \rho_s(\mathbf{y})$. The proof is due to Noah Stephens-Davidowitz.

Lemma 9. *Let $\mathcal{L}$ be a rank-n lattice.* $\sum_{\mathbf{y}\in\mathcal{L}} \rho_s(\mathbf{y}) = 1 + 2^{-O(n)}$ *as long as* $\lambda_1 > Cs \cdot \sqrt{n/2\pi e}$ *for some absolute constant* $C \approx 3$.

Proof. Using a "Lebesgue integral trick" (mentioned in class), we have

$$\begin{aligned}\sum_{\mathbf{y}\in\mathcal{L}} \rho_s(\mathbf{y}) &= \int_0^1 \left|\{\mathbf{y} \in \mathcal{L} : \rho_s(\mathbf{y}) \geq t\}\right| dt \\ &= \int_0^1 \left|\{\mathbf{y} \in \mathcal{L} : e^{-\pi\|\mathbf{y}\|^2/s^2} \geq t\}\right| dt\end{aligned}$$

Now, we do a change of variables $t = e^{-\pi r^2/s^2}$, we get:

$$\begin{aligned}&= \frac{2\pi}{s^2} \int_0^\infty \left|\{\mathbf{y} \in \mathcal{L} : \|\mathbf{y}\| \leq r\}\right| r e^{-\pi r^2/s^2} dr \\ &\leq \frac{2\pi}{s^2} \cdot (\int_0^{\lambda_1} + \int_{\lambda_1}^\infty) \left|\{\mathbf{y} \in \mathcal{L} : \|\mathbf{y}\| \leq r\}\right| r e^{-\pi r^2/s^2} dr \\ &\leq (1 - e^{-\pi\lambda_1^2/s^2}) + \frac{2\pi}{s^2} \int_{\lambda_1}^\infty \left|\{\mathbf{y} \in \mathcal{L} : \|\mathbf{y}\| \leq r\}\right| r e^{-\pi r^2/s^2} dr \\ &\leq 1 + \frac{2\pi}{s^2} \int_{\lambda_1}^\infty \left(\frac{3r}{\lambda_1}\right)^n r e^{-\pi r^2/s^2} dr \\ &\leq 1 + \frac{2\pi C^n}{s^2\lambda_1^n} \int_{\lambda_1}^\infty r^{n+1} e^{-\pi r^2/s^2} dr\end{aligned}$$

where $C = 3$. After another change of variables ($w = \pi r^2/s^2$), we can bound this by

$$1 + \left(\frac{sC}{\lambda_1\sqrt{\pi}}\right)^n \Gamma(n/2)$$

where $\Gamma(\cdot)$ is the gamma function. Applying the bound on gamma functions, we get

$$1 + \left(\frac{sC}{\lambda_1}\sqrt{\frac{n}{2\pi e}}\right)^n$$

As long as $\lambda_1 > Cs \cdot \sqrt{n/2\pi e}$, we get a sum that is exponentially close to 1. □

Finally, applying this to our scenario, where the lattice is $\mathcal{L}^*$ and the function is $\rho_{1/s}$, we get that the sum $\sum_{\mathbf{y}\in\mathcal{L}^*} \rho_{1/s}(\mathbf{y})$ is exponentially close to 1 as long as

$$s \geq \lambda_n(\mathcal{L}) \cdot C \cdot \sqrt{\frac{n}{2\pi e}}$$

Indeed, if s is so large, we have $\lambda_1(\mathcal{L}^*) \geq 1/\lambda_n(\mathcal{L}) \geq \frac{C}{s} \cdot \sqrt{\frac{n}{2\pi e}}$ where the first inequality is by Lemma 2.

2 Worst-case to Average-case Reduction for SIS

The reduction is due to Ajtai originally, but our presentation follows the work of Micciancio and Regev, and borrows from Regev's lecture notes.

We first illustrate the intuition behind the worst-case to average-case reduction by showing how to reduce the approximate-SIVP problem to a variant of SIS over the torus $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, $\mathsf{SIS}_{\mathbb{T}}$. $\mathsf{SIS}_{\mathbb{T}}$ is exactly as in SIS, except that you are given a matrix $\mathbf{A} \in \mathbb{T}^{n\times m}$ and you are asked to find a small integer linear combination that sums to zero. That is, find $\mathbf{x} \in \mathbb{Z}^m$ such that $\mathbf{Ax} = 0$ and $||\mathbf{x}||$ is "small".

