COS 598D - Lattices

scribe: Srdjan Krstic

Introduction

In the first part we will give a brief introduction to lattices and their relevance in some topics in computer science. Then we show some specific problems and results, namely the LLL algorithm and a construction of a collision-resistant hash function from lattices, in particular from worst-case hardness of the SIVP problem.

The first two sections are largely based on the first two lectures by Oded Regev, course "Lattices in Computer Science, Fall 2004. The third part is somewhat based on the paper by Oded Regev and Daniele Micciancio - *Worst-case to Average-case Reductions based on Gaussian Measures*, SIAM Journal on Computing 37(1) pp. 267-302, 2007. and to a larger extent on the paper by O. Goldreich, S. Goldwasser, and S. Halevi - *Collision-Free Hashing from Lattice Problems*, ECCC, TR96-042, 1996.

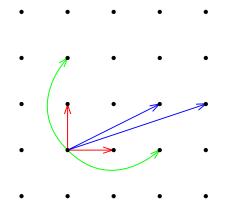
1 Basic properties of lattices

1.1 Definitions

Definition 1. For a given set of *n* linearly independent vectors $b_1, b_2, \ldots, b_n \in \mathbb{R}^m$, we define a lattice $L(b_1, \ldots, b_n)$ as a set of vectors

$$\{\sum a_i b_i | a_i \in \mathbb{Z}\}.$$

The vectors b_1, \ldots, b_n represent the basis of the lattice. Equivalently we can represent the basis as a matrix B whose columns are the basis vectors. We will mostly be interested in those lattices where n = m, i.e. full-rank lattices.



In the example above, the two red vectors span the lattice containing all integer points in \mathbb{Z}^2 , and so do the blue vectors. The two green vectors, however, don't span the same lattice. In the case such as the one above with red and blue vectors, we say that two bases are *equivalent* if they span the same lattice. Two bases are equivalent if and only if we can obtain one from the other using only the three types of operations on its vectors:

1.
$$b_i \rightarrow -b_i$$

2. $b_i \rightarrow b_i + kb_j, k \in \mathbb{Z}$
3. $b_i \leftrightarrow b_j$

Equivalently, if we represent the basis as a matrix B, to bases B and B' are equivalent if and only if there exists a matrix $U \in \mathbb{Z}^{n \times n}$, such that $\det(U) = \pm 1$ and B' = UB. Hence it follows that if two bases B and B' are equivalent we have $|\det(B)| = |\det(B')|$, and we define that to be $\det(L)$, the determinant of the lattice spanned by B or B'.

Definition 2. Given a lattice basis B, we define the fundamental parallelepiped $\mathcal{P}(B)$ as

$$\mathcal{P}(B) = \{\sum a_i b_i | a_i \in \mathbb{R}, a_i \in [0, 1)\}$$

Notice that the whole lattice can be "tiled" with copies of the fundamental parallelepiped. Also, the fundamental parallelepiped doesn't contain any lattice points, except its origin. The volume of $\mathcal{P}(B)$ is equal to $|\det(B)|$, and hence by definition equal to $\det(L)$.

One of the properties of a lattice of particular interest is the length of the shortest nonzero vector in the lattice. Thus we define it as

$$\lambda_1 = \min_{v \in L \setminus \{0^n\}} ||v||,$$

where by the length $|| \cdot ||$ of a vector we will assume Euclidean l_2 norm. We can also define λ_i for i > 1 as follows:

$$\lambda_i(L) = \min_{\substack{v_1 \dots v_i \in L, \\ \text{lin. ind.}}} \max_{1 \le j \le i} ||v_j||.$$

Equivalently we can define the ith successive minimum as

$$\lambda_i(L) = \inf\{r | \dim(\operatorname{span}(L \cap \mathbf{B}(0, r))) \ge i\},\$$

where $\mathbf{B}(0, r)$ represents a closed ball centered at 0 of radius r.

The notion of the successive minima in lattices gives rise to several interesting problems, such as:

SVP: Shortest vector problem - compute λ_1 CVP: Closest vector problem - given $y \notin L$, find $v \in L$ that minimizes ||v - y||

Finding solutions to these problems is hard, but we can achieve meaningful bounds.

1.2 Some properties

Theorem 1 (Blichfield). Given a measurable set $S \subseteq \mathbb{R}^n$ and a full-rank lattice L, such that vol(S) > det(L), there exist $x, y \in S$, $x \neq y$, such that $x - y \in L$.

