SINGULAR SYMMETRIC MATRICES

ALEX ESKIN AND YONATAN R. KATZNELSON

ABSTRACT. The varieties, $\mathcal{S}_{n,r}$, of $n\times n$ symmetric matrices of rank r are studied, with the purpose of estimating the function $N(T,n,r)$, which counts the number of integer points whose norm is bounded by T. When $r=n-1$, we obtain the asymptotic formula

$$N(T,n,n-1)=a(n)T^{n(n-1)/2}\log T+O(T^{n(n-1)/2}(\log T)^{1-\epsilon(n)}).$$

For $n-1>r\geq 3$ we obtain upper and lower bounds of the form

$$c_1(n,r)T^{nr/2}<N(T,n,r)<c_2(n,r)T^{nr/2},$$

for T sufficiently large, where $c_1,c_2>0$. The case $r=2$ has been resolved by W. Schmidt.

1. Introduction

Let $\mathcal{S}_{n,r}$ be the affine variety of $n\times n$ symmetric matrices of rank r;

$$\mathcal{S}_{n,r}=\{A\in\mathbb{C}^{n\times n}:A={}^{\mathrm{t}}A,\ \mathrm{rank}(A)=r\}.$$

We denote by $\mathcal{S}_{n,r}(\mathbb{Z})$ the integral points of $\mathcal{S}_{n,r}$. Our purpose here is to analyze the counting function

$$N(T,n,r)=\#\{A\in\mathcal{S}_{n,r}(\mathbb{Z}):\|A\|<T\}.$$

$\|A\|$ denotes the standard Euclidean norm of a matrix in the space $\mathbb{R}^{n\times n}$; $\|A\|^2=\sum_{ij}a_{ij}^2$.

There is a natural correspondence between symmetric integral matrices of rank 2 and integral quadratic forms that are decomposable over the field $\mathbb{Q}$ of rational numbers. In this context Schmidt [11] proves that

$$N(T,n,2)=\alpha(n)T^n\log T+\begin{cases}O\left(T^3\log^{1/2}T\right) & : \quad n=3\\ O\left(T^n\right) & : \quad n\geq 4,\end{cases}\tag{1}$$

with $\alpha(n)>0$. In his work, Schmidt considers the slightly different norm $\|A\|_1=\max_{ij}|a_{ij}|$. We will give an indication of his argument a little further on.

The principal results here are the following two theorems

Theorem 1.1. For $n\geq 4$

$$N(T,n,n-1)=\frac{n(n-1)\mathcal{V}(n)\mathcal{V}(n(n-1)/2)}{2^{2+(n-1)(n-2)/4}\zeta(n)}T^{n(n-1)/2}\log T+O\left(T^{n(n-1)/2}(\log T)^{1-\epsilon(n)}\right),\tag{2}$$

as T tends to infinity. Here $\epsilon(n)=\frac{(n-2)}{n-2+(n-3)(n^2+2)/3}$, which is strictly positive for $n\geq 3$.

Eskin supported in part by a fellowship from the Sloan foundation, and Department of Education grant no. P200A10022.A01.

Here and throughout the paper, $\mathcal{V}(k)$ is the volume of the unit ball in k dimensions, and $\zeta(k)$ is the Riemann zeta function. The notations '$f(x) \ll g(x)$' and '$f(x) = O(g(x))$' mean that there exists a positive constant c such that $|f(x)| < c|g(x)|$, and '$f(x) \approx g(x)$' means $f(x) \ll g(x) \ll f(x)$. Throughout the paper, the constants implied by these notations will depend only on n and/or r, unless explicitly stated otherwise.

Theorem 1.2. For $n \geq 4$ and $n-2 \geq r \geq 3$, and T sufficiently large

$$N(T,n,r) \approx T^{nr/2}, \tag{3}$$

It is interesting that while there is a factor of $\log T$ in the main terms of both $N(T,n,2)$ and $N(T,n,n-1)$, there is no such factor in the main term of $N(T,n,r)$ for $n-2 \geq r \geq 3$.

Rather than specify a rank, we could simply count the number of integral symmetric matrices of determinant 0. The corresponding question for determinant $k \neq 0$ has been dealt with in a general setting by Duke, Rudnick and Sarnak [2], Eskin and McMullen [3], and Eskin, Mozes and Shah [4]. If we set

$$N^k(T,n) = \{A \in \mathcal{S}_{n,n}(\mathbb{Z}) : \|A\| < T,\ \det(A) = k\},$$

then one has

$$N^k(T,n) \sim c(n,k) T^{n(n-1)/2},$$

as T tends to infinity. Once again we note that there is no log factor. To obtain these results, one applies harmonic analysis to homogeneous quotients of affine symmetric spaces. These methods do not apply to the case of determinant 0 however, in part because the variety of singular symmetric matrices is highly singular as a variety. Indeed the singular set of the variety of symmetric matrices of determinant 0 is precisely $\mathcal{S}_{n,n-2}$. This naturally leads one to consider the different singular ranks separately.

To obtain our results we apply and refine the methods of [7, 8]. In those papers the problem of counting integral matrices of fixed (singular) rank was considered, without the further restriction of symmetry. The basic observation exploited there, and here, is that the rows of a rank r integral matrix span an r dimensional lattice in $\mathbb{Z}^n$. This suggests the following notation. For an r dimensional lattice $\Lambda \subset \mathbb{R}^n$ we set

$$S(\Lambda) = \{A \in \mathbb{R}^{n\times n} : A = {}^{\mathrm{t}}A, \quad \text{rows of } A \text{ are in } \Lambda\}.$$

Several facts are clear: $S(\Lambda)$ is a lattice; if $\dim(\Lambda) = r$ and $A \in S(\Lambda)$ then $\mathrm{rank}(A) \leq r$; if $A \in \mathcal{S}_{n,r}(\mathbb{Z})$ then there exists an r dimensional $\Lambda \subset \mathbb{Z}^n$ such that $A \in S(\Lambda)$, e.g. the lattice spanned by the rows of A; $\Lambda_1 \subset \Lambda_2$ implies that $S(\Lambda_1) \subset S(\Lambda_2)$. This last comment indicates that we should restrict our attention to *primitive* sublattices of $\mathbb{Z}^n$.

Definition 1.1. An r dimensional lattice $\Lambda \subset \mathbb{Z}^n$ is **primitive** if it is not properly contained in any other r dimensional sublattice of $\mathbb{Z}^n$.

For any lattice Ξ in $\mathbb{R}^n$ or $\mathbb{R}^{n\times n}$ we set

$$N(T,\Xi) = \#(\Xi \cap B(T)),$$

where $B(T)$ is the ball of radius T centered at the origin in the subspace spanned by Ξ over $\mathbb{R}$. For an r dimensional lattice $\Lambda \in \mathbb{Z}^n$ we write

$$N(T,S(\Lambda)) = N_1(T,S(\Lambda)) + N_2(T,S(\Lambda))$$

where N_1 counts the matrices of rank exactly r and N_2 counts those of rank $< r$.

Proposition 1.1.

$$N(T,n,r) = \sum_{\dim(\Lambda)=r}{}' N_1(T, S(\Lambda)),$$

where the prime indicates that the sum extends over primitive lattices in $\mathbb{Z}^n$.

Proof: This amounts to the observation that if $A \in S(\Lambda_1) \cap S(\Lambda_2)$, for two distinct primitive r dimensional lattices Λ_1 and Λ_2, then A has rank strictly less than r. We remark that the sum on the right is in fact finite since $N_1(T, S(\Lambda)) = 0$ for all but finitely many lattices, as we shall see in §4. □

Notation: Throughout this paper i, j, k, m, n, r, s will denote integers. Positive constants may be denoted c, c_1, c_2, etc., when they are absolute, or depend at most on the dimensions involved. When constants depend on a function or a parameter, we will denote them by c_f, $c_f(\rho)$, etc., or state the dependence explicitly. We use lower case greek letters to denote arbitrary real numbers as the need arises, where ε will generally denote a small positive number. l and u will frequently be used to denote vectors in $\mathbb{R}^n$.

The authors wish to thank Z. Rudnick, P. Sarnak and W. Schmidt for their comments and suggestions.

2. Lattices

We collect here some basic facts about lattices, and set some notation. The proofs of facts stated here without reference may be found in [1, 5, 12].

We denote lattices in $\mathbb{R}^n$ by Λ. The dimension of Λ is the dimension of the subspace, E_Λ, spanned by Λ over $\mathbb{R}$. Every r dimensional lattice Λ has an integral basis, $\{u_1, \dots, u_r\}$, which spans Λ over $\mathbb{Z}$. The *determinant* of Λ, $d(\Lambda)$, is the r dimensional volume of the parallelepiped spanned by an integral basis of Λ in E_Λ; it is independent of the choice of basis. Denote by L the symmetric matrix whose coefficients are the pairwise inner products (u_i, u_j) of the elements of a basis for Λ, then

$$d(\Lambda)^2 = |\det(L)|. \tag{4}$$

For vectors other than matrices, we will denote the standard Euclidean norm by $|\cdot|$.

Definition 2.1. A basis $\{l_1, \dots, l_r\}$ of Λ is **reduced** if it minimizes the product $\prod_{i=1}^r |l_i|$, among all integral bases of Λ.

This is not the classical definition, but is equivalent to that definition, and somewhat easier to state. By the fundamental theorem of Minkowski there exist constants $\gamma_r > 0$ (depending only on r) such that

$$|l_1| \cdots |l_r| \geq d(\Lambda) \geq \gamma_r |l_1| \cdots |l_r|, \tag{5}$$

for any r-dimensional lattice Λ with reduced basis $\{l_1, \dots, l_r\}$. The left hand inequality is trivial, being satisfied by any basis. Thus, in some sense, a reduced basis is 'almost orthogonal'.