How would such a reduction look like? On the one hand, the reduction has to generate a *uniformly random* $\mathsf{SIS}_{\mathbb{T}}$ instance from a given lattice $\mathcal{L}$; therefore, the SIS instance "forgets" the lattice $\mathcal{L}$ that was used to generate it. On the other hand, a solution to the SIS instance has to somehow be mapped back to a non-trivially short vector in $\mathcal{L}$. This (apparent) conundrum is common to all worst-case to average-case reductions, and the answer is that the reduction knows some information connecting the lattice to the SIS instance which, together with the SIS solution, helps it generate short vectors in $\mathcal{L}$.

The reduction first generates a random vector $\mathbf{v} \in \mathcal{P}(\mathbf{B})$ in the parallelpiped associated to the given basis. It does so by sampling a vector $\mathbf{x} \leftarrow \rho_s$ from the (zero-centered) Gaussian with standard deviation parameter $s \geq \eta_\varepsilon(\mathcal{L})$, the smoothing parameter for some negligible function $\varepsilon = \varepsilon(n)$, and setting

$$\mathbf{v} = \mathbf{x} \pmod{\mathcal{P}(\mathbf{B})}$$

By the smoothing lemma, $\mathbf{v}$ is (close to) random over the paraellelepiped. The first column of the SIS matrix $\mathbf{A}$ is then set to

$$\mathbf{a} = \mathbf{B}^{-1}\mathbf{v} \in \mathbb{T}^n$$

which is (close to) random over $[0, 1)^n$. Repeat this process independently m times to generate the statistically close to uniform $\mathsf{SIS}_{\mathbb{T}}$ matrix $\mathbf{A} \in \mathbb{T}^{n\times m}$ where

$$\mathbf{A} = \mathbf{B}^{-1}\mathbf{V}$$

Call the Gaussian matrix corresponding to $\mathbf{V}$ as $\mathbf{X}$. The reduction will keep $\mathbf{X}$ to itself.

Assume now that there is a $\mathsf{SIS}_{\mathbb{T}}$ algorithm that gives us a non-zero integer vector $\mathbf{x} \in \mathbb{Z}^m$ such that $\mathbf{Ax} \in \mathbb{Z}^n$. (this is what it means for $\mathbf{Ax}$ to be $\mathbf{0} \pmod 1$.) Then we know that $\mathbf{B}^{-1}\mathbf{Vx} \in \mathbb{Z}^n$ and therefore, $\mathbf{Vx} \in \mathcal{L}(\mathbf{B})$ is a lattice vector. Now, since $\mathbf{X} \equiv \mathbf{V} \pmod{\mathcal{P}(\mathbf{B})}$, we know that $\mathbf{Xx} \in \mathcal{L}(\mathbf{B})$ is also a lattice vector.

We now argue that it is short. We know that $||\mathbf{Xx}|| \approx s||\mathbf{x}||\sqrt{n} \approx \lambda_n\sqrt{mn}$. Here, the first equality is because each column of $\mathbf{X}$ is a continuous Gaussian with parameter s and therefore $\mathbf{Xx}$ has parameter $s||\mathbf{x}||$ and therefore length $s||\mathbf{x}||\sqrt{n}$ w.h.p. The second equality is using the smoothing lemma, substituting λ_n for s upto logarithmic factors and $\sqrt{m}$ as the norm of $\mathbf{x}$, assuming it is a 0-1 vector.

This seems to work, except that we are uncomfortable working with real numbers. Furthermore, it is unclear that a "random" matrix $\mathbf{A} \in \mathbb{T}^{n\times m}$ will have an SIS solution at all. We therefore discretize.

Discretization. Consider splitting each entry into a multiple of $1/q$ (for some sufficiently large value of q that we will set shortly) and an error term. That is,

$$\mathbf{A} = \mathbf{Q} + \mathbf{E} \pmod 1$$

where $q\mathbf{Q} \in \mathbb{Z}^{n\times m}$ and $||\mathbf{E}||_\infty \le 1/2q$.

Our first try is to feed the SIS algorithm with the matrix $q\mathbf{Q}$ which is uniformly random mod q. The adversary returns an $\mathbf{x}$ such that $q\mathbf{Qx} = 0 \pmod q$. This gives us

$$\mathbf{0} = \mathbf{Qx} = (\mathbf{A} - \mathbf{E})\mathbf{x} = \mathbf{B}^{-1}(\mathbf{V} - \mathbf{BE})\mathbf{x} = \mathbf{B}^{-1}(\mathbf{X} - \mathbf{BE})\mathbf{x} \pmod 1$$

and therefore, $(\mathbf{X} - \mathbf{BE})\mathbf{x}$ is a lattice vector. We would, in analogy to before, show that these are short lattice vectors.