Proof: Recall that the fundamental parallelepiped can be used to "tile" the lattice. Formalizing this argument, let the basis of L be B. For $x \in L$, let $\mathcal{P}_x(B) = \{x + y | y \in \mathcal{P}(B)\}$. Then $\mathcal{P}_x(B) \cap \mathcal{P}_z(B) = \emptyset$ for $x, z \in L, x \neq z$, and $\bigcup_{x \in L} \mathcal{P}_x(B) = \mathbb{R}^n$. Now define $S_x = S \cap \mathcal{P}_x(B)$. Then $S = \bigcup_{x \in L} S_x$, so $\operatorname{vol}(S) = \sum_{x \in L} \operatorname{vol}(S_x)$. Now define $\overline{S_x} = \{y - x | y \in S_x\}$. Then $\overline{S_x} \subseteq \mathcal{P}_x(B)$ and $\operatorname{vol}(\overline{S_x}) = \operatorname{vol}(S_x)$, and hence

$$\sum_{x \in L} \operatorname{vol}(\overline{S_x}) = \sum_{x \in L} \operatorname{vol}(S_x) = \operatorname{vol}(S) > \det L = \operatorname{vol}(\mathcal{P}(B)).$$

Thus, there exist some $x, y \in L, x \neq y$, such that $\overline{S_x} \cap \overline{S_y} \neq \emptyset$. So, $\exists x, y, z \in L$ such that $\overline{S_x} \cap \overline{S_y} \ni z$. Then, z + x and z + y are two points in S, and $(z + x) - (z + y) = x - y \in L$.

Theorem 2 (Minkowski). Let S be a centrally symetric convex set with $vol(S) > 2^n det(L)$. Then there exists $x \in S \cap (L \setminus \{0^n\})$.

Proof: Let $\frac{S}{2} = \{x | 2x \in S\}$. Then we have $\operatorname{vol}(\frac{S}{2}) = \frac{\operatorname{vol}(S)}{2^n}$, so $\operatorname{vol}(\frac{S}{2}) > \operatorname{det}(L)$. By Blichfield's theorem then $\exists x, y \in \frac{S}{2}, x \neq y$, such that $x - y \in L$. Notice $x - y = \frac{2x + (-2y)}{2}$. We know $2x, 2y \in S$, and by central symmetry of $S, -2y \in S$, so $\frac{2x + (-2y)}{2} \in S$, by convexity of S, and so $x - y \in S \cap (L \setminus \{0^n\})$.

Corollary: $\lambda_1 \leq \sqrt{n} (\det(L))^{1/n}$.

Proof: Look at the cube $C = [-\det(L)^{1/n}, \det(L)^{1/n}]^n$. By Minkowski's theorem, there is a point $x \in L \setminus \{0^n\}$ that is inside the cube, and so is $||x|| \leq \sqrt{n} \det(L)^{1/n}$.

Minkowski's theorem thus gives us an approximation for the SVP problem. But this bound is not very tight; However, there is a surprising result by Lenstra and Schnorr from 1990, that using only the result of Minkowski's theorem finds an n-approximation for the SVP problem. The theorem follows:

Theorem 3 (Lenstra-Schnorr). If we have an algorithm that finds a vector $v \in L \setminus \{0^n\}$, such that $||v|| \leq f(n)(\det(L))^{1/n}$, for some non-decreasing function f(n), then we have a $f(n)^2$ approximation for the SVP problem.

Proof: For a more detailed proof refer to Oded Regev's lecture notes

(http://www.cs.tau.ac.il/~odedr/goto.php?name=ln_dual&link=teaching/lattices_fall_2004/ln/DualLattice.pdf) We will assume familiarity with dual lattices. See next lecture for reference on dual lattices. We apply the algorithm to L(B) and the dual lattice $L(B)^*$, to obtain vectors u and v, respectively. We know that $||u|| \leq f(n)(\det(L(B)))^{1/n}$ and $||u|| \leq f(n)(\det(L(B))^*)^{1/n} = f(n)(\det(L(B)))^{-1/n}$, so $||u||||v|| \leq f(n)^2 = g(n)$. Thus we constructed an algorithm that given a basis B, outputs two vectors $u \in L(B)$ and $v \in L(B)^*$, such that $||u||||v|| \leq g(n)$ for some non-decreasing function g(n). We show that this algorithm gives us a g(n) approximation for SVP. Given a lattice, use the algorithm to obtain a pair of vectors u_1, v_1 . We assume, wlog, that v_1 is primitive as we can easily force it to be. Then, let L' be the projection of L^* on the subspace of $\operatorname{span}(L^*)$ orthogonal to v_1 . Then we recursively apply the above process for L', and back from recursion we get vectors $u_2, \ldots, u_n, v'_2, \ldots, v'_n$. Then we define v_i 's as $v_i = v'_i + \alpha_i v_i$ for the unique $\alpha_i \in (-\frac{1}{2}, \frac{1}{2}]$ for which $v_i \in L^*$. It can be shown that the set of vectors v_1, \ldots, v_n is a basis of L^* . For all $i, ||u_i|| ||\tilde{v}_i|| \leq g(n-i+1) \leq g(n)$. Then let b'_1, \ldots, b'_n be a dual basis of v_1, \ldots, v_n , which is thus a basis of L, and let $b_i = b'_{n-i+1}$ so that we have