Definition 2.2. The **successive minima**, $\mu_1, \dots, \mu_r$, of an r dimensional lattice Λ are defined as follows. For $1 \leq k \leq r$, μ_k is the infimum over all positive numbers x such that the set $B(x) \cap \Lambda$ contains k linearly independent elements. When more than one lattice is being considered, we will sometimes specify by writing $\mu_j(\Lambda)$ for the j^{th} successive minimum of Λ.

The successive minima of Λ are closely related to the lengths of a reduced basis, $\{l_1, \dots, l_r\}$, for Λ. Indeed we have $|l_1| = \mu_1$ and for $j = 2 \dots r$

$$|l_j| \geq \mu_j \gg |l_j|, \tag{6}$$

where the implied constants depend only on j, and thus are uniformly bounded [12]. We adopt the convention that all bases are ordered from shortest vector to longest vector. The following lemma is essentially a reformulation of Minkowski's theorem.

Lemma 2.1. Let $\{l_1, \dots, l_r\}$ be a reduced basis for Λ, then

$$\left|\sum_{j=1}^{r} a_j l_j\right|^2 \geq \delta_r \delta_{r-1} \cdots \delta_1 \sum_{j=1}^{r} a_j^2 |l_j|^2,$$

where $\delta_j = 1 - (1-\gamma_j^2)^{1/2} > 0$, and depends only on j.

Proof: The proof is by induction. We first remark that if $\{l_1, \dots, l_r\}$ is a reduced basis, then for $1 \leq k \leq r$, $\{l_1, \dots, l_k\}$ is a reduced basis for the lattice, Λ_k, which they span [12]. Clearly $|a_1 l_1|^2 = a_1^2 |l_1|^2$ (and $\delta_1 = 1$), so assume that for $2 \leq k \leq r$

$$\left|\sum_{j=1}^{k-1} a_j l_j\right|^2 \geq \delta_{k-1} \cdots \delta_1 \sum_{j=1}^{k-1} a_j^2 |l_j|^2.$$

It is sufficient to prove that $|u + a l_k|^2 \geq \delta_k(|u|^2 + a^2|l_k|^2)$, for any $u \in \Lambda_{k-1}$. By Minkowski's theorem, (5), it follows that

$$d(\Lambda_k) \geq \gamma_k |l_k||l_{k-1}| \cdots |l_1| \geq \gamma_k |l_k| d(\Lambda_{k-1}).$$

¿From elementary geometry we have $d(\Lambda_k) = |l_k| d(\Lambda_{k-1}) \sin\theta_k$, where θ_k is the angle between l_k and the subspace $E_{\Lambda_{k-1}}$. Hence $\sin\theta_k \geq \gamma_k$ and therefore $\cos\theta_k \leq (1-\gamma_k^2)^{1/2} = 1 - \delta_k$. Thus for any $u \in \Lambda_{k-1}$

$$|(u, l_k)| \leq (\cos\theta_k)|u||l_k| \leq (1-\delta_k)|u||l_k|.$$

But now we are done by the triangle inequality. □

Lemma 2.2. Let Λ be an r dimensional lattice with reduced basis $\{l_1, \dots, l_r\}$. Let $\{u_1, \dots, u_r\}$ be another basis of Λ satisfying

$$d(\Lambda) \geq C|u_1| \cdots |u_r|,$$

for some $C > 0$. Then for $1 \leq j \leq r$

$$|u_j| \approx |l_j|,$$

with the implied constants depending only on r and C.

Proof: Set $\beta_j = |l_j|/|u_j|$, then from the definition of the successive minima and (6) it follows that $\alpha \geq \beta_j > 0$ where α depends only on r. Let j_0 be an arbitrary index between 1 and r, then

$$\beta_{j_0} \alpha^{r-1} \geq \beta_{j_0} \left(\prod_{j \neq j_0} \beta_j\right) \geq C,$$

where the right hand inequality follows from our assumption and the fact that $\prod |l_j| \geq d(\Lambda)$ always. Hence $\beta_{j_0} \geq \alpha^{1-r} C$. □

In what follows we will be dealing with lattices (of symmetric matrices) for which we will have bases that are not necessarily reduced, but which satisfy the condition of the lemma above. Thus the purpose of the lemma is to give us estimates on the successive minima of such lattices. We also note that the inequality of Lemma 2.1 remains valid for a basis satisfying the condition of Lemma 2.2, the only difference being that the product of δ's is replaced by a possibly smaller positive constant that may depend on the constant C appearing there, but remains independent of Λ.

Analyzing the counting functions $N(T,n,r)$ depends on obtaining good estimates for the simpler counting functions $N(T,\Lambda)$, for various lattices Λ. When the successive minima (equivalently, the lengths of the elements of a reduced basis) of the lattice Λ are uniformly bounded then an elementary argument yields the following

Proposition 2.3. Let Λ be an r dimensional lattice with reduced basis $\{l_1,\ldots,l_r\}$, and assume that $|l_r|\le c$. Then for $T\ge 1$

$$N(T,\Lambda)=\frac{\mathcal{V}(r)T^r}{d(\Lambda)}+O\left(\frac{T^{r-1}}{d(\Lambda)}\right),$$

where the implied constant depends on c and r.

Proof: Denote by P the fundamental parallelepiped spanned by $\{l_1,\ldots,l_r\}$. Let $\{u_1,\ldots,u_N\}$, $N=N(T,\Lambda)$, be the points of $\Lambda\cap B(T)$, and set

$$P(T)=\bigcup_{j=1}^{N}(u_j+P).$$

The symmetric difference $B(T)\triangle P(T)$ is contained in the set of points whose distance from the boundary $\partial(B(T))$ is less than the diameter of P. The diameter of P is less than $r\cdot|l_r|\le r\cdot c$, so

$$|\mathrm{vol}(B(T))-\mathrm{vol}(P(T))|\ll T^{r-1}.$$

Now, $\mathrm{vol}(P(T))=d(\Lambda)\cdot N(T,\Lambda)$, so

$$d(\Lambda)\cdot N(T,\Lambda)=\mathrm{vol}(B(T))+O(T^{r-1}),$$

and dividing by $d(\Lambda)$ completes the proof, since $\mathrm{vol}(B(T))=\mathcal{V}(r)T^r$. □

In the more general case, when the successive minima of the lattices may be quite large, Schmidt [10] provides the following

Lemma 2.4. Let Λ be an r dimensional lattice with reduced basis $\{l_1,\ldots,l_r\}$. Then

$$N(T,\Lambda)=\frac{\mathcal{V}(r)T^r}{d(\Lambda)}+O\left(\sum_{j=1}^{r}\frac{T^{r-j}}{|l_1|\cdots|l_{r-j}|}\right),$$

where the empty product corresponding to $j=r$ on the right hand side is set equal to 1. The implied constant depends only on r.

We will be concerned mostly with sublattices of $\mathbb{Z}^n$, and in this context we require a bit more terminology. Let Λ be a primitive r-dimensional sublattice of $\mathbb{Z}^n$. The *dual* lattice $\Lambda^\perp\subset\mathbb{Z}^n$ is the primitive $n-r$-dimensional lattice of integer vectors that are orthogonal to E_Λ. The *polar* lattice Λ^p is the lattice of all (rational) vectors in E_Λ having integral inner product with the elements of Λ. Denote by π_Λ the orthogonal projection onto $E_\Lambda^\perp$. The *factor* lattice $\Lambda^\star=\mathbb{Z}^n/\Lambda$ is the image of $\mathbb{Z}^n$ by π_Λ. We have the following basic facts

Lemma 2.5. For a primitive r dimensional $\Lambda\subset\mathbb{Z}^n$, the following identities hold:

- $d(\Lambda^\perp)=d(\Lambda)$.
- $d(\Lambda^p)=d(\Lambda)^{-1}$.

- $\Lambda^\star = (\Lambda^\perp)^p$. Thus $d(\Lambda^\star) = d(\Lambda)^{-1}$.

Proof: See [10]. □

We will be summing various functions over sets of primitive r-dimensional lattices, often defined in terms of a bound on the size of their determinant. We will use the method of summation by parts for this, and so require an estimate for the number of such lattices. Let $P(X, n, r)$ denote the number of primitive r dimensional sublattices of $\mathbb{Z}^n$, with determinant bounded by X. The following result of Schmidt is fundamental.

Theorem [10]

$$P(X, n, r) = a(n, r)X^n + O\left(X^{n-b(n,r)}\right), \tag{7}$$

where

$$a(n, r) = \frac{1}{n}\binom{n}{r}\frac{\mathcal{V}(n)\mathcal{V}(n-1)\cdots\mathcal{V}(n-r+1)}{\mathcal{V}(1)\mathcal{V}(2)\cdots\mathcal{V}(r)}\cdot\frac{\zeta(2)\zeta(3)\cdots\zeta(r)}{\zeta(n)\zeta(n-1)\cdots\zeta(n-r+1)},$$

(with the empty product of ζ's, corresponding to $r = 1$, set equal to 1), and where

$$b(n, r) = \max(1/r, 1/(n-r)).$$

Definition 2.3. For a number $c > 0$, we say that an r-dimensional lattice Λ is c-**regular** if

$$\mu_1(\Lambda) \geq c \cdot d(\Lambda)^{1/r}.$$

Denote by $P(X, n, r, c)$ the number of primitive c-regular r-dimensional sublattices of $\mathbb{Z}^n$ with determinant bounded by X. A modification of Schmidt's proof yields the following

Proposition 2.6. Let $n \geq 2$ and $n - 1 \geq r \geq 1$. There exist constants $c(n, r) > 0$ and $d(n, r) > 0$ such that for X sufficiently large and $c(n, r) \geq c > 0$, $P(X, n, r, c) \geq d(n, r)X^n$.

In other words, a positive proportion of r-dimensional sublattices of $\mathbb{Z}^n$ are $c(n, r)$-regular. This proposition is used in establishing the lower bound in Theorem 1.2. We begin with two lemmas.