$$||\mathbf{Xx} - \mathbf{BEx}|| \le ||\mathbf{Xx}|| + ||\mathbf{BEx}|| \le s||\mathbf{x}||\sqrt{n} + \frac{||\mathbf{x}||_1}{q} \cdot \max_i ||\mathbf{b}_i||_2$$

So, this does not give us short vectors, rather it reduces the length of the longest vector in the basis by a factor of $q/||\mathbf{x}||_1 \ge q/m$ (roughly, assuming SIS produces 0-1 vectors). So, as long as $q \gg m \approx n\log q$, we get an improvement. Repeat this iteratively many times to get to roughly $s\sqrt{mn} \approx \lambda_n\sqrt{mn} \approx \lambda_n \cdot \tilde{O}(n)$.

We are stuck at solving n-approximate SIVP given a solver for SIS. Can we improve this?

Open Problem 3.1. Show a reduction from $\sqrt{n}$-SIVP (or better) to average-case SIS.

In the regime of exponential reductions, we show such reductions in a recent joint work with Brakerski and Stephens-Davidowitz.

Another question is to improve the values of q for which one can show SIS average-case hard.

Open Problem 3.2. Show a reduction from approximate SIVP to SIS with modulus $q = O(1)$.

Why do we get a non-zero vector, again? There is one important issue that we overlooked. We showed that the reduction produces a short lattice vector, but why is the vector non-zero? Relatedly, when the reduction produces many shorter vectors that form a new basis to iterate on, why do we have the guarantee that we get n linearly independent vectors from the reduction?

We will now show non-zero-ness formally, but here is the intuition: we need to think of the SIS algorithm as the adversary who is trying to send us a vector $\mathbf{x}$ which is somehow cleverly designed so that $(\mathbf{X}-\mathbf{BE})\mathbf{x}$ is the zero vector. What does the SIS algorithm see? It possibly sees $\mathbf{V}=\mathbf{X} \pmod{\mathcal{P}(\mathbf{B})}$ but (a) it never sees $\mathbf{X}$ itself; and (b) given $\mathbf{V}$, there are multiple possible values of $\mathbf{X}$, which is a consequence of smoothing-type arguments. In other words, the adversary is trying to force $(\mathbf{X}-\mathbf{BE})\mathbf{x}$ to be $\mathbf{0}$, but it does not know what $\mathbf{X}$ is. We then argue that information-theoretically, it cannot succeed.

We omit the formal argument, but refer the reader to Regev's lecture notes for the full proof.

2.1 Other Open Problems

Vinod finds it rather bothersome that Ajtai's reduction (and essentially every other known reduction) that demonstrates average-case hardness of SIS starts from the *SIVP* problem, rather than the more natural *SVP*. This motivates the following open problem.

Open Problem 3.3**.** Show a reduction from worst-case SVP to (average-case) SIS.

In fact, he would ideally like a reduction from worst-case SIS to average-case SIS, bypassing lattices altogether. Indeed, observe that solving SIS is the same as finding a short vector in a lattice (namely, the lattice $\Lambda^{\perp}(\mathbf{A}) := \{\mathbf{x}\in\mathbb{Z}^m : \mathbf{A}\mathbf{x}=0 \pmod q\}$). However, even viewing through these lens, what we have demonstrated is an algorithm that finds vectors of length related to λ_n, and not λ_1. That is, if the worst-case SIS lattice has a short vector but no n linearly independent short vectors, then the reduction will miss finding the short vector (‼) We view this as a deficiency in our understanding of SIS and worst-case to average-case reductions. Therefore, a related problem is:

Open Problem 3.4**.** Show a reduction from worst-case SIS to average-case SIS without going through lattices.

3 Bounded Distance Decoding and LWE

The bounded distance decoding (BDD) problem is a promise variant of the closest vector problem (CVP) on lattices, where the target point is guaranteed to be so close to the lattice that there is a *unique* closest vector. In other words, in the c-BDD problem for a $c\in[0,1/2)$, one is given a basis $\mathbf{B}\in\mathbb{Z}^{m\times m}$ of a lattice $\mathcal{L}(\mathbf{B})$ and a target vector $\mathbf{t}\in\mathbb{Z}^m$ such that $\mathsf{dist}(t,\mathcal{L}(\mathbf{B}))\leq c\cdot\lambda_1(\mathcal{L}(\mathbf{B}))$, and the goal is to find the lattice vector that is closest to $\mathbf{t}$.

