Let $X_n = SL_n\mathbb{Z}\backslash SL_n\mathbb{R}$ be the space of n-dimensional lattices L, with the probability Haar measure μ. The goal of this notes is to estimate the average number of LLL bases of lattices in X_n.

Our strategy is in two steps: first we will slightly tweak the LLL condition into a form that is more amenable to direct computation, and estimate the average number of this "modified LLL" bases; then we use this result to deal with real LLL.

Let's quickly recall the definition of an LLL basis. For vectors $x_1, \dots, x_n \in \mathbb{R}^n$, let x_i^* be the component of x_i orthogonal to $\text{span}(x_1, \dots x_{i-1})$. Fix constants $1/2 < \eta \leq 1$ and $1/2 \leq \delta < \eta$. We say $x_1, \dots, x_n$ form an LLL basis with factor (η, δ) if

(i) $\det(x_1 \dots x_n) = 1$.
(ii) $|\mu_{i,j}| \leq \delta$ for all $j < i$, where $\mu_{i,j} := \langle x_i, x_j^* \rangle / \|x_j^*\|^2$.
(iii) $\eta \|x_i^*\| \leq \|x_{i+1}^* + \mu_{i+1,i} x_i^*\|$ for all $i = 1, \dots, n-1$.

In practice, one takes η and δ arbitrarily close to 1 and $1/2$, respectively. In this paper, we will take $\eta = 1$ and $\delta = 1/2$ (in doing so we give up the polynomial-time performance of the LLL algorithm, but since the running time is not of our immediate concern, it does not matter).

It is condition (iii) that we would like to modify. Let's say $x_1, \dots, x_n$ form a *Siegel basis* with factor (η, δ) if they satisfy (i), (ii), and

(iii') $\eta \|x_i^*\| \leq \|x_{i+1}^*\|$ for all $i = 1, \dots, n-1$.

For Siegel reduction, we will set $\delta = 1/2$ unless otherwise stated, and leave $1/2 < \eta \leq 1$ arbitrary.

Average number of Siegel bases. Let

$$\rho(x_1, \dots x_n) = \begin{cases} 1 & \text{if } \{x_1, \dots, x_n\} \text{ is a Siegel basis} \\ 0 & \text{otherwise.} \end{cases}$$

Then the average number of Siegel bases can be expressed as the integral

$$\int_{X_n} \sum \left[\begin{matrix} x_1, \dots, x_n \\ \text{forms a basis of } L \end{matrix} \right] \rho(x_1, \dots, x_n) d\mu.$$

By an integration formula of Schmidt, this equals

$$(1) \quad \prod_{j=2}^{n} \frac{1}{\zeta(j)} \int \dots \int \rho(x_1, \dots, x_{n-1}, t_1 x_1 + \dots + t_{n-1} x_{n-1} + x) dt_1 \dots dt_{n-1} dx_1 \dots dx_{n-1}.$$

Here $x = x(x_1, \dots, x_{n-1})$ is any vector such that the determinant of the $n \times n$ matrix formed by $x_1, \dots, x_{n-1}, x$ is 1. Each dx_i is an integration over $\mathbb{R}^n$, of course, and each dt_i is an integration over $\mathbb{R}$. $\zeta(s)$ is the Riemann zeta function.

One can show that the integral in (1) equals

$$\int_\Delta \prod_{i=1}^{n-1} S_{n+1-i}(r_i) r_i^{n-i} dr_i$$

where the domain of integration Δ equals

$$\Delta = \{(r_1, \dots, r_{n-1}) : \eta r_i \leq r_{i+1} \text{ for } i = 1, \dots n-1\}$$

(we understand $r_n := (r_1 \dots r_{n-1})^{-1}$), and $S_j(x)$ is the surface area of a sphere in $\mathbb{R}^j$ of radius x. Therefore we can rewrite (1) as

$$\prod_{j=2}^{n} \frac{S_j(1)}{\zeta(j)} \int_\Delta \prod_{i=1}^{n-1} r_i^{2(n-i)} dr_i. \tag{2}$$

Remark. One could arrive at the same formula (2) by using the Iwasawa decomposition of $SL_n\mathbb{R}$ and the expression for μ with respect to this decomposition. With this approach, we understand (2) as the measure of a Siegel set in $SL_n\mathbb{R}$.