Lemma 2.7. Let $1 \leq r - 1 \leq n - 2$ and $\Lambda_1 \subset \mathbb{Z}^n$ be a primitive c_1-regular $r - 1$-dimensional lattice. Assume that $\Lambda_2 \subset \mathbb{Z}^n$ is primitive r-dimensional, $\Lambda_1 \subset \Lambda_2$ and $c_2 d(\Lambda_1)^{r/(r-1)} \leq d(\Lambda_2) \leq c_3 d(\Lambda_1)^{r/(r-1)}$, with $c_3 > c_2 > 0$. Then Λ_2 is c_4-regular with $c_4 > 0$ depending only on the dimensions, c_1, c_2 and c_3.

Proof: If $\mu_1(\Lambda_1) = \mu_1(\Lambda_2)$ then the lemma follows from the regularity of Λ_1 and the bounds on $d(\Lambda_2)$. If not then there is a vector $l_0 \in \Lambda_2 \setminus \Lambda_1$ with $|l_0| = \mu_1(\Lambda_2) < \mu_1(\Lambda_1)$. Now l_0 and a reduced basis for Λ_1 span a sublattice of finite index in Λ_2, whence by Minkowski's theorem $|l_0|\mu_1(\Lambda_1)\cdots\mu_{k-1}(\Lambda_1) \gg d(\Lambda_2)$. Dividing both sides by the product of the successive minima of Λ_1, together with the bounds on $d(\Lambda_2)$ yields the desired lower bound for $|l_0|$. □

Lemma 2.8. Let $\Lambda \subset \mathbb{Z}^n$ be a primitive c_1-regular r-dimensional lattice, $n - 1 \geq r > 1$. The number of primitive $r - 1$-dimensional sublattices $\Lambda' \subset \Lambda$ satisfying

$$d(\Lambda') \leq c_2 d(\Lambda)^{(r-1)/r} \tag{8}$$

is bounded. The bound depends only on n, r, c_1 and c_2.

Proof: For a constant $c > 0$, the number of vectors in Λ of length less than $c \cdot d(\Lambda)^{1/r}$ is uniformly bounded in terms of c, c_1 and r. This follows immediately from Lemma 2.4 and the fact that $|l_j| \geq c_1 d(\Lambda)^{1/r}$ for every element l_j of a reduced basis of Λ.

If $\Lambda' \subset \Lambda$ satisfies (8), it follows that Λ' is itself c_3-regular, with $c_3 \geq c_1/c_2^{1/(r-1)}$. Hence the elements of a reduced basis for Λ' are all bounded in length by $c \cdot d(\Lambda)^{1/r}$, where $c = \gamma_{r-1} c_2 / c_1^{r-2}$. Thus by the first paragraph, the number of such sublattices is bounded independently of Λ. $\square$

Proof (of 2.6): The proof is by induction on r. Fix $n \geq 2$. When $r = 1$ the claim is trivial since every 1-dimensional lattice is 1-regular. Assume now that the proposition holds for $r - 1 \leq n - 2$. We choose X sufficiently large so that $P(X^{(r-1)/r}/2, n, r-1, c) \geq d(n, r-1)(X^{(r-1)/r}/2)^n$ when $c(n, r-1) \geq c > 0$.

Given a primitive c_1-regular ($c_1 = c(n, r-1)$) $\Lambda \subset \mathbb{Z}^n$ of dimension $r - 1$, (for which we assume $d(\Lambda) < X^{(r-1)/r}$), we wish to construct primitive c_2-regular r-dimensional lattices $\Lambda' \subset \mathbb{Z}^n$ containing Λ. For $v \in \mathbb{Z}^n \setminus \Lambda$ we denote by $\Lambda \oplus v$ the lattice spanned by Λ and v. It is immediate that $\Lambda \oplus v = \Lambda \oplus w$ if and only if $\pi_\Lambda(v) = \pm \pi_\Lambda(w)$. Furthermore $\Lambda \oplus v$ is a primitive lattice if and only if $\pi_\Lambda(v)$ is a primitive vector in $\Lambda^\star$ (see [10]). Finally, we note that $d(\Lambda \oplus v) = |\pi_\Lambda(v)| \cdot d(\Lambda)$.

Thus to each pair $\{u, -u\}$ of primitive vectors in $\Lambda^\star$, there corresponds a primitive r-dimensional lattice $\Lambda_u = \Lambda \oplus v_u$, where $v_u \in \pi_\Lambda^{-1}(u) \subset \mathbb{Z}^n$. If we specify that

$$d(\Lambda)^{1/(r-1)} < |u| < 2d(\Lambda)^{1/(r-1)}, \tag{9}$$

then $d(\Lambda)^{r/(r-1)} < d(\Lambda_u) < 2d(\Lambda)^{r/(r-1)}$, and by Lemma 2.7 Λ_u is c_2-regular. Let $E(T, n, r-1, c_1)$ denote the set of primitive c_1-regular $r-1$-dimensional lattices $\Lambda \subset \mathbb{Z}^n$, with $d(\Lambda) < T$, and let $N_p(T, \Lambda^\star)$ be the number of primitive points in $\Lambda^\star \cap B(T)$. If $d(\Lambda) < X^{(r-1)/r}/2$ and $u \in \Lambda^\star$ satisfies (9), then $d(\Lambda_u) < X$. Thus setting $E_X = E(X^{(r-1)/r}/2, n, r-1, c_1)$ we have

$$P(X, n, r, c_2) \gg \sum_{\Lambda \in E_X} [N_p(2d(\Lambda)^{1/(r-1)}, \Lambda^\star) - N_p(d(\Lambda)^{1/(r-1)}, \Lambda^\star)] \tag{10}$$

since by Lemma 2.8 we are overcounting by at most a fixed constant factor independent of X.

We require an estimate for the number of primitive vectors in $\Lambda^\star$ satisfying the condition in (9). A theorem of Mahler [9] implies that for $1 \leq j \leq k$

$$\mu_j(\Lambda) \ll \mu_{k-j+1}(\Lambda^p)^{-1} \tag{11}$$

for any k-dimensional lattice Λ. (The constant implied by Mahler's result is $k!$. This has been improved to k^2 by Lenstra and Schnorr [5].) In particular, this implies

$$\mu_1(\Lambda^\star) \gg \mu_{n-r+1}^{-1}(\Lambda^\perp) \gg d(\Lambda^\perp)^{-1} = d(\Lambda)^{-1}. \tag{12}$$

Furthermore, it follows from (11) that the successive minima of $\Lambda^\star = (\Lambda^\perp)^p$ are bounded independently of $\Lambda \subset \mathbb{Z}^n$, since $\mu_1(\Lambda^\perp) \geq 1$. So by Proposition 2.3 for a variable $\tau > 1$

$$N(\tau, \Lambda^\star) = \mathcal{V}(n-r+1)\tau^{n-r+1}/d(\Lambda^\star) + O(\tau^{n-r}/d(\Lambda^\star)). \tag{13}$$

On the other hand

$$N(\tau, \Lambda^\star) = \sum_{j=1}^{\xi \tau d(\Lambda)} N_p(\tau/j, \Lambda^\star),$$

since $N_p(\tau/j, \Lambda^\star) = 0$ when $j > \xi \tau d(\Lambda) \geq \mu_1(\Lambda^\star)$, where ξ is the constant (which is independent of Λ) implied by (12). Now, by the Möbius inversion principle [6]

$$N_p(\tau, \Lambda^\star) = \sum_{j=1}^{\xi \tau d(\Lambda)} \mu(j) N(\tau/j, \Lambda^\star). \tag{14}$$

Here, $\mu(j)$ is the Möbius function and shouldn't be confused with the successive minima of the lattice. Inserting (13) in (14) we have

$$\begin{aligned} N_p(\tau,\Lambda^\star) &= \frac{\mathcal{V}(n-r+1)\tau^{n-r+1}}{d(\Lambda^\star)}\left(\sum_{j=1}^{\xi\tau d(\Lambda)}\mu(j)j^{r-n-1}+O\left(\tau^{-1}\sum_{j=1}^{\xi\tau d(\Lambda)}j^{r-n}\right)\right) \\ &= \frac{\mathcal{V}(n-r+1)\tau^{n-r+1}}{\zeta(n-r+1)d(\Lambda^\star)}(1+O(\tau^{-1}\log(\tau d(\Lambda)))), \end{aligned} \tag{15}$$

since $n-r\geq 1$.

Now, for $d(\Lambda)$ sufficiently large and $\tau\geq d(\Lambda)^{1/(r-1)}$, the 'O' term on the right of (15) is less than 1/3 in absolute value, whence

$$\frac{4}{3}\frac{\mathcal{V}(n-r+1)\tau^{n-r+1}}{\zeta(n-r+1)d(\Lambda^\star)}>N_p(\tau,\Lambda^\star)>\frac{2}{3}\frac{\mathcal{V}(n-r+1)\tau^{n-r+1}}{\zeta(n-r+1)d(\Lambda^\star)}.$$

Thus, for $d(\Lambda)$ sufficiently large

$$\begin{aligned} N_p(2d(\Lambda)^{1/(r-1)},\Lambda^\star)-N_p(d(\Lambda)^{1/(r-1)},\Lambda^\star) &> \frac{4\mathcal{V}(n-r+1)}{3\zeta(n-r+1)}\frac{d(\Lambda)^{\frac{n-r+1}{r-1}}}{d(\Lambda^\star)} \\ &= \frac{4\mathcal{V}(n-r+1)}{3\zeta(n-r+1)}d(\Lambda)^{\frac{n}{r-1}}, \end{aligned}$$

since $d(\Lambda^\star)=d(\Lambda)^{-1}$, by Lemma 2.5, and $n-r+1\geq 2$. Inserting this estimate in (10) we have for X sufficiently large

$$P(X,n,r,c_2)\gg\sum_{\Lambda\in E_X}d(\Lambda)^{\frac{n}{r-1}}\gg X^n,$$

using summation by parts and the induction hypothesis for $r-1$. □

3. Symmetric Matrices

For (row) vectors $u,v\in\mathbb{R}^n$, we define their *tensor product*

$$u\otimes v\stackrel{def}{=}{}^{t}u\cdot v.$$

The operation on the right is just matrix multiplication, resulting in an $n\times n$ matrix. A direct computation shows that

$$(u\otimes v,u'\otimes v')=(u,u')(v,v'), \tag{16}$$

where $(\cdot,\cdot)$ is the standard inner product in the appropriate Euclidean space. Note that $u\otimes u$ and $u\otimes v+v\otimes u$ are both symmetric matrices.