BDD and LWE are very closely related as the reader may have noticed already. In particular, LWE can be seen as an average-case version of BDD in the following way. Define the LWE lattice

$$\Lambda(\mathbf{A}) := \{\mathbf{z}\in\mathbb{Z}^m : \exists\, \mathbf{s}\in\mathbb{Z}_q^n \text{ s.t. } \mathbf{z}=\mathbf{s}^T\mathbf{A} \pmod q\}$$

(Note that $q\mathbb{Z}^m\subseteq\Lambda(\mathbf{A})\subseteq\mathbb{Z}^m$.) It is not hard to show that the minimum distance of $\Lambda(\mathbf{A})$ for a uniformly random matrix $\mathbf{A}\in\mathbb{Z}_q^{n\times m}$ is $c'q^{1-n/m}$ with high probability. (We will leave this calculation as an exercise.)

LWE is then the regime where the secret $\mathbf{s}$ (which defines the closest vector) is uniquely determined given $\mathbf{s}^T\mathbf{A}+\mathbf{e}^T$.

3.1 Discrete Gaussians Strike Again

As we saw in the last lecture, the Gaussian function

$$\rho_s(\mathbf{x}) := e^{-\pi\|\mathbf{x}\|^2/s^2}$$

from $\mathbb{R}^n$ to $\mathbb{R}$ can be turned into a probability distribution over $\mathbb{R}^n$ by normalizing with $\int_{\mathbb{R}^n} \rho_s(\mathbf{x})d\mathbf{x} = s^n$. Henceforth, we will call this the (n-dimensional) Gaussian distribution N_s. Thus,

$$N_s(\mathbf{x}) = \frac{1}{s^n} \cdot e^{-\pi\|\mathbf{x}\|^2/s^2}$$

Given a lattice $\mathcal{L}$, we will define the discrete Gaussian distribution $D_{\mathcal{L},s}$ as the probability distribution that assigns the value 0 to all $\mathbf{x} \notin \mathcal{L}$ and the values

$$D_{\mathcal{L},s}(\mathbf{x}) = \frac{\rho_s(\mathbf{x})}{\rho_s(\mathcal{L})}$$

for every $\mathbf{x} \in \mathcal{L}$. Here, $\rho_s(\mathcal{L}) := \sum_{\mathbf{v}\in\mathcal{L}} \rho_s(\mathbf{v})$.

The latter definition can be generalized to any discrete set; for example, we will let $D_{\mathcal{L}+\mathbf{c},s}$ denote the discrete Gaussian over the lattice coset $\mathcal{L}+\mathbf{c} = \{\mathbf{v}+\mathbf{c} : \mathbf{v} \in \mathcal{L}\}$ which assigns the Gaussian mass (normalized appropriately) to each vector in $\mathcal{L} + \mathbf{c}$ and 0 to all other vectors.

We will also define off-centered versions of these quantities $\rho_{s,\mathbf{c}}$, $N_{s,\mathbf{c}}$ and $D_{\mathcal{L},s,\mathbf{c}}$; for example, $\rho_{s,\mathbf{c}}(\mathbf{x}) := e^{-\pi\|\mathbf{x}-\mathbf{c}\|^2/s^2}$, and so on.

When s exceeds the smoothing parameter of the lattice $\eta_\varepsilon(\mathcal{L})$, the discrete Gaussian over $\mathcal{L}$ starts having a number of nice regularity properties that make it behave essentially as if it were a continuous Gaussian distribution. Some examples follow.

Lemma 10. *For any $\mathbf{c} \in \mathbb{R}^n$, and $s \geq \eta_\varepsilon(\mathcal{L})$,*

$$\rho_s(\mathcal{L} + \mathbf{c}) \in [1 - 2\varepsilon, 1 + 2\varepsilon] \cdot \rho_s(\mathcal{L})$$

Proof. Let $\mathbf{c}'$ denote the shortest vector in the lattice coset $\mathcal{L} + \mathbf{c}$. Then,

$$\begin{aligned}
\rho_s(\mathcal{L} + \mathbf{c}) &= \rho_{s,-\mathbf{c}}(\mathcal{L}) \\
&= \det(\mathcal{L}^*) \cdot \widehat{\rho_{s,-\mathbf{c}}}(\mathcal{L}^*) \\
&= \det(\mathcal{L}^*) \cdot \sum_{\mathbf{z}\in\mathcal{L}^*} \widehat{\rho_{s,-\mathbf{c}}}(\mathbf{z}) \\
&= \det(\mathcal{L}^*) \cdot \sum_{\mathbf{z}\in\mathcal{L}^*} e^{2\pi i\langle\mathbf{c},\mathbf{z}\rangle} \rho_{1/s}(\mathbf{z}) \\
&= \det(\mathcal{L}^*) \cdot \left(1 + \sum_{\mathbf{z}\in\mathcal{L}^*\setminus\{\mathbf{0}\}} e^{2\pi i\langle\mathbf{c},\mathbf{z}\rangle} \rho_{1/s}(\mathbf{z})\right) \\
&\in [1 - \varepsilon, 1 + \varepsilon] \cdot \det(\mathcal{L}^*)
\end{aligned}$$