We reduced the problem to bounding

$$\int_\Delta \prod_{i=1}^{n-1} r_i^{2(n-i)} dr_i \tag{3}$$

from both sides. For the lower bound, first rewrite (3) as

$$\int_{r_{n-1}=0}^{\infty} \int_{r_{n-2}=0}^{\eta^{-1} r_{n-1}} \dots \int_{r_2=0}^{\eta^{-1} r_3} \int_{r_1=0}^{\min(\eta^{-1} r_2, \eta^{-1}(r_{n-1}^2 r_{n-2} \dots r_2)^{-1})} r_1^{2(n-1)} r_2^{2(n-2)} \dots r_{n-1}^2 dr_1 \dots dr_{n-1}. \tag{4}$$

By using $\eta r_i \le r_{i+1}$ repeatedly we find that

$$r_2 \le \eta^{-n+2} r_{n-1}, \eta^{\frac{1}{2}n^2 - \frac{5}{2}n + 5} r_{n-1}^{1-n} \le (r_{n-1}^2 r_{n-2} \dots r_2)^{-1}.$$

Hence by solving

$$\eta^{-n+2} r_{n-1} \le \eta^{\frac{1}{2}n^2 - \frac{5}{2}n + 5} r_{n-1}^{1-n}$$

for r_{n-1}, we find a number $\alpha < 1$ such that whenever $r_{n-1} < \alpha$,

$$r_2 \le (r_{n-1}^2 r_{n-2} \dots r_2)^{-1}$$

holds within the implied domain of the integration above. We could take $\alpha = \eta^{\frac{1}{2}n - \frac{3}{2}}$, for example.

Now (4) is the sum of two terms

$$\int_{r_{n-1}=0}^{\alpha} \int_{r_{n-2}=0}^{\eta_{-1} r_{n-1}} \dots \int_{r_2=0}^{\eta_{-1} r_3} \int_{r_1=0}^{\eta^{-1} r_2} r_1^{2(n-1)} r_2^{2(n-2)} \dots r_{n-1}^2 dr_1 \dots dr_{n-1} \tag{5}$$

and

$$\int_{r_{n-1}=\alpha}^{\infty} \int_{r_{n-2}=0}^{\eta_{-1} r_{n-1}} \dots \int_{r_2=0}^{\eta_{-1} r_3} \int_{r_1=0}^{\min(\eta^{-1} r_2, \eta^{-1}(r_{n-1}^2 r_{n-2} \dots r_2)^{-1})} r_1^{2(n-1)} r_2^{2(n-2)} \dots r_{n-1}^2 dr_1 \dots dr_{n-1}. \tag{6}$$

Clearly (6) is nonnegative, so (5) yields a lower bound on (3), which equals

$$\frac{\eta^{-\frac{1}{6}(4n^3 - 9n^2 - n + 6)} \alpha^{n^2 - 1}}{(n^2 - 1)(n^2 - 2^2) \dots (n^2 - (n-1)^2)}.$$

With the chosen value of α above, this becomes

$$\frac{\eta^{-\frac{1}{6}(n^3 + 2n - 3)}}{(n^2 - 1)(n^2 - 2^2) \dots (n^2 - (n-1)^2)}.$$

It remains to give an upper bound of (3). We will temporarily use the notation

$$\oint_a^b f(x) dx = \begin{cases} \int_a^b f(x) dx & \text{if } a \le b \\ 0 & \text{if } a > b. \end{cases}$$

We can rewrite (3) as

$$\int_{r_1=0}^{\infty}\int_{r_2=\eta r_1}^{\infty}\dots\int_{r_{n-2}=\eta r_{n-3}}^{\infty}\oint_{r_{n-1}=\eta r_{n-2}}^{(\eta r_1\dots r_{n-2})^{-\frac{1}{2}}} r_1^{2(n-1)}\dots r_{n-1}^2 dr_1\dots dr_{n-1}.$$

It is necessary to circle the last integral because $\eta r_{n-2}\le r_{n-1}$ and $r_n\ge \eta r_{n-1}(\Leftrightarrow (\eta r_1\dots r_{n-2})^{-\frac{1}{2}}\ge r_{n-1})$ must be satisfied simultaneously for $(r_1,\dots,r_{n-1})$ to be an element of Δ.