$$\min ||\tilde{b}_i|| = \min \frac{1}{||\tilde{v}_i||} \ge \frac{1}{g(n)} \min ||u_i||.$$

Then,

$$\min ||u_i|| \le g(n) \min ||\tilde{b}_i|| \le g(n)\lambda_1(L).$$

Therefore, outputting the shortest u_i guarantees us a g(n) approximation for SVP.

2 LLL (Lenstra, Lenstra, Lovasz)

2.1 Gram-Schmidt orthogonalization

The Gram-Schmidt procedure takes a set of n linearly independent vectors and produces another set of n linearly independent vectors which span the same subspace and are orthogonal to each other. For vectors b_1, \ldots, b_n their Gram-Schmidt orthogonalization are vectors $\tilde{b}_1, \ldots, \tilde{b}_n$, where

$$\begin{split} b_1 &= b_1 \\ \tilde{b}_2 &= b_2 - \alpha_{2,1} \tilde{b}_1 \\ \vdots \\ \tilde{b}_i &= b_i - \sum_{j < i} \alpha_{i,j} \tilde{b}_j, \\ \text{where } \alpha_{i,j} = \frac{\langle b_i, \tilde{b}_j \rangle}{\langle \tilde{b}_j, \tilde{b}_j \rangle} \end{split}$$

This way we have a set of *n* linearly independent vectors where $\langle \tilde{b}_i, \tilde{b}_j \rangle = 0$ for $i \neq j$, i.e. all vectors are orthogonal to each other, and for any $i, 1 \leq i \leq n$, $\operatorname{span}(b_1, \ldots, b_i) = \operatorname{span}(\tilde{b}_1, \ldots, \tilde{b}_i)$.

Expressed in the normalized Grahm-Schmidt basis, the matrix whose columns are $b_1 \dots b_n$ will be:

From this representation we can immediately calculate the determinant of the lattice as

$$\det(L) = \prod_{i=1}^{n} ||\tilde{b}_i||$$

Another property that follows is

$$\lambda_i \ge \min_j ||\tilde{b}_j||$$

We sketch the proof of this property. Notice it suffices to show that $\lambda_1 \geq \min_j ||\tilde{b}_j||$. Let x be an arbitrary nonzero integer vector, $x \in \mathbb{Z}^n$, and we show that $||Bx|| \geq \min_j ||\tilde{b}_j||$. Let i be the largest index for which $x_i \neq 0$. Then

$$|\langle Bx, \tilde{b}_i \rangle| = |\langle \sum_{j=1}^i x_j b_j, \tilde{b}_i \rangle| = |x_i| \langle \tilde{b}_i, \tilde{b}_i \rangle = |x_i| ||\tilde{b}_i||^2$$

The first equality holds because looking at the above representation of B we see that for j < i, $\langle b_j, \tilde{b}_i \rangle = 0$, and the second equality is due to the fact that $\langle b_i, \tilde{b}_i \rangle = \langle \tilde{b}_i, \tilde{b}_i$. On the other side we can bound $|\langle Bx, \tilde{b}_i \rangle|$ by $|\langle Bx, \tilde{b}_i \rangle| \leq ||Bx|| ||\tilde{b}_i||$, and combining these two inequalities we get

$$||Bx|| \ge |x_i|||\tilde{b}_i|| \ge ||\tilde{b}_i|| \ge \min_j \tilde{b}_j.$$

Due to this fact and the definition of λ 's, we can also observe that

$$\max_{i} ||b_i|| \ge \lambda_n \ge ||b_n||$$

2.2 LLL-reduced bases

Definition 3. We call a basis $B = b_1, \ldots, b_n$ *LLL-reduced* if it satisfies the two following constraints:

1. $\forall i, ||\tilde{b}_i|| \le 4||\tilde{b}_{i+1}||$

2.
$$\forall j < i, |\alpha_{i,j}| \le \frac{1}{2}$$

Remark: This is a special case of an LLL-reduced basis. In general, we could define a δ -LLL-reduced basis by changing property 1 to $\forall i, \delta ||\tilde{b}_i||^2 \leq ||\alpha_{i+1}, i\tilde{b}_i + \tilde{b}_{i+1}||^2$, where $\frac{1}{4} < \delta < 1$. Our definition is a special case for $\delta = \frac{5}{16}$.