The claim follows. □

A direct corollary is the following statement about discrete Gaussians modulo sublattices. It says that if you choose a vector from a discrete Gaussian over a dense (rank n) lattice $\mathcal{L}$ and reduce it modulo a sparser (also rank n) lattice $\mathcal{L}' \subseteq \mathcal{L}$, you get a uniformly random element of the finite group $\mathcal{L}/\mathcal{L}'$. This will be instantiated later in the lecture where $\mathcal{L}$ will be an arbitrary lattice and $\mathcal{L}' = q\mathcal{L}$ will be a scaling of it. Here, $\mathcal{L}/\mathcal{L}' \cong \mathbb{Z}_q^n$.

Lemma 11 (Discrete+Continuous Convolution)**.** *Let $\mathcal{L}$ be a lattice. Consider the distribution obtained by sampling a vector* $\mathbf{v}$ *from the discrete Gaussian $D_{\mathcal{L},s}$ and a vector* $\mathbf{w}$ *from the continuous Gaussian N_r and adding them together, where $s, r \geq \eta_\varepsilon(\mathcal{L}) \cdot \sqrt{2}$ (where ε is a negligible function of n). Then, the resulting distribution is statistically close to the continuous Gaussian $N_{\sqrt{r^2+s^2}}$.*

Proof. Consider the distribution Y obtained by adding up the two vectors. Let $t = \sqrt{r^2 + s^2}$.

$$\begin{aligned}
Y(\mathbf{x}) &= \sum_{\mathbf{v}\in\mathcal{L}} \Pr_{D_{\mathcal{L},s}}[\mathbf{v}] \cdot \Pr_{N_r}[\mathbf{x}-\mathbf{v}] \\
&= \frac{1}{\rho_s(\mathcal{L}) \cdot r^n} \sum_{\mathbf{v}\in\mathcal{L}} \rho_s(\mathbf{v}) \cdot \rho_r(\mathbf{x}-\mathbf{v}) \\
&= \frac{1}{\rho_s(\mathcal{L}) \cdot r^n} \sum_{\mathbf{v}\in\mathcal{L}} e^{-\pi\|\mathbf{v}\|^2/s^2} \cdot e^{-\pi\|\mathbf{x}-\mathbf{v}\|^2/r^2} \\
&= \frac{1}{\rho_s(\mathcal{L}) \cdot r^n} \sum_{\mathbf{v}\in\mathcal{L}} e^{-\pi\left(\|\mathbf{v}\|^2\cdot(t^2/r^2s^2) - 2\langle \mathbf{x},\mathbf{v}\rangle/r^2 + \|\mathbf{x}\|^2/r^2\right)} \\
&= \frac{e^{-\pi\|\mathbf{x}\|^2\cdot\frac{1}{r^2}\cdot(1-\frac{s^2}{t^2})}}{\rho_s(\mathcal{L}) \cdot r^n} \sum_{\mathbf{v}\in\mathcal{L}} e^{-\pi\left(\|\mathbf{v}\|^2\cdot(t^2/r^2s^2) - 2\langle \mathbf{x},\mathbf{v}\rangle/r^2 + \|\mathbf{x}\|^2\cdot(s^2/t^2r^2)\right)} \\
&= \frac{e^{-\pi\|\mathbf{x}\|^2/t^2}}{\rho_s(\mathcal{L}) \cdot r^n} \sum_{\mathbf{v}\in\mathcal{L}} e^{-\pi\|\mathbf{v}-s^2/t^2\cdot\mathbf{x}\|^2/(rs/t)^2} \\
&= \frac{\rho_t(\mathbf{x})}{t^n} \cdot \frac{t^n}{r^n} \cdot \frac{\rho_{rs/t,s^2/t^2\cdot\mathbf{x}}(\mathcal{L})}{\rho_s(\mathcal{L})} \\
&\in \Big[1-\varepsilon, 1+\varepsilon\Big] \cdot \frac{\rho_t(\mathbf{x})}{t^n} \cdot \frac{t^n}{r^n} \cdot \frac{\rho_{rs/t}(\mathcal{L})}{\rho_s(\mathcal{L})}
\end{aligned}$$

where we used Lemma 10 on the numerator since $rs/t \geq \eta_\varepsilon(\mathcal{L})$.