For the last integral to be nontrivial, $\eta r_{n-2}\le(\eta r_1\dots r_{n-2})^{-\frac{1}{2}}$ must hold, which is equivalent to $r_{n-2}\le(\eta^3 r_1\dots r_{n-3})^{-\frac{1}{3}}$. Hence (3) equals

$$\int_{r_1=0}^{\infty}\int_{r_2=\eta r_1}^{\infty}\dots\oint_{r_{n-2}=\eta r_{n-3}}^{(\eta^3 r_1\dots r_{n-3})^{-\frac{1}{3}}}\int_{r_{n-1}=\eta r_{n-2}}^{(\eta r_1\dots r_{n-2})^{-\frac{1}{2}}} r_1^{2(n-1)}\dots r_{n-1}^2 dr_1\dots dr_{n-1}.$$

Again for the second last integral to be nontrivial, $\eta r_{n-3}\le(\eta^3 r_1\dots r_{n-3})^{-\frac{1}{3}}$ must hold, which is equivalent to $r_{n-3}\le(\eta^6 r_1\dots r_{n-4})^{-\frac{1}{4}}$. Repeating this process, we conclude that (3) equals

$$\int_{r_1=0}^{\eta^{-\frac{n-1}{2}}}\dots\int_{r_{n-i}=\eta r_{n-i-1}}^{\eta^{-\frac{i}{2}}(r_1\dots r_{n-i-1})^{-\frac{1}{i+1}}}\dots\int_{r_{n-1}=\eta r_{n-2}}^{(\eta r_1\dots r_{n-2})^{-\frac{1}{2}}} r_1^{2(n-1)}\dots r_{n-1}^2 dr_1\dots dr_{n-1}.$$

This integral is bounded from above by

$$\int_{r_1=0}^{\eta^{-\frac{n-1}{2}}}\dots\int_{r_{n-i}=0}^{\eta^{-\frac{i}{2}}(r_1\dots r_{n-i-1})^{-\frac{1}{i+1}}}\dots\int_{r_{n-1}=0}^{(\eta r_1\dots r_{n-2})^{-\frac{1}{2}}} r_1^{2(n-1)}\dots r_{n-1}^2 dr_1\dots dr_{n-1},$$

which can be computed by elementary means, and equals

$$\frac{\eta^{-\frac{1}{6}(n^3+2n-3)}}{3\cdot\dots(i+1+1/i)\dots(n+1/(n-1))}.$$

Summarizing our results so far,

Theorem 1. *The average number of Siegel bases is*

$$\prod_{j=2}^{n}\frac{S_j(1)}{\zeta(j)}\cdot C\cdot\eta^{-\frac{1}{6}(n^3+2n-3)},$$

where C is some constant between $\prod_{i=1}^{n-1}(n^2-i^2)^{-1}$ and $\prod_{i=1}^{n-1}(i+1+1/i)^{-1}$.

Average number of LLL bases. In this section, we will use Theorem 1 to approximate the average number of the real LLL bases. The idea is simple: since Theorem 1 gives us an estimate on the measure of a Siegel set of $X_n=SL_n\mathbb{Z}\backslash SL_n\mathbb{R}$, we will approximate the set of all LLL bases by Siegel sets and use that estimate.