Corollary 1: If b_1, \ldots, b_n is LLL-reduced, then

$$||b_1|| \ge \lambda_1 \ge 4^{-(n-1)}||b_1||,$$

where the first inequality represents the simple fact that b_1 is a vector in the lattice and hence has to be $\geq \lambda_1$, and the second one follows directly from property 1.

Corollary 2:

$$\max_{i} ||b_i|| \ge \lambda_n \ge \frac{1}{n4^n} \max_{i} ||b_i||$$

By Corollary 1, if we could transform an arbitrary basis into an LLL-reduced one, we would obtain a 4^{-n} approximation for the SVP problem by simply returning b_1 .

2.3 Algorithm

Step 1 (ensuring property 2): Given a basis $B = b_1 \dots b_n$, perform Gram-Schmidt orthogonalization. For each b_i , we know $b_i = \tilde{b}_i + \sum_{j < i} \alpha_{i,j} \tilde{b}_j$. Then let $a_{i,j} = n + \beta_{i,j}$, where $n \in \mathbb{Z}$ and $-\frac{1}{2} < b_j \leq \frac{1}{2}$. Now we change b_i by $b_i - nb_j$. Here is the pseudocode:

for i = 2 to n for j = i - 1 downto 1 $b_i \leftarrow b_i - nb_j$

The crucial part for ensuring correctnes here is the reverse order of the inner loop. Consider the orthonormal basis we get after normalizing Gram-Schmidt vectors, and look at B expressed in that basis:

($ \tilde{b}_1 $	*		*)
	0	$ \tilde{b}_2 $:
	÷	:	·	*
(0	0		$ \tilde{b}_n $)

The matrix is upper-right triangular, and thus by doing the inner loop in the reverse order we will not affect the coefficients to the right of the current column, so after finishing all the * elements, which are of type $\alpha_{i,j}||\tilde{b}_j||$ will become $\leq \frac{1}{2}||\tilde{b}_j||$.

Step 2 (ensuring property 1): If there exists *i* such that $||\tilde{b}_i|| > 4||\tilde{b}_{i+1}||$, swap b_i with b_{i+1} , recompute Gram-Schmidt basis, repeat step 1 and step 2. If this sequence ever terminates, we will indeed have an *LLL*-reduced basis.

Algorithm Analysis: Define a potential function

$$\Phi = \prod_{i=1}^{n} \det(b_1 \dots b_i) = \prod_{i=1}^{n} \prod_{j=1}^{i} ||\tilde{b}_i||$$

In step 1 of the algorithm, the Gram-Schmidt basis doesn't change since we only do operations of type $b_i \leftarrow b_i + kb_j$. It only changes in step 2, where two consecutive columns are swapped. But then if we swapped columns i and i + 1, notice that for no other $j \neq i$, $\det(b_1 \dots, b_j)$ changed. Thus the only difference is in $\det(b_1, \dots, b_i)$ term. Denote the initial potential with Φ , the new potential with Φ' , initial i^{th} Gram-Schmidt vector with \tilde{b}_i , and the new one with \tilde{b}'_i . Then, $\Phi/\Phi' = ||\tilde{b}_i||/||\tilde{b}'_i||$. We have

$$||\tilde{b}'_i|| = ||\alpha_{i+1,i}\tilde{b}_i + \tilde{b}_{i+1}|| \le \frac{1}{2}||\tilde{b}_i|| + ||\tilde{b}_{i+1}|| \le \frac{1}{2}||\tilde{b}_i|| + \frac{1}{4}||\tilde{b}_i|| = \frac{3}{4}||\tilde{b}_i||$$

Thus, $\Phi/\Phi' \ge 4/3$, and the algorithm finishes in polynomial time.