By Proposition 12, we have $\frac{\rho_{rs/t}(\mathcal{L})}{\rho_s(\mathcal{L})} \in [1-2\varepsilon, 1+2\varepsilon] \cdot (r/t)^n$. Put together with the above, we have

$$Y(\mathbf{x}) \in \Big[1-3\varepsilon, 1+3\varepsilon\Big] \cdot N_t(\mathbf{x})$$

from which it follows that the statistical distance between the two distributions in question is at most 3ε. □

Proposition 12. *Assume that $s_1, s_2 \geq \eta_\varepsilon(\mathcal{L})$. Then,*

$$\frac{\rho_{s_1}(\mathcal{L})}{\rho_{s_2}(\mathcal{L})} \in [1-2\varepsilon, 1+2\varepsilon] \cdot \left(\frac{s_1}{s_2}\right)^n$$

Proof. We have

$$\rho_s(\mathcal{L}) = \det(\mathcal{L}^*) \cdot s^n \rho_{1/s}(\mathcal{L}^*) \in [1-\varepsilon, 1+\varepsilon] \cdot s^n \cdot \det(\mathcal{L}^*)$$

where the first equality uses Poisson summation and the fact that $\widehat{\rho_s} = s^n\rho_{1/s}$, and the second the definition of the smoothing parameter and the fact that $s \geq \eta_\varepsilon(\mathcal{L})$. Thus,

$$\frac{\rho_{s_1}(\mathcal{L})}{\rho_{s_2}(\mathcal{L})} \in [1-2\varepsilon, 1+2\varepsilon] \cdot \left(\frac{s_1}{s_2}\right)^n$$

□

3.2 Poor Person's Discrete Gaussian Sampling

For the first step of our reduction in the next section, we need an algorithm to sample from the discrete Gaussian distribution $D_{\mathcal{L},s}$ given s and some basis $\mathbf{B}$ of $\mathcal{L}$. Clearly, this is hard to do if $s < 1/\sqrt{n} \cdot \max_i \|\mathbf{b}_i\|$ as it will then give us a way to make the vectors of $\mathbf{B}$ shorter, a computationally hard problem. However, one can hope that for significantly larger s, this is possible. Indeed, Gentry, Peikert and Vaikuntanathan, following an algorithm of Klein, show such a (polynomial-time) algorithm with $s \geq \omega(\sqrt{\log n}) \cdot \max_i \|\mathbf{b}_i\|$ (in fact, something slightly stronger but it will not matter to us). Their algorithm samples from a distribution that is negligibly close (in statistical distance) to the discrete Gaussian.

Here, we will make do with something significantly weaker.

We will show a *very* simple algorithm SimpleDGS that samples from the discrete Gaussian $D_{\mathcal{L},s}$ where $s \geq 2^n \cdot \max_i \|\mathbf{b}_i\|$. The algorithm simply samples a vector $\mathbf{v} \leftarrow N_s$ from the continuous Gaussian distribution with parameter s and "rounds" it modulo the parallelepiped $\mathcal{P}(\mathbf{B})$. That is, output

$$\mathbf{v}' = \mathbf{B}\lfloor \mathbf{v} \rceil \in \mathcal{L}(\mathbf{B})$$

To show that this is statistically close to $D_{\mathcal{L},s}$, we calculate the two probabilities:

- $\Pr[\mathbf{w} \sim D_{\mathcal{L},s}] = c \cdot \rho_s(\mathbf{w})$ for some constant normalization factor c.
- $\Pr[\mathbf{w} \sim \text{SimpleDGS}] = c' \cdot \int_{\mathbf{x} \in \mathcal{P}(\mathbf{B})} \rho_s(\mathbf{w} + \mathbf{x}) d\mathbf{x}$.

The intuition is that $\rho_s(\mathbf{w} + \mathbf{x})$ is very close to $\rho_s(\mathbf{w})$ for all the *typical* vectors, that is, vectors of length at most $s\sqrt{n}$. Indeed,

$$\rho_s(\mathbf{w} + \mathbf{x}) = \rho_s(\mathbf{w}) \cdot e^{-\pi(2\langle \mathbf{w}, \mathbf{x} \rangle + \|\mathbf{x}\|^2)/s^2}$$

It suffices to show that $|2\langle \mathbf{w}, \mathbf{x} \rangle + \|\mathbf{x}\|^2)/s^2|$ is very small. Note that this quantity is at most $(2\|\mathbf{w}\|\|\mathbf{x}\| + \|\mathbf{x}\|^2)/s^2$ by Cauchy-Schwartz. Since $\|\mathbf{w}\| \approx s\sqrt{n}$ is the length of the typical vectors (Exercise: Check this!) and $s \gg 2^n \max_i \|\mathbf{b}_i\| \geq 2^n \|\mathbf{x}\|$, we are done.