Let's start with a description of the approximation. Choose $0\le\delta_0<\dots<\delta_k\le 1/2$ and a function σ from $\{1,\dots,n-1\}$ to $\{0,\dots,k-1\}$, and define $\eta_i=\sqrt{1-\delta_i^2}$. We say $x_1,\dots,x_n\in\mathbb{R}^n$ form a (lower) approximate LLL—aLLL for short—basis with factor (η_i,δ_i,σ) if the following conditions are satisfied:

(i) $\det(x_1\dots x_n)=1$.
(ii) $|\mu_{i,j}|\le 1/2$ for all $j<i-1$, where $\mu_{i,j}:=\langle x_i,x_j^*\rangle/\|x_j^*\|^2$.
(iii) $|\mu_{i+1,i}|\in[\delta_{\sigma(i)},\delta_{\sigma(i)+1}]$.
(iv) $\eta_{\sigma(i)}\|x_i^*\|\le\|x_{i+1}^*\|$ for all $i=1,\dots,n-1$.

We will let $\delta_i = i/2k$ unless otherwise mentioned. It is easy to see that

$$\sum_{\sigma} (\text{the average number of aLLL bases with factor } (\eta_i, \delta_i, \sigma)), \tag{7}$$

where σ runs over all functions from $\{1, \dots, n-1\}$ to $\{0, \dots, k-1\}$, approaches the average of LLL bases from below as $k \to \infty$.

By a slight modification of the proof of Theorem 1, we can prove that the summand in (7) is bounded from below by

$$\prod_{j=2}^{n} \frac{S_j(1)}{\zeta(j)} \cdot \frac{1}{k^{n-1}} \int_{r_{n-1}=0}^{\alpha} \cdots \int_{r_i=0}^{\eta_{\sigma(i)}^{-1} r_{i+1}} \cdots \int_{r_1=0}^{\eta_{\sigma(1)}^{-1} r_2} r_1^{2(n-1)} \dots r_{n-1}^2 dr_1 \dots dr_{n-1}, \tag{8}$$

where $\alpha = (\eta_{\sigma(1)} \eta_{\sigma(2)}^2 \dots \eta_{\sigma(n-2)}^{n-2})^{\frac{1}{n-1}} \eta_{\sigma(n-1)}^{-\frac{1}{n+1}}$ this time.

The integral in (8) is equal to

$$\prod_{i=1}^{n-1} (n^2 - i^2)^{-1} \cdot \prod_{i=1}^{n-2} \eta_{\sigma(i)}^{-i(n-i-1)} \cdot \eta_{\sigma(n-1)}^{-(n-1)}.$$

Therefore (7) is bounded from below by

$$\prod_{j=2}^{n} \frac{S_j(1)}{\zeta(j)} \cdot \prod_{i=1}^{n-1} (n^2 - i^2)^{-1} \sum_{\sigma} \frac{1}{k^{n-1}} \prod_{i=1}^{n-2} \eta_{\sigma(i)}^{-i(n-i-1)} \cdot \eta_{\sigma(n-1)}^{-(n-1)}.$$

We notice that the summation above is a product of Riemann sums

$$\sum_{j=0}^{k-1} \frac{1}{k} \eta_j^{-i(n-i-1)} = \sum_{j=0}^{k-1} \frac{1}{k} \sqrt{1 - \left(\frac{j}{2k}\right)^2}^{-i(n-i-1)}$$

for $i = 1, \dots, n-2$, and

$$\sum_{j=0}^{k-1} \frac{1}{k} \sqrt{1 - \left(\frac{j}{2k}\right)^2}^{-(n-1)}.$$

Therefore, taking $k \to \infty$, we have proved

Theorem 2. *The average number of the LLL bases of an n-dimensional lattice is bounded from below by*

$$\prod_{j=2}^{n} \frac{S_j(1)}{\zeta(j)} \cdot \prod_{i=1}^{n-1} (n^2 - i^2)^{-1} \cdot \prod_{i=1}^{n-2} \int_{-\frac{1}{2}}^{\frac{1}{2}} \sqrt{1 - x^2}^{-i(n-i-1)} dx \cdot \int_{-\frac{1}{2}}^{\frac{1}{2}} \sqrt{1 - x^2}^{-(n-1)} dx. \tag{9}$$

The upper bound can be obtained similarly, and equals

$$\prod_{j=2}^{n} \frac{S_j(1)}{\zeta(j)} \cdot \prod_{i=1}^{n-1} a_i^{-1} \int_{-\frac{1}{2}}^{\frac{1}{2}} \sqrt{1 - x^2}^{-i \sum_{j=i}^{n-1} \frac{a_j}{j+1}} dx \tag{10}$$

where $a_i = i + 1 + i^{-1}$.