Proposition 3.1. For $n\geq 3$

$$N(T,n,1)=\frac{\mathcal{V}(n)\zeta(n/2)}{2\zeta(n)}T^{n/2}+\left\{\begin{array}{lcl} O\left(T\log T\right) & : & n=3 \\ O\left(T^{(n-1)/2}\right) & : & n\geq 4. \end{array}\right.$$

Proof: If $A \in \mathcal{S}_{n,1}(\mathbb{Z})$ then A has at least one nonzero row, and all other rows are multiples of this row. This and the condition of symmetry together imply that $A = m \cdot (u \otimes u)$ for a primitive $u \in \mathbb{Z}^n$ and $m \in \mathbb{Z}$. Furthermore, for two primitive vectors $u, v \in \mathbb{Z}^n$, $m_1(u \otimes u) = m_2(v \otimes v)$ if and only if $m_1 = m_2$ and $u = \pm v$. By (16) $\|A\| = m|u|^2$, and so

$$\begin{aligned} N(T,n,1) &= \sum_{m<T} P((T/m)^{1/2}, n, 1) \\ &= a(n,1)T^{n/2} \sum_{m<T} m^{-n/2} + O\left(T^{(n-1)/2} \sum_{m<T} m^{(1-n)/2}\right). \end{aligned}$$

The first sum in the second row yields $\zeta(n/2) + O(T^{1-(n/2)})$, while the second sum is $O(\log T)$ when $n = 3$ and is uniformly bounded when $n \geq 4$. The constant in the main term comes from the formula (7) in Schmidt's theorem. □

The proof of Schmidt's result for $N(T,n,2)$ follows from the similar observation that if $A \in \mathcal{S}_{n,2}(\mathbb{Z})$ then $A = u \otimes v + v \otimes u$, for two noncolinear vectors $u, v \in \mathbb{Z}^n$. This leads to evaluating a sum of the form

$$\sum_{|u||v|<T} 1,$$

over vectors in $\mathbb{Z}^n$. This gives a main term on the order of $T^n \log T$. The factor of $\log T$ can be seen to come from the behavior of the divisor function $d(n)$.

Recall the notation from §1. For $\Lambda \subset \mathbb{Z}^n$ we denote by $S(\Lambda)$ the lattice of symmetric matrices in $\mathbb{Z}^{n \times n}$ whose rows are in Λ. The first order of business is determining an integral basis for $S(\Lambda)$ and to do so we fix a reduced basis of Λ, $\{l_1, \dots, l_r\}$. For $1 \leq i \leq r$ set

$$L_{ii} = l_i \otimes l_i$$

and for $1 \leq i < j \leq r$ set

$$L_{ij} = l_i \otimes l_j + l_j \otimes l_i.$$

Lemma 3.2.

$$\begin{aligned} (L_{ii}, L_{jj}) &= (l_i, l_j)^2 \\ (L_{ii}, L_{st}) &= 2(l_i, l_s)(l_i, l_t) \\ (L_{ij}, L_{st}) &= 2[(l_i, l_s)(l_j, l_t) + (l_i, l_t)(l_j, l_s)]. \end{aligned}$$

In particular, for $1 \leq i \leq j \leq r$

$$|l_i||l_j| \leq |L_{ij}| \leq 2|l_i||l_j|.$$

Proof: This follows immediately from (16). □

Proposition 3.3. The set $\{L_{ij} : 1 \leq i \leq j \leq r\}$ is an integral basis for $S(\Lambda)$.

Proof: Assume that $A = \sum a_{ij} L_{ij} = 0$, denote the rows of A by A_k. It follows that for $1 \leq k \leq n$, $A_k = \sum_1^r \alpha_k^j l_j = 0$, where

$$\alpha_k^j = \sum_{s=1}^r a_{js} l_s^k,$$

and l_s^k is the k^{th} component of l_s. By the linear independence of $\{l_s\}_1^r$, we see that $\alpha_k^j = 0$ for $1 \leq k \leq n$ and $1 \leq j \leq r$. Thus for any $1 \leq j \leq r$, it follows from the definition of α_k^j above and the linear independence of the $\{l_s\}$ that $a_{js} = 0$, $1 \leq s \leq r$. Thus $\{L_{ij}\}$ is an independent set.

Next assume that $A \in S(\Lambda)$ so

$$A = \sum_{j=1}^r u_j \otimes l_j,$$

with $u_j \in \mathbb{Z}^n$. Thus $A = U \cdot L$, where U is the $n \times r$ matrix with the ${}^t u_j$'s as columns and L is the $r \times n$ matrix with the l_j's as rows. Now, since $\mathrm{rank}(L) = r$ it follows that $\mathrm{rank}(A) = \mathrm{rank}(U)$. Hence the subspace E_1 (over $\mathbb{R}$) spanned by the $u_j' s$ is the same as the subspace spanned by the rows of A, since the columns of A are linear combinations of the ${}^t u_j' s$, and A is symmetric. But the rows of A are contained in Λ so $E_1 \subset E_\Lambda$, (the subspace spanned over $\mathbb{R}$ by Λ). Finally, since Λ is primitive $\Lambda = E_\Lambda \cap \mathbb{Z}^n$ and in particular each u_j is in Λ. Writing u_i as a linear combination of the l_j's we obtain

$$A = \sum_{i,j=1}^{r} a_{ij} l_i \otimes l_j = \sum_{i,j=1}^{r} a_{ij} l_j \otimes l_i,$$

where the right hand equality follows once again from symmetry. Linear independence implies that $a_{ij} = a_{ji}$ so that we may collect terms and conclude that

$$A = \sum_{1 \le i \le j \le r} a_{ij} L_{ij}.$$

It follows that $S(\Lambda)$ has dimension $r(r+1)/2$. □

We now calculate the determinant $d(S(\Lambda))$. Denote by $d\mu$ the standard Lebesgue measure on $\mathcal{S}_{n,n}(\mathbb{R})$. For $g \in GL(n, \mathbb{R})$ let $F_g : \mathcal{S}_{n,n} \to \mathcal{S}_{n,n}$ be the linear map $A \to gA\,{}^t g$. By linearity ${F_g}^*(d\mu) = \delta(g) d\mu$ for some scalar function $\delta(g)$. We have $\delta(gh) = \delta(g)\delta(h)$ so that δ is a homomorphism from $GL(n, \mathbb{R})$ to $\mathbb{R}$. Such a homomorphism must be trivial on $SL(n, \mathbb{R})$, since this group is simple, and hence $\delta(g) = \det(g)^\alpha$ for some α.

Proposition 3.4.

$$\delta(g) = \det(g)^{n+1}.$$

Proof: Let g be a diagonal matrix and compute. □

Lemma 3.5.

$$d(S(\Lambda)) = 2^{r(r-1)/4} d(\Lambda)^{r+1}.$$

Proof: Let $\{l_1, \ldots, l_r\}$ be an integral basis for Λ and let $\{f_{r+1}, \ldots, f_n\}$ be an orthonormal basis for the orthogonal complement in $\mathbb{R}^n$ of the subspace spanned by Λ. Denote by $\tilde{\Lambda}$ the lattice spanned by $\{l_1, \ldots, l_r, f_{r+1}, \ldots, f_n\}$. Let $g \in GL(n, \mathbb{R})$ be a linear map that takes $\tilde{\Lambda}$ to $\mathbb{Z}^n$ then it follows that

$$|\det(g)|^{-1} = d(\tilde{\Lambda}) = d(\Lambda).$$

Furthermore, F_g takes $S(\tilde{\Lambda})$ to $\mathcal{S}_{n,n}(\mathbb{Z})$, and so by the previous lemma,

$$d(S(\tilde{\Lambda})) = d(\Lambda)^{n+1} d(\mathcal{S}_{n,n}(\mathbb{Z})) = 2^{n(n-1)/4} d(\Lambda)^{n+1}. \tag{17}$$

The second equality follows from (4) and Lemma 3.2 applied to $\mathcal{S}_{n,n}(\mathbb{Z})$. Applying (4) to $S(\tilde{\Lambda})$ yields

$$d(S(\tilde{\Lambda})) = 2^{n(n-1)/4 - r(r-1)/4} d(S(\Lambda)) d(\Lambda)^{n-r},$$

and combining this with (17) completes the proof. □

Comment:

$$d(\Lambda)^{r+1} \gg |l_1|^{r+1} \cdots |l_r|^{r+1} \gg \prod_{1 \le i \le j \le r} |L_{ij}|,$$

so that the basis we have constructed satisfies the condition of Lemma 2.2 applied to the lattice $S(\Lambda)$. Thus the successive minima of $S(\Lambda)$ are, to within a uniform multiplicative factor, the pairwise products of the successive minima of the lattice Λ.

4. Upper and Lower Bounds

In order to evaluate the sum in Proposition 1.1, it is first necessary to determine the set of primitive r dimensional lattices in $\mathbb{Z}^n$ for which $N_1(T, S(\Lambda)) > 0$.

Lemma 4.1. Let $\Lambda \subset \mathbb{Z}^n$ be a primitive r dimensional lattice with reduced basis $\{l_1, \ldots, l_r\}$. Then $N_1(T, S(\Lambda)) > 0$ implies that

$$|l_i||l_j| \ll T \tag{18}$$

for all pairs (i, j) such that $i + j \leq r + 1$.