3 SIVP_{α} - short independent vectors problem

Input: basis B Output: $v_1, \ldots v_n \in L$ - linerly independent vectors such that $||v_i|| \leq \alpha \lambda_n$. This gives us a bound on both λ_n and λ_1 . It also implies existance of some cryptographic primitives, according to the two following theorems which we state without proof:

Theorem 4 (Ajtai '96). There exists c such that if solving $SIVP_{n^c}$ is hard, then there exists a one-way function.

Proof: See M. Ajtai - Generating hard instances of lattice problems (STOC '96). The proof is a reduction to modular subset-sum problem.

Following Ajtai's ideas, Goldreich, Goldwasser and Halevi showed that essentially the same construction gives us a collision-resistant hash function.

Theorem 5 (Goldreich, Goldwasser, Halevi '96). There exists c such that if solving $SIVP_{n^c}$ is hard, then there exists collision-resistant hash function.

For the algorithm we are about to present, we assume access to a "Collision-Finder" oracle. Given m vectors of length n, $(a_1, \ldots a_m)$, and a number $q \in \mathbb{Z}$, the oracle returns an m-length vector b such that $\sum b_i a_i = 0^n \pmod{q}$. Now we can present idea behind the algorithm:

Input:

- LLL-reduced basis B
- Collision-Finder oracle, which succeeds with probability n^{-c}
- parameter \hat{s} , $2s(L) < \hat{s} < 4s(L)$ (where s(L) is the smoothing parameter of the lattice, defined below)

Output:

• vector v or "fail"

We will need to prove the three following properties of this algorithm:

- 1. Conditioned on not "fail":
 - (a) $v \in L$
 - (b) with probability $1 \operatorname{negl}(n)$, $||v|| \le O(n^3 \lambda_n)$
- 2. not "fail" happens with probability $\frac{n^{-c}}{2}$
- 3. for all fixed hyperplanes \mathcal{H} of dimension dim $\leq n-1$, $Pr[v \notin \mathcal{H}] \geq n^{-d}$, for some constant d

We can deal with the probabilities by repeating the process n^{c+d+2} times which would give us enough results for v. We now present an implementation of this idea and prove that it follows the three required propreties above. First we need another definition: **Definition 4.** A smoothing parameter s(B) of a lattice basis B is the smallest s such that

$$|G_s^n \mod \mathcal{P}(B) - \mathcal{U}_{\mathcal{P}(B)}| \le n^{-\log(n)}$$

where G_s^n is an *n*-dimensional Gaussian with mean 0 and standard deviation s.

And we use its following two properties:

Theorem 6 (Ajtai).

$$\frac{\lambda_n}{100} \le s(B) \le n\lambda_n$$

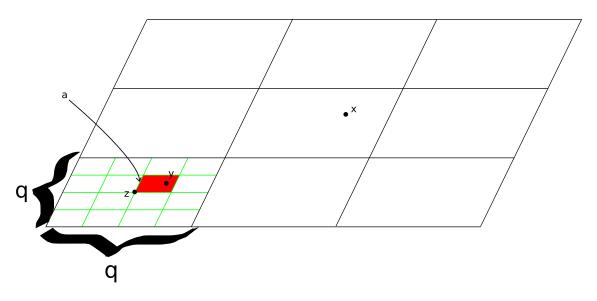
The proof of this theorem will be covered in next week's class.

Fact:

$$\Pr_{x \leftarrow G_s^n} \big[||x|| > 2s \big] \le 2^{-\Omega(n)}$$

Now we can formulate the actual algorithm:

In other words, we split the fundamental parallelepiped into $q \times q$ cells. Then for any given x_i , we take y_i to be x_i modulo the fundamental parallelepiped, a_i to be the cell in which y_i falls, and z_i to be the corner of that cell.



And it remains to prove the three claims from above. In order to do that, we announce our choice of $m = 8n^2$ and $q = 2^{4n}$:

1. (a)

$$v = \sum b_i(x_i - y_i + z_i) = \sum b_i(x_i - y_i) + \sum b_i z_i$$

The $\sum b_i(x_i - y_i)$ falls in L, so we need only consider $\sum b_i z_i$:

$$\sum b_i z_i = \sum b_i \frac{Ba_i}{q} = \frac{B}{q} \sum b_i a_i.$$

But $\sum b_i a_i = 0^n \pmod{q}$, so $\sum b_i a_i = qu$, for some integer vector u. Thus,

$$\sum b_i z_i = \frac{B}{q} q u = B u \in L$$

(b)

$$||v|| = ||\sum b_i(x_i - y_i + z_i)||$$

By triangle inequality,

$$||v|| \le \sum ||b_i|| ||x_i - y_i + z_i||$$

 b_i 's are bounded by 1, so with probability $1 - \operatorname{negl}(n)$,

$$||v|| \le \sum ||x_i|| + \sum ||-y_i + z_i||$$

Because of theorem 6 and the Gaussian distribution, $\sum ||x_i|| \leq m \cdot s = O(n^3 \lambda_n)$. The other term, $\sum ||-y_i+z_i||$ by the definition of y's and z's, we can bound by $\frac{m \cdot \operatorname{diam}(\mathcal{P}(B))}{q} = \operatorname{negl}(n)$, so ||v|| is bounded by $O(n^3 \lambda_n)$