Distribution of LLL bases. Recall that (7) is bounded from below by a constant (depending on n) times

$$\sum_{\sigma} \frac{1}{k^{n-1}} \prod_{i=1}^{n-1} \eta_{\sigma(i)}^{-\alpha(i)}, \tag{11}$$

where the summation is taken over all maps $\sigma : \{1, \ldots, n-1\} \to \{0, \ldots, k-1\}$, and

$$\alpha(i) := \begin{cases} i(n-i-1) & \text{if } i = 1, \ldots, n-2 \\ n-1 & \text{if } i = n-1. \end{cases} \tag{12}$$

The aim of this section is to show that, for all sufficiently large n and a suitable choice of k (k increasing with n), the main contribution to (11) comes from the single summand with $\sigma(i) = k-1$ for all i. In fact, the contributions from all the other σ's combined are arbitrarily small in proportion to the main term.

By combining this result and its "upper" approximate LLL counterpart (obtained by practically the same argument), this shows that for almost all LLL bases $\{x_1, \ldots, x_{n-1}\}$, the absolute value of $\mu_{i+1,i} = \langle x_{i+1}, x_i^* \rangle / \|x_i^*\|^2$ is near $1/2$ for all i. A little more thoughts reveal that for almost all such LLL bases $\|x_{i+1}^*\| / \|x_i^*\|$ is near $\sqrt{3}/2$ as well. Therefore we conclude that for almost all LLL bases the projections of x_i and x_{i+1} onto the orthogonal complement of $\mathrm{span}(x_1, \ldots, x_{i-1})$ have about the same lengths and form a 60 degree angle.

Let $0 \le a < 1$. Observe that for $n^a \le i \le n - n^a$, $\alpha(i) = O(n^{1+a})$. For other values of i, $\alpha(i) \ge O(n)$.

There are less than k^{n-1} σ's for which $\sigma(i) < k-1$ for at least one $n^a \le i \le n - n^a$. The ratio of the contributions from these σ's to our proclaimed main term is at most

$$k^{n-1} \left(\frac{\eta_{k-2}}{\eta_{k-1}} \right)^{-n^{1+a}}. \tag{13}$$

And there are less than k^{2n^a} σ's for which $\sigma(i) = k-1$ for all $n^a \le i \le n - n^a$ but not all $\sigma(i)$ equals $k-1$. The ratio of the contributions from these σ's to the main term is at most

$$k^{2n^a} \left(\frac{\eta_{k-2}}{\eta_{k-1}} \right)^{-n}. \tag{14}$$

It therefore suffices to show that for a judicious choice of k and a, both (13) and (14) converge to zero as $n \to \infty$. Indeed, since

$$\frac{\eta_{k-2}}{\eta_{k-1}} = \sqrt{1 + \frac{2k-3}{3k^2 + 2k - 1}}$$

and so

$$\left(\frac{\eta_{k-2}}{\eta_{k-1}} \right)^{3k} \to e \text{ as } k \to \infty,$$

(13) and (14) are close to, respectively,

$$k^{n-1} e^{-\frac{1}{3} n^{1+a}/k}, k^{2n^\alpha} e^{-\frac{1}{3} n/k}.$$

Setting $k = n^{1/3}$ and $a = 1/2$, we see that (13) and (14) are close to, respectively,

$$n^{\frac{1}{3}(n-1)} e^{-\frac{1}{3} n^{1+1/6}}, n^{\frac{2}{3} n^{1/2}} e^{-\frac{1}{3} n^{2/3}},$$

both of which converge to zero as $n \to \infty$, as desired.