Proof: Let $A \in S(\Lambda)$, with $|A| \leq T$. We write

$$A = \sum_{1 \leq i \leq j \leq r} a_{ij} L_{ij},$$

with $a_{ij} \in \mathbb{Z}$. From Lemma 2.1 it follows that $a_{ij} \neq 0$ implies $|L_{ij}| \leq c_1 T$, for an appropriate c_1, which depends only on r. Let $c_2 = 2c_1$, and assume that $|l_i||l_j| > c_2 T$ for some pair of indices (i, j), with $i + j = r + 1$. Now for $s \geq i$ and $t \geq j$ we have

$$|L_{st}| \geq |l_s||l_t| \geq |l_i||l_j| \geq 1/2|L_{ij}| > c_1 T,$$

so that $a_{st} = 0$ for all such pairs. We now show that it follows from this that $\operatorname{rank}(A) \leq r - 1$, and hence that $N_1(T, S(\Lambda)) = 0$.

Since this is now a question of linear dependence, we may assume that Λ is in fact spanned by the first r standard unit vectors, and thus the last $n - r$ rows and columns of A are all 0's. Hence the rank of A is the rank of the first $r \times r$ minor. From our assumption it follows that the last $r - i + 1 = j$ rows of this minor have 0's in their last $r - j + 1 = i$ columns, so that they span at most an $r - i = j - 1$ dimensional space. Thus these last j rows are linearly dependent and $\operatorname{rank}(A) \leq r - 1$. □

Corollary 4.2. If $N_1(T, S(\Lambda)) > 0$ then $d(\Lambda) \ll T^{r/2}$.

Proof: Assume that $N_1(T, S(\Lambda)) > 0$. By the previous lemma this implies that

$$(c_2 T)^r \geq \prod_{i+j=r+1} |l_i||l_j| = (|l_1| \cdots |l_r|)^2 \geq d(\Lambda)^2,$$

with c_2 as above. □

We now establish an upper bound for $N(T, n, r)$. But first, some more notation. Let $E(T, n, r)$ be the set of primitive r dimensional lattices $\Lambda \subset \mathbb{Z}^n$, for which $N_1(T, S(\Lambda)) > 0$. In particular, the reduced bases of such lattices satisfy (18). Let $E^\star(T, n, r)$ be the set of primitive r dimensional sublattices of $\mathbb{Z}^n$ that satisfy $d(\Lambda) \leq (c_2 T)^{r/2}$. For fixed $c > 0$ let $E(T, n, r, c)$ be the subset of $E^\star$ that are c-regular. By Corollary 4.2 $E(T, n, r) \subset E^\star(T, n, r)$, and by choosing c sufficiently small, we may assume that $E(T, n, r, c) \subset E(T, n, r)$.

Theorem 4.1. For $n \geq 4$ and T sufficiently large,

$$N(T, n, n-1) = O\left(T^{n(n-1)/2} \log T\right),$$

and for $n - 2 \geq r \geq 3$

$$N(T, n, r) = O\left(T^{nr/2}\right).$$

Proof: By Proposition 1.1, and the definitions above

$$N(T,n,r) = \sum_{E(T,n,r)} N_1(T,S(\Lambda)) < \sum_{E(T,n,r)} N(T,S(\Lambda)).$$

By Lemma 2.4 and Lemma 3.5

$$N(T,S(\Lambda)) = \frac{\mathcal{V}(r(r+1)/2)T^{r(r+1)/2}}{2^{r(r-1)/4}d(\Lambda)^{r+1}} + O\left(\sum_{k=1}^{\varrho} \frac{T^{\varrho-k}}{\nu_1\nu_2\cdots\nu_{\varrho-k}}\right), \tag{19}$$

where $\varrho = r(r+1)/2$. In this equation on the right hand side and in the following equation on the left hand side, the term in the sum corresponding to $k=\varrho$ is set equal to 1. The ν_j's are the successive minima of $S(\Lambda)$. By the comments at the end of §3, these successive minima are essentially the pairwise products $|l_i||l_j|$, for a reduced basis $\{l_1,\ldots,l_r\}$ of Λ. For a fixed Λ

$$\sum_{k=1}^{\varrho} \frac{T^{\varrho-k}}{\nu_1\nu_2\cdots\nu_{\varrho-k}} \ll \prod_{1\le i\le j\le r}{}' \max\left(1, \frac{T}{|l_i||l_j|}\right), \tag{20}$$

where the prime on the product indicates that the factor '$\max(1,T/|l_r|^2)$' is omitted (this corresponds to omitting the largest successive minimum). The product is a combinatorial manipulation which captures the largest term in the sum. It follows that

$$N(T,n,r) \ll T^{r(r+1)/2} \sum_{E(T,n,r)} d(\Lambda)^{-(r+1)} + \mathcal{E}_1(T,n,r), \tag{21}$$

where $\mathcal{E}_1$ is the sum over $E(T,n,r)$ of the product in (20). For the sake of continuity, we defer the evaluation of all the remainder terms to §7. It is established there that for $r \ge 3$, $\mathcal{E}_1(T,n,r) = O(T^{nr/2})$. Since $E(T,n,r) \subset E^\star(T,n,r)$

$$\sum_{E(T,n,r)} d(\Lambda)^{-(r+1)} < \sum_{E^\star(T,n,r)} d(\Lambda)^{-(r+1)},$$

and we evaluate the sum on the right by applying summation by parts together with Schmidt's formula (7). For $r = n-1$ this yields

$$\sum_{E^\star(T,n,n-1)} d(\Lambda)^{-n} = \frac{n(n-1)\mathcal{V}(n)}{4\zeta(n)} \log T + O(1), \tag{22}$$

and for $n-2 \ge r \ge 3$

$$\sum_{E^\star(T,n,r)} d(\Lambda)^{-(r+1)} \ll T^{r(n-r-1)/2}.$$

Substituting these estimates in (21), together with the bound on $\mathcal{E}_1$, completes the proof. □

Lemma 4.3. Let $\Lambda \subset \mathbb{Z}^n$ be an r-dimensional lattice with reduced basis $\{l_1,\ldots,l_r\}$. Then

$$N_2(T,S(\Lambda)) = O\left(\frac{T^{r(r-1)/2}}{|l_1|^{r(r-1)}} \log_+ \left(T/|l_1|^2\right)\right),$$

where the implied constant depends on n and r, and $\log_+ x$ denotes $\max(1,|\log x|)$.

Proof: Let $A \in S(\Lambda)$ with $T > ||A||$. We write $A = \sum_{ij} a_{ij} L_{ij}$ and conclude from Lemma 2.1 that

$$T^2 > ||A||^2 \gg \sum_{ij} a_{ij}^2 ||L_{ij}||^2.$$

Since $||L_{ij}||^2 \geq |l_1|^4$, it follows that

$$T/|l_1|^2 \gg \left(\sum_{ij} a_{ij}^2 \right)^{1/2}. \tag{23}$$

Now assume that $\text{rank}(A) < r$, and consider the $r \times r$ symmetric matrix $\tilde{A} = (a_{ij})$ (setting $a_{ji} = a_{ij}$ for the missing entries). Then $\text{rank}(A) = \text{rank}(\tilde{A})$. To see this we may once again assume that Λ is spanned by the first r standard unit vectors, as in the proof of Lemma 4.1, since this is simply a question of linear dependence. But in that case the first $r \times r$ minor of A is $\tilde{A}$, and all other entries are 0. Hence $\tilde{A}$ is singular, so Theorem 4.1 may be applied if $r > 3$ and Schmidt's result (1) may be applied if $r = 3$, yielding the desired upper bound in both cases. □

Theorem 4.2. For $n \geq 4$

$$N(T, n, n-1) \gg T^{n(n-1)/2} \log T,$$

and for $n - 2 \geq r \geq 3$

$$N(T, n, r) \gg T^{nr/2}.$$

The implied constants depend on n and r.

Proof: We will produce the lower bound by considering only c-regular lattices (with $c = c(n, r)$, see Proposition 2.6), and furthermore we will restrict to $E(\varepsilon_0 T, n, r, c)$ for a suitably chosen $\varepsilon_0 > 0$.

If $\Lambda \in E(T, n, r, c)$ then the successive minima, $\{\nu_j\}$, of the lattice $S(\Lambda)$ are all on the order of $d(\Lambda)^{2/r}$, to within a multiplicative factor that is independent of Λ. Thus if $T/d(\Lambda)^{2/r}$ is sufficiently large, the main term in (19) dominates the error term. Choosing $\varepsilon_1 > 0$ sufficiently small we have

$$N(T, S(\Lambda)) > \frac{2}{3} \frac{\mathcal{V}(r(r+1)/2) T^{r(r+1)/2}}{2^{r(r-1)/4} d(\Lambda)^{r+1}}, \tag{24}$$

for all $\Lambda \in E(\varepsilon_1 T, n, r, c)$. If Λ is c-regular then $|l_1|^2 \geq c^2 d(\Lambda)^{2/r}$, and so by Lemma 4.3

$$N_2(T, S(\Lambda)) \ll \frac{T^{r(r-1)/2}}{d(\Lambda)^{r-1}} \log_+ \left(T/d(\Lambda)^{2/r} \right).$$

Hence, we may choose $\varepsilon_2 > 0$ sufficiently small so that

$$\frac{1}{3} \frac{\mathcal{V}(r(r+1)/2) T^{r(r+1)/2}}{2^{r(r-1)/4} d(\Lambda)^{r+1}} > N_2(T, S(\Lambda))$$

when $\Lambda \in E(\varepsilon_2 T, n, r, c)$. Setting $\varepsilon_0 = \min(\varepsilon_1, \varepsilon_2)$ guarantees that

$$N_1(T, S(\Lambda)) = N(T, S(\Lambda)) - N_2(T, S(\Lambda)) > \frac{1}{3} \frac{\mathcal{V}(r(r+1)/2) T^{r(r+1)/2}}{2^{r(r-1)/4} d(\Lambda)^{r+1}},$$

for all $\Lambda \in E(\varepsilon_0 T, n, r, c)$. We insert this last estimate in the sum in Proposition 1.1 to obtain

$$N(T, n, r) \gg T^{r(r+1)/2} \sum_{E(\varepsilon_0 T, n, r, c)} d(\Lambda)^{-(r+1)} \gg \begin{cases} T^{nr/2} & : \ n - 2 \geq r \geq 3 \\ T^{n(n-1)/2} \log T & : \ r = n - 1, \end{cases}$$

summing by parts and applying Proposition 2.6. □

Corollary 4.4. For $n-2 \geq r \geq 3$ and T sufficiently large

$$N(T,n,r) \approx T^{nr/2},$$

where the implied constants depend on n and r.