A remark to a reader who might be wondering if this algorithm in fact performs better, i.e., with a smaller s, and if the large s is merely an artifact of our analysis. To show that it is not, the reader is recommended to let $\mathcal{L} = \mathbb{Z}$ and show that for small s, the rounded continuous Gaussian (our distribution) and the discrete Gaussian over $\mathbb{Z}$ are in fact statistically far.

3.3 From (Worst-case) BDD to (Average-case) LWE

We show the reduction from the worst-case bounded distance decoding problem, which we saw was morally the same as the LWE problem, to the average-case LWE problem.

We will produce LWE samples where the LWE noise are drawn from a continuous Gaussian. It is easy to discretize it and make the noise comes from the rounded continuous Gaussian distribution.

Claim 13. *The vectors $\mathbf{a}_i$ are statistically close to uniformly random in $\mathbb{Z}_q^n$ and independent.*

Proof. By inspection, we see that the probability of getting $\mathbf{a}_i$ is the probability that the discrete Gaussian $D_{\mathcal{L}^*,s}$ lands up in the set $q\mathcal{L}^* + \mathbf{B}^*\mathbf{a}_i$. This is precisely

$$\frac{\rho_s(q\mathcal{L}^* + \mathbf{c})}{\sum_{\mathbf{c}} \rho_s(q\mathcal{L}^* + \mathbf{c})} \tag{2}$$

Regev's BDD to LWE Reduction

Input: Lattice basis $\mathbf{B} \in \mathbb{Z}^{n \times n}$, $\mathbf{t} = \mathbf{B}\mathbf{s} + \mathbf{e} \in \mathbb{Z}^n$.
(*For simplicity, we will assume that* $\|\mathbf{e}\|$ *is known.*)
Output: LWE instance $\mathbf{A} \in \mathbb{Z}_q^{n \times m}$, $y \in \mathbb{Z}_q^m$.

Repeat m times:

- Let $q \geq 2^{2n}$, where $s \geq q\sqrt{2} \cdot \eta_\varepsilon(\mathcal{L}^*)$ and $r \geq \sqrt{2} \cdot \|\mathbf{x}\| \cdot \eta_\varepsilon(\mathcal{L}^*)$.
- Sample a vector $\mathbf{v}_i \leftarrow D_{\mathcal{L}^*,s}$.
- Compute

$$\mathbf{a}_i := (\mathbf{B}^*)^{-1}\mathbf{v}_i = \mathbf{B}^T\mathbf{v}_i \pmod q \text{ and } b_i := \mathbf{t}^T\mathbf{v}_i + e_i' \pmod q$$

where $e_i' \leftarrow N_r$.

Run the LWE algorithm on input $(\mathbf{A}, \mathbf{b})$ where the columns of $\mathbf{A}$ are the $\mathbf{a}_i$, and output what it outputs.

Since $s \geq q\eta_\varepsilon(\mathcal{L}^*) = \eta_\varepsilon(q\mathcal{L}^*)$, we know by Lemma 10 that

$$\sum_{\mathbf{c}} \rho_s(q\mathcal{L}^* + \mathbf{c}) \in [1 - 2\varepsilon, 1 + 2\varepsilon] \cdot \rho_s(q\mathcal{L}^*) \cdot q^n$$

and

$$\rho_s(q\mathcal{L}^* + \mathbf{c}) \in [1 - 2\varepsilon, 1 + 2\varepsilon] \cdot \rho_s(q\mathcal{L}^*)$$

therefore, the ratio in equation 2 is in the range $\frac{1}{q^n} \cdot [1 - 4\varepsilon, 1 + 4\varepsilon]$. Consequently, the statistical distance is at most 4ε. □

Claim 14. *$b_i = \mathbf{s}^T\mathbf{a}_i + e_i$ and e_i is statistically close to a (1-dimensional) continuous Gaussian N_t where $t = \|\mathbf{x}\| \cdot \sqrt{2}\eta_\varepsilon(\mathcal{L}^*)$.*

Proof. For the reduction and the proof, we will assume that $\|\mathbf{e}\|$ is known. This assumption can be removed with more care; we refer to Regev's 2005 paper for more details.