2. It suffices to show that $\Delta(a, \hat{a}) < \operatorname{negl}(n)$, where $\Delta(\cdot, \cdot)$ represents the statistical distnace between the two distributions and $\hat{a} = \{\hat{a}_1, \ldots, \hat{a}_m\}$ uniform over \mathbb{Z}_q^n . Let

$$f: \mathcal{P}(B) \to \mathbb{Z}_q^n$$
$$f(x) = \lfloor qB^{-1}x \rfloor$$
$$\hat{a} = (f(G_{\hat{s}}^n \mod \mathcal{P}(B)), \dots f(G_{\hat{s}}^n \mod \mathcal{P}(B)))$$
$$a = (f(\mathcal{U}_{\mathcal{P}(B)}), \dots, f(\mathcal{U}_{\mathcal{P}(B)}))$$

The statistical distance between a and \hat{a} is m times the distance between each corresponding entries, since they are independent of each other, and that is $m \cdot \operatorname{negl}(n) = \operatorname{negl}(n)$.

3. For any fixed hyperplane \mathcal{H} of dimension $\leq n-1$, we want to bound $Pr[v \notin \mathcal{H}]$. That is,

$$\Pr_{x_i}\left[\sum b_i(x_i - y_i + z_i) \notin \mathcal{H}\right]$$

Denote $-y_i + z_i$ by $f(x_i)$. By averaging argument, there exists some *i* for which $b_i \neq 0$ with probability at least 1/m. Without loss of generality, assume this index is i = 1 and the

corresponding coordinate $b_1 = 1$. Then we can fix x_2, \ldots, x_m such that $Pr[b_1 = 1] > \frac{1}{m^3}$. Then we can look only at

$$\Pr_{x}\left[x - f(x \mod \mathcal{P}(B)) \notin \mathcal{H}\right]$$

We define the following three sets:

Good = {
$$x|x - f(x \mod \mathcal{P}(B)) \notin \mathcal{H} \land b_i = 1$$
}
Bad = { $x|x - f(x \mod \mathcal{P}(B)) \in \mathcal{H} \land b_i = 1$ }
Bad' = { $x|x \in \text{Bad} \land \frac{||x||}{s} \le 10 \log(m)$ }

We want to show that Good has enough probability of happenning, i.e. $Pr[\text{Good}] \ge n^{-d}$. By definitions we see that

 $Pr[\text{Bad} \setminus \text{Bad'}] \le \text{negl}$

Hence, if Pr[Bad'] is also negligible, that would immediately imply that Good is noticable. Otherwise, we claim that $Pr[\text{Good}] \geq Pr[\text{Bad}]/n^d$, which would finish the proof. Consider a mapping defined in the following way: let $u \in L$ be the shortest point not in \mathcal{H} . By definition, $||u|| \leq \lambda_n$. Then the mapping sends $x \to u + x$. If $x \in \text{Bad}$, then $(u + x) \in \text{Good}$. Otherwise, $x - f(x \mod \mathcal{P}(B)) \in \mathcal{H}$, and $x + u - f((x + u) \mod \mathcal{P}(B)) \in \mathcal{H}$, so $u \in \mathcal{H}$, which cannot be. And now look at the distributions of x and u + x:

$$A: G_{\hat{s}}^{n}(x) = \alpha(\hat{s}) \cdot e^{-(\frac{||x||}{\hat{s}})^{2}}$$
$$B: G_{\hat{s}}^{n}(x+u) = \alpha(\hat{s}) \cdot e^{-(\frac{||x||^{2}}{\hat{s}^{2}} + \frac{2||x||||u||}{\hat{s}^{2}} + \frac{||u||^{2}}{\hat{s}^{2}})}$$

and we can bound $\frac{||u||}{\hat{s}}$ terms by O(1), so $\frac{A}{B} < \frac{1}{n^d}$, for some constant d.