Proof: This is Theorem 1.2, and it follows from Theorems 4.1 and 4.2 . □

5. Proof of Theorem 1.1

In this section we restrict our attention to the case $r = n-1$, and we use the shorthand $E = E(T,n,n-1)$, $E^\star = E^\star(T,n,n-1)$.

Definition 5.1. A lattice $\Lambda \in E$ is said to have δ **bounded quotients** if $|l_j|/|l_{j+1}| \geq \delta$ for $1 \leq j \leq n-2$. Here $\delta > 0$ and $\{l_1, \dots, l_{n-1}\}$ is a reduced basis for Λ.

We will denote the subset of these lattices by $E_\delta = E_\delta(T, n-1)$. We introduce two more sums:

$$\mathcal{E}_2 = \mathcal{E}_2(T, n-1, \delta) = T^{n(n-1)/2} \sum_{E^\star \setminus E_\delta} d(\Lambda)^{-n}, \tag{25}$$

$$\mathcal{E}_3 = \mathcal{E}_3(T, n-1, \delta) = \sum_{E_\delta} N_2(T, S(\Lambda)) \tag{26}$$

When we come to bounding the remainder terms, δ will be chosen to balance the sizes of $\mathcal{E}_2$ and $\mathcal{E}_3$.

By Proposition 1.1, we have $N(T,n,n-1) = \sum_E N_1(T, S(\Lambda))$, but this is not a convenient expression to evaluate. Since $N_1 = N - N_2$, and we have a reasonably convenient expression for N, it essential to control N_2. To do this we restrict to E_δ:

$$\begin{aligned} N(T,n,n-1) &= \sum_E N_1(T,S(\Lambda)) \\ &= \sum_{E_\delta} N_1(T,S(\Lambda)) + \sum_{E\setminus E_\delta} N_1(T,S(\Lambda)) \\ &= \sum_{E_\delta} N(T,S(\Lambda)) + O\left(\mathcal{E}_3 + \sum_{E\setminus E_\delta} N(T,S(\Lambda))\right) \\ &= \frac{\mathcal{V}(n(n-1)/2)T^{n(n-1)/2}}{2^{(n-1)(n-2)/4}} \sum_{E_\delta} d(\Lambda)^{-n} + O\left(\mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3\right) \\ &= \frac{\mathcal{V}(n(n-1)/2)T^{n(n-1)/2}}{2^{(n-1)(n-2)/4}} \sum_{E^\star} d(\Lambda)^{-n} + O\left(\mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3\right). \end{aligned}$$

The sum on the left of the third line is replaced by the sum on the fourth line $+O(\mathcal{E}_1)$. The sum on the right of the third line is replaced by $O(\mathcal{E}_1 + \mathcal{E}_2)$. Finally, the transition from the fourth to the fifth line contributes one more $O(\mathcal{E}_2)$.

To evaluate the main term in the last line above, we once more apply summation by parts together with Schmidt's theorem for $P(T^{(n-1)/2}, n, n-1)$. This yields

$$N(T,n,n-1) = \frac{n(n-1)\mathcal{V}(n)\mathcal{V}(n(n-1)/2)}{2^{2+(n-1)(n-2)/4}\zeta(n)} T^{n(n-1)/2}\log T + O\left(T^{n(n-1)/2} + \mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3\right). \tag{27}$$

In the last section we will show that $\mathcal{E}_1(T,n,n-1) = O(T^{n(n-1)/2})$ and that δ can be chosen so that $\mathcal{E}_2(T,n,n-1)+\mathcal{E}_3(T,n,n-1) = O(T^{n(n-1)/2}(\log T)^{1-\frac{n-2}{n-2+\beta(n)}})$, where $\beta(n) = (n-3)(n^2+2)/3$ is strictly positive for $n \geq 4$. This, together with (27), completes the proof of Theorem 1.1 . □

6. Sums and Integrals

To bound the remainder terms, we will be considering multiple sums ranging over certain bounded sets of integer vectors in $\mathbb{R}^n$. To estimate such sums, it is convenient to be able to compare them to integrals which can then be computed directly. To do this the functions that are being summed, as well as the ranges of summation, need to be reasonably well behaved.

In this section we will first describe a class of functions and domains for which this type of estimate works. We will then give a more specific version of the estimate, (Lemma 6.7), for the specific domains we use in the next section.

Let $\mathcal{P}$ be the set of strictly positive continuous functions, f, on $(0,\infty)$ satisfying the following condition. For each $\rho \geq 1$ there exists a $c_f(\rho)$, $\infty > c_f(\rho) > 0$, such that

$$f(y) \leq c_f(\rho) f(x), \tag{28}$$

when $\rho x \geq y \geq x \geq 1/2$. (The lower bound 1/2, is artificial. Any $\varepsilon > 0$ would do as well.) Note that if (28) holds for any one fixed $\rho_0 > 1$, then it holds for all $\rho \geq 1$, with $c_f(\rho) \leq c_f(\rho_0)^{\kappa(\rho)}$, and $\kappa(\rho) = \log_{\rho_0}(\rho) + 1$.

Let $\mathcal{X}$ be the set of strictly positive functions on $(0,\infty)$ satisfying

$$f(y) \leq c_f f(x), \tag{29}$$

when $x+1 \geq y \geq x \geq 1/2$. Here also, $\infty > c_f > 0$. Then $\mathcal{P} \subset \mathcal{X}$, since we may take $c_f = c_f(3)$. The set $\mathcal{P}$ consists of functions whose behavior is 'no worse' than that of a power function, and $\mathcal{X}$ consists of functions whose behavior is 'no worse' than that of an exponential function. More precisely

Proposition 6.1. With $\mathcal{P}, \mathcal{X}$ as above we have

(a) If f is strictly positive, continuous and nonincreasing then $f \in \mathcal{P}$.
(b) $f(x) = x^\alpha \in \mathcal{P}$, for $\alpha > 0$.
(c) $f(x) = \log_+^\beta(x) \in \mathcal{P}$, for $\beta > 0$, where $\log_+(x) = \max(1, |\log x|)$.
(d) If $f, g \in \mathcal{P}$ $(\in \mathcal{X})$ then $h = \max(f,g) \in \mathcal{P}$ $(\in \mathcal{X})$.
(e) If $f, g \in \mathcal{P}$ $(\in \mathcal{X})$ then $h = f \cdot g \in \mathcal{P}$ $(\in \mathcal{X})$.
(f) $f(x) = c^x \in \mathcal{X}$, for $c > 0$.

Proof: All the functions listed above are strictly positive and continuous on $[1/2,\infty)$. It remains to show the existence of an appropriate $c_f(\rho)$ (c_f) for each.

(a) Clear, with $c_f(\rho) \leq 1$ for all $\rho > 1$.
(b) For $\rho x \geq y \geq x$, we have $y^\alpha = (y/x)^\alpha x^\alpha \leq \rho^\alpha x^\alpha$.
(c) For $\rho x \geq y \geq x$, we have $\log_+^\beta y = (\log_+ y/\log_+ x)^\beta \log_+^\beta x < (\log\rho + 1)^\beta \log_+^\beta x$.
(d) If, e.g. $f(y) \geq g(y)$, then

$$h(y) = f(y) \leq c_f(\rho) f(x) \leq c_f(\rho) h(x),$$

since $h(x) \geq f(x)$ by defintion. It follows that $c_h(\rho) = \max(c_f(\rho), c_g(\rho))$. The proof is identical for $\mathcal{X}$, with c_f replacing $c_f(\rho)$.
(e) It is immediate that $c_{fg}(\rho) \leq c_f(\rho) c_g(\rho)$, and $c_{fg} \leq c_f c_g$.
(f) $c^y \leq c^{x+1} = c \cdot c^x$, for $x+1 \geq y$. □

Remark: It is also easy to see that if $cg(x) < f(x) < g(x)$ on $(0, \infty)$, with $c > 0$ and $g \in \mathcal{P}$ $(g \in \mathcal{X})$, then $f \in \mathcal{P}$ $(f \in \mathcal{X})$.

For the purposes of replacing sums by integrals, the class $\mathcal{X}$ is sufficient. The class $\mathcal{P}$ however, has the following important property, which does not extend to all of $\mathcal{X}$. It will allow us to replace sums by integrals iteratively.

Proposition 6.2. Assume that $g(x)$ is an increasing positive function, and that $h(x) \in \mathcal{P}$. We also assume that $h(x) > 2g(x)$, unless h is nonincreasing, in which case we only require $h(x) > g(x)$. If $f(t, x)$ belongs to $\mathcal{P}$ as a function of each variable separately, with constants $c_{f,1}(\rho)$ for the first variable, and $c_{f,2}(\rho)$ for the second variable, then

$$F(x) = \int_{g(x)}^{h(x)} f(t, x)\, dt \ \in \mathcal{P}.$$

The constant $c_F(\rho)$ depends only on $c_h(\rho), c_{f,1}(\rho)$ and $c_{f,2}(\rho)$.