Start by noting that

$$\begin{aligned} b_i &= \mathbf{t}^T\mathbf{v}_i + e_i' \pmod q \\ &= (\mathbf{s}^T\mathbf{B}^T + \mathbf{e}^T)\mathbf{B}^*\mathbf{a}_i + e_i' \pmod q \\ &= \mathbf{s}^T\mathbf{B}^T\mathbf{B}^{-T}\mathbf{a}_i + \mathbf{e}^T\mathbf{v}_i + e_i' \pmod q \\ &= \mathbf{s}^T\mathbf{a}_i + e_i \pmod q \end{aligned}$$

where the second equality follows from the definition of $\mathbf{t} := \mathbf{B}\mathbf{s} + \mathbf{e}$ and that of $\mathbf{a}_i$, and $e_i := \mathbf{e}^T\mathbf{v}_i + e_i'$. It remains to analyze the distribution of e_i.

First, e'_i is distributed like $\mathbf{e}^T\mathbf{w}_i$ where $\mathbf{w}_i$ is a continuous Gaussian with parameter $\sqrt{2}\eta_\varepsilon(\mathcal{L}^*)$. Thus,

$$e'_i = \mathbf{e}^T(\mathbf{v} + \mathbf{w}) = \mathbf{e}^T\mathbf{w}'$$

where $\mathbf{w}'$ is distributed like $N_{s'}$ by Lemma 11 with

$$s' \approx q \cdot \|\mathbf{x}\| \cdot \eta_\varepsilon(\mathcal{L}^*) \leq cq\lambda_1(\mathcal{L})\eta_\varepsilon(\mathcal{L}^*) \in cq \cdot [1, \sqrt{n}]$$

by Banaszczyk's theorem. In the worst case, if $c \ll 1/\sqrt{n}$, this gives us an LWE distribution with meaningfully bounded error. □

In summary, the reduction solves $1/\sqrt{n}$-BDD assuming an LWE solver with a constant factor noise-to-modulus ratio.

3.4 From (Worst-case) SIVP to (Worst-case) BDD

3.4.1 A Classical Reduction

We now present a classical reduction from gapSVP to BDD due to Peikert. We contrast this with Regev's quantum reduction from SIVP to BDD.

The advantage of Peikert's reduction, of course, is that it is classical. However, it is a reduction from a decision problem (gapSVP) to a search problem (BDD), as opposed to Regev's quantum reduction that reduces from search SIVP. For classes of lattices such as ideal lattices, the gapSVP problem for small factors turns out to be easy making the (analogous) reduction vacuous, so it is important to find a reduction starting from a *search* problem. Thus, the following question is wide open.

Open Problem 4.1. Show a (worst-case) reduction from SIVP (or SVP or CVP) to BDD.

We sketch the idea behind Peikert's reduction which in turn draws inspiration from a beautiful *coAM* protocol for gapSVP due to Goldreich and Goldwasser. Let $\mathcal{L}$ be the input lattice with the promise that $\lambda_1(\mathcal{L}) \leq 1$ or $\lambda_1(\mathcal{L}) > \gamma$. Assume that we have access to a c-BDD solver, namely an algorithm that returns the closest lattice vector given the promise that the target point is within distance $c \cdot \lambda_1(\mathcal{L})$ from the lattice. The reduction works as follows.

- Pick a random lattice point $\mathbf{z} \in \mathcal{L}$ and add a random point $\mathbf{e}$ from a ball of radius $c \cdot \gamma$.
- Run the BDD solver with input $\mathbf{t} := \mathbf{z} + \mathbf{e}$.
- If the BDD solver produces a vector $\mathbf{z}' = \mathbf{z}$, output NO ("large λ_1") else output NO ("small λ_1").

On the one hand, if $\lambda_1(\mathcal{L}) > \gamma$, then the distance of $\mathbf{t}$ from the lattice is at most $c \cdot \lambda_1(\mathcal{L})$ and thus it satisfies the BDD promise. Consequently, the BDD solver will return $\mathbf{z}$. On the other hand, if $\lambda_1(\mathcal{L}) \leq 1$, the (uniform distribution on the) balls centered at $\mathbf{z}$ and $\mathbf{z} + \mathbf{u}$ where $\|\mathbf{u}\| = \lambda_1(\mathcal{L})$ are statistically close, *if* $c\gamma \geq \sqrt{n}$. Therefore, a c-BDD algorithm helps us solve $\sqrt{n}/c$-gapSVP.

Putting this together with the worst-case to average-case reduction, we get a $O(n)$-gapSVP algorithm given an LWE solver with constant noise-to-modulus ratio.