Proof: Let $\rho x \geq y \geq x \geq 1/2$. If $h(x)$ is nonincreasing then

$$F(y) = \int_{g(y)}^{h(y)} f(t, y)\, dt \leq c_{f,2}(\rho) \int_{g(x)}^{h(x)} f(t, x)\, dt = c_{f,2}(\rho) F(x).$$

In the general case we have $\gamma = c_h(\rho) \geq 1$, and

$$F(y) = \int_{g(y)}^{h(y)} f(t, y)\, dt \leq c_{f,2}(\rho) \int_{g(x)}^{\gamma h(x)} f(t, x)\, dt.$$

Consider the linear map of the interval $[g(x), \gamma h(x)]$ onto the interval $[g(x), h(x)]$, defined by $\tau = \lambda t + \mu$, where $\lambda = (h(x) - g(x))/(\gamma h(x) - g(x))$, and $\mu = g(x) - \lambda g(x)$. We have $\lambda \leq 1$, and thus $\mu \geq 0$, since $g(x) \leq h(x) \leq \gamma h(x)$. Also

$$\lambda^{-1} = \frac{\gamma h(x) - g(x)}{h(x) - g(x)} \leq \frac{2\gamma h(x)}{h(x)} = 2\gamma,$$

since $0 < 2g(x) \leq h(x)$. Making the change of variable $\tau = \lambda t + \mu$ gives

$$F(y) \leq \lambda^{-1} c_{f,2}(\rho) \int_{g(x)}^{h(x)} f(\lambda^{-1}(\tau - \mu), x)\, d\tau \leq 2\gamma c_{f,2}(\rho) \int_{g(x)}^{h(x)} f(\lambda^{-1}(\tau - \mu), x)\, d\tau.$$

Now, $\tau \leq \lambda^{-1}(\tau - \mu) \leq 2\gamma\tau$, for $g(x) \leq \tau \leq h(x)$, since $\mu \geq 0$, and $\lambda^{-1} \leq 2\gamma$. Thus

$$f(\lambda^{-1}(\tau - \mu), x) \leq c_{f,1}(2\gamma) f(\tau, x),$$

for τ in this range. Hence $F(y) \leq 2\gamma c_{f,2}(\rho) c_{f,1}(2\gamma) F(x)$. $\square$

We now come to our basic lemma relating sums to integrals. For a domain $D \subset \mathbb{R}^n$, and constants $B > A > 0$ let

$$D(A, B) = \{u \in D \cap \mathbb{Z}^n : A \leq |u| < B\},$$

and let

$$\nu_D(x) = |D(x, x + 1)|.$$

Lemma 6.3. Let $D \subset \mathbb{R}^n$, and for $B > A \geq 1$ and $f \in \mathcal{X}$ set

$$S(f; D, A, B) = \sum_{D(A,B)} f(|l|).$$

Then for any $\nu \in \mathcal{X}$ for which $\nu_D(x) \leq \nu(x)$ on $[A, B]$

$$S(f; D, A, B) < 2c_f^2 c_\nu \int_{A-1/2}^{B} f(t)\nu(t)\, dt.$$

Proof: Let $k \geq 0$ be the integer satisfying $A + k < B \leq A + k + 1$, then

$$S(f; D, A, B) \leq \sum_{i=0}^{k} S(f; D, A+i, A+i+1). \tag{30}$$

Since $f \in \mathcal{X}$ and $\nu \geq \nu_D$ on this range, it follows that

$$S(f; D, A+i, A+i+1) \leq c_f f(A+i)\nu_D(A+i) \leq c_f f(A+i)\nu(A+i), \tag{31}$$

for $0 \leq i \leq k$. By Proposition 6.1, $f \cdot \nu \in \mathcal{X}$, and so $f(A+i)\nu(A+i) \leq c_f c_\nu f(x)\nu(x)$, for $x \in (A+i-1/2, A+i)$. Thus

$$f(A+i)\nu(A+i) \leq 2c_f c_\nu \int_{A+i-1/2}^{A+i} f(t)\nu(t)\, dt.$$

This, together with (30) and (31) completes the proof. □

We now describe the type of domain which we will be considering in the next section. These domains are defined in terms of basic geometric conditions satisfied by reduced bases.

Definition 6.1. For a vector $u \in \mathbb{R}^n$ let

$$H(u) = \{v \in \mathbb{R}^n : |(u, v)| \leq |u|^2/2\}.$$

Equivalently, $H(u)$ is the set of points in $\mathbb{R}^n$ whose distance from the subspace $u^\perp$ is less than $|u|/2$. We remark that $H(0) = \mathbb{R}^n$.

Lemma 6.4. Let l_i and l_j be elements in the same reduced basis for a lattice Λ, with $|l_j| \geq |l_i|$ say. Then $l_j \in H(l_i)$.

Proof: Since l_j can be replaced by $l_j - l_i$ in a basis for Λ, it follows that

$$|l_j|^2 + |l_i|^2 - 2(l_j, l_i) = |l_j - l_i|^2 \geq |l_j|^2.$$

Hence the lemma, replacing l_j by $-l_j$ if necessary. □

Given an ordered set of integer vectors $\{l_1, \ldots, l_j\}$, we will use the shorthand

$$H_j = H(l_1) \cap \ldots \cap H(l_j),$$

and for the sake of consistency, we let $H_0 = \mathbb{R}^n$. From the definition it follows that $H(u)$ always contains the ball $B(|u|/2)$, and in particular H_j contains $B(1/2)$, for any given set of integer vectors. Finally we note that if $\{l_1, \ldots, l_r\}$ is a reduced basis for some lattice Λ, then $l_{j+1} \in H_j$, $1 \leq j < r$, by Lemma 6.4. We will use the following simple fact that we state as a lemma for reference.

Lemma 6.5. Let $H \subset \mathbb{R}^m$ be a convex set containing the ball $\{x : |x| < 1/2\}$, and let $\tilde{H} = \{x : d(x, H) < 1/2\}$. Then

$$\tilde{H} \subset \tau H = \{\tau x : x \in H\},$$

where $\tau = 1 + \sqrt{2}$.

Proof: Assume $u \in \tilde{H} \setminus \bar{H}$, and let $x \in \partial H$ be the closest point to u. Thus, in particular we have

$$|x - u| < 1/2. \tag{32}$$

If u is a multiple of x then $|\tau x| > |2x| \geq |u|$ by (32), since $|x| \geq 1/2$ by our assumption. But then we are done, since τH is also convex. Thus we may assume that x and u are independent (and $m \geq 2$), and we restrict our attention to the 2 dimensional plane they span.

We may now assume that x lies on the positive x-axis, and that u lies in the first quadrant. It then follows that u is contained in the half disk of radius $1/2$, centered at x, that lies in the first quadrant. But this half disk is contained in the triangle with vertices $(0,0), (0, \tau/2)$ and $(\tau x, 0)$, as follows immediately from elementary geometry. Since this triangle is contained in τD by convexity, we are done. □

For parameters $B > A > 0$ we set, as before,

$$H_j(A,B) = \{u \in H_j \cap \mathbb{Z}^n : A \leq |u| < B\}.$$

Lemma 6.6. Let $\{l_1, \dots, l_j\}$ be a reduced basis for a lattice $\Lambda \subset \mathbb{Z}^n$, with $j < n$, and assume that $B - 1 \geq A \geq 1$, then

$$|H_j(A,B)| \ll |l_1| \cdots |l_j| (B^{n-j} - A^{n-j}),$$

where for $j = 0$, the empty product of $|l_i|$'s is set equal to 1. The implied constant depends only on n and j.

Proof: For the case $j = 0$, the lemma follows immediately from Proposition 2.3, since $B \geq A + 1$. So assume $n > j \geq 1$.

We enumerate the points in $H_j(A,B)$: $u_1, \dots, u_N$, $N = |H_j(A,B)|$. By considering the disjoint collection of balls of radius $1/2$ centered at these points, we see that

$$N < 2^n \mathcal{V}(n)^{-1} \operatorname{vol}_n(\tilde{H}), \tag{33}$$

where $\tilde{H} = \{x \in \mathbb{R}^n : d(x, H_j(A,B)) < 1/2\}$.

Let π be the projection onto the subspace, E, spanned by $l_1, \dots, l_j$. From the definition of H_j, it follows that $\pi[H_j]$ is a parallelepiped symmetric about the origin in E which contains $B(1/2)$. A simple induction shows that $\operatorname{vol}_j(\pi[H_j]) = |l_1| \cdots |l_j| / (\sin\theta_2 \cdots \sin\theta_j)$, where θ_k is the angle between l_k and the subspace spanned by $l_1, \dots, l_{k-1}$. Since the l_k's form a reduced basis, it follows that these angles and hence their sin's are bounded uniformly away from 0, so that $\operatorname{vol}_j(\pi[H_j]) \ll |l_1| \cdots |l_j|$. Now, $\pi[\tilde{H}] \subset \{x \in E : d(x, \pi[H_j]) < 1/2\}$, so by Lemma 6.5 $\pi[\tilde{H}] \subset (1 + \sqrt{2})\pi[H_j]$, and

$$\operatorname{vol}_j(\pi[\tilde{H}]) \ll \operatorname{vol}_j(\pi[H_j]) \ll |l_1| \cdots |l_j|. \tag{34}$$

For $u \in \tilde{H}$ we write $u = u_1 + u_2$ with $u_1 \in E$ and $u_2 \perp E$. If $|u_1| > A - 1/2$ then $|u_2| < ((B + 1/2)^2 - (A - 1/2)^2)^{1/2}$, otherwise $((A - 1/2)^2 - |u_1|^2)^{1/2} \leq |u_2| < ((B + 1/2)^2 - |u_2|^2))^{1/2}$. In either case, for any $u \in \pi[\tilde{H}]$

$$\operatorname{vol}_{n-j}(\pi^{-1}[u] \cap \tilde{H}) \ll (B + 1/2)^{n-j} - (A - 1/2)^{n-j} \ll B^{n-j} - A^{n-j}.$$

Combining this with (33) and (34) we have

$$N \ll \operatorname{vol}_n(\tilde{H}) = \int_{\pi[\tilde{H}]} \operatorname{vol}_{n-j}(\pi^{-1}[u] \cap \tilde{H}) \, du \ll |l_1| \cdots |l_j| (B^{n-j} - A^{n-j}).$$ □

Lemma 6.7. If $\{l_1, \ldots, l_j\}$, with $j < n$, is a reduced basis for a lattice $\Lambda \subset \mathbb{Z}^n$, $f \in \mathcal{X}$ and $B > A \geq 1$, then

$$S(f; H_j, A, B) = \sum_{H_j(A,B)} f(|l|) = O\left(|l_1| \cdots |l_j| \int_{A-1/2}^{B} f(t) t^{n-j-1}\, dt\right),$$

where for $j = 0$ the empty product before the integral is set equal to 1. The implied constant depends only on c_f, n and j.

Proof: By the preceding lemma

$$\nu_{H_j}(x) \leq c(n,j)|l_1| \cdots |l_j|((x+1)^{n-j} - x^{n-j}) < 2^{n-j} c(n,j)|l_1| \cdots |l_j| x^{n-j-1},$$

for $x \geq 1$, by the binomial formula. Since $\nu(x) = x^{n-j-1} \in \mathcal{X}$ (where c_ν depends only on n and j), the result now follows from Lemma 6.3. □

7. Remainder Terms

We come finally to the remainder terms. Each of these is given as a sum over a set of primitive lattices in $\mathbb{Z}^n$. For $\mathcal{E}_1$ and $\mathcal{E}_3$, the function being summed is explicitly given as a positive function of the lengths of a reduced basis for the lattice, one basis for each lattice. In the case of $\mathcal{E}_2$, the function is $d(\Lambda)^{-n}$, which is comparable to $(|l_1| \cdots |l_{n-1}|)^{-n}$, where $\{l_1, \ldots, l_{n-1}\}$ is a reduced basis of Λ.

The procedure we follow in all three cases is the same. We have

$$\sum_{\tilde{E}} f(|l_1|, \ldots, |l_r|) < \sum_{H_0(A_1,B_1)} \sum_{H_1(A_2,B_2)} \cdots \sum_{H_{r-1}(A_r,B_r)} f(|l_1|, \ldots, |l_r|).$$

$\tilde{E}$ is a finite set of primitive r dimensional sublattices of $\mathbb{Z}^n$, defined in terms of bounds on the lengths of their reduced basis elements, and the limits B_j and A_j are determined from this definition. In the sum on the left, the function f is evaluated at precisely one reduced basis for each lattice, and the inequality is simply a reflection of the fact that we increase our estimate when we sum over *all* possible reduced bases for each lattice.

Using the results of the previous section, we then replace each of the sums by the corresponding integral, working from the inside out. At each stage we assume that the vectors $l_1, \ldots, l_j$ that define H_j do form a reduced basis, and thus Lemma 6.7 applies. But when summing over H_j, we no longer place this restriction on l_{j+1}, and simply sum over all integer vectors in the given range. Once again, we see that this can only increase the estimate.

Proposition 7.1. For $n - 1 \geq r \geq 3$

$$\mathcal{E}_1(T, n, r) = O(T^{nr/2}).$$

Proof: Recall that

$$\mathcal{E}_1(T, n, r) = \sum_{E(T,n,r)} \prod_{1 \leq i \leq j \leq r}{}' \max(1, T/|l_i||l_j|),$$

where for each $\Lambda \in E(T, n, r)$, $\{l_1, \ldots, l_r\}$ is a fixed reduced basis for Λ, and the prime means that the factor $\max(1, T/|l_r|^2)$ is omitted from the product.

If $\{l_1, \ldots, l_r\}$ is the reduced basis of $\Lambda \in E(T, n, r)$, then $|l_j||l_{r-j+1}| \leq c_r T$, and, as always, $|l_1| \leq \ldots \leq |l_r|$. Implementing the plan outlined above, we obtain

$$\mathcal{E}_1(T, n, r) \ll \sum_{H_0(1,T^{1/2})} \sum_{H_1(A_2,B_2)} \cdots \sum_{H_{r-1}(A_r,B_r)} \prod_{1 \leq i \leq j \leq r}{}' \max(1, T/|l_i||l_j|), \tag{35}$$

where $A_1 = 1$, $A_{j+1} = |l_j|$ for $1 \le j < r$, $B_j = T^{1/2}$ for $j \le (r+1)/2$, and $B_j = T/|l_{r-j+1}|$ for $j > (r+1)/2$ (having 'absorbed' c_r in T).

The function of $|l_r|$ appearing in the innermost sum,

$$\prod_{1 \le i \le r-1} \max(1, T/|l_i||l_r|),$$

is in $\mathcal{P}$ by Proposition 6.1. We may therefore invoke Lemma 6.7 and replace the innermost sum by the integral

$$|l_1| \cdots |l_{r-1}| \int_{A_r/2}^{B_r} t_r^{n-r} \prod_{1 \le i \le r-1} \max(1, T/|l_i| t_r)\, dt_r.$$

We note that since $A_r \ge 1$, $A_r - 1/2 \ge A_r/2$, and by decreasing the lower limit of integration, we only increase the estimate. The function of $|l_{r-1}|$ appearing in the sum over H_{r-2} is

$$|l_{r-1}| \prod_{1 \le i \le r-1} \max(1, T/|l_i||l_{r-1}|) \int_{A_r/2}^{B_r} t_r^{n-r} \prod_{1 \le i \le r-1} \max(1, T/|l_i| t_r)\, dt_r.$$

This function is in $\mathcal{P}$ by Propositions 6.1 and 6.2. Thus we may apply Lemma 6.7 again and replace the sum over H_{r-2} by the integral

$$|l_1| \cdots |l_{r-2}| \int_{A_{r-1}/2}^{B_{r-1}} \int_{A_r/2}^{B_r} t_{r-1}^{n-r+2} t_r^{n-r} \prod_{r-1 \le j \le r} \max(1, T/t_{r-1} t_j) \prod_{1 \le i \le r-2} \max(1, T/|l_i| t_j)\, dt_r\, dt_{r-1}.$$

The resulting function of $|l_{r-2}|$ appearing in the next sum is also in $\mathcal{P}$, as before, and Lemma 6.7 may be invoked again. Continuing in this manner, all the sums are replaced by the corresponding integrals, resulting in the estimate

$$\mathcal{E}_1(T, n, r) \ll \int_{1/2}^{T^{1/2}} \int_{t_1/2}^{T^{1/2}} \cdots \int_{t_{r-1}/2}^{T/t_1} t_1^{n+r-2} t_2^{n+r-4} \cdots t_r^{n-r} {\prod_{1 \le i \le j \le r}}' \max(1, \frac{T}{t_i t_j})\, dt_r \ldots dt_2\, dt_1.$$

We make the substitutions $t_j = T^{1/2} u_j$ to obtain $\mathcal{E}_1(T, n, r) \ll T^{nr/2} I(n, r, T^{-1/2}/2)$, where

$$(36)\quad I(n, r, \varepsilon) = \int_{\varepsilon}^{1} \int_{u_1/2}^{1} \cdots \int_{u_{r-1}/2}^{1/u_1} u_1^{n+r-2} u_2^{n+r-4} \cdots u_r^{n-r} {\prod_{1 \le i \le j \le r}}' \max\left(1, \frac{1}{u_i u_j}\right) du_r \ldots du_1.$$

The lower limit of the j^{th} integral is $u_{j-1}/2$, for $2 \le j \le r$. The upper limit of the j^{th} integral is 1 for $1 \le j \le (r+1)/2$, and is $1/u_{r-j+1}$ for $(r+1)/2 < j \le r$. The proposition will follow from

Lemma 7.2. For $n > r \ge 3$ and $\varepsilon > 0$, $I(n, r, \varepsilon) = O(1)$.

Proof: For $u_1, \ldots, u_r$ in the range of integration, we have $u_j u_{r-j+1} \le 1$, and so $u_1 u_2 \cdots u_r \le 1$. It follows then that $I(n+1, r, \varepsilon) \le I(n, r, \varepsilon)$, since the integrand for $n+1$ is the integrand for n multiplied by $u_1 \cdots u_r$. It therefore suffices to prove the assertion of the lemma for $r = n-1$. This will be done by induction.

To begin, we have

$$I(3, 2, \varepsilon) = \int_{\varepsilon}^{1} \int_{u_1/2}^{1/u_1} du_2 du_1 < |\log \varepsilon|,$$

and

$$
\begin{aligned}
I(4,3,\varepsilon) &= \int_\varepsilon^1 \int_{u_1/2}^1 \int_{u_2/2}^{1/u_1} u_1 \max(1, 1/u_2u_3)\, du_3 du_2 du_1 \\
&< \int_\varepsilon^1 \int_{u_1/2}^1 \int_{u_2/2}^{1/u_2} u_1 u_2^{-1} u_3^{-1}\, du_3 du_2 du_1 + \int_\varepsilon^1 \int_{u_1/2}^1 \int_{1/u_2}^{2/u_1} u_1\, du_3 du_2 du_1 \\
&\ll \int_\varepsilon^1 u_1 \log^2 u_1 + 1\, du_1 = O(1).
\end{aligned}
$$

We now assume that $I(m, m-1, \varepsilon) = O(1)$ for $4 \le m < n$, and estimate

(37)

$$
I(n, n-1, \varepsilon) = \int_\varepsilon^1 \int_{u_1/2}^1 \cdots \int_{u_{n-2}/2}^{1/u_1} u_1^{2n-3} u_2^{2n-5} \cdots u_{n-1} \sideset{}{'}\prod_{1\le i\le j\le n-1} \max(1, \frac{1}{u_i u_j})\, du_{n-1} \dots du_2 du_1,
$$

beginning with the innermost integral.

Lemma 7.3. Let $u_1, \dots, u_{n-2}$ be in the range of integration of the integral above (as described in the paragraph following (36), then

$$
J = J(u_1, \dots, u_{n-2}) = \int_{u_{n-2}/2}^{1/u_1} u \prod_{1\le i\le n-2} \max(1, \frac{1}{u_1 u})\, du \ll u_1^{3-n} (\log_+ u_1)(u_2 \cdots u_{n-2})^{-1},
$$

where in this case we set $\log_+ u = \max(1, |\log u|)$.

proof: Let $k \ge 2$ be the largest index such that $u_k u_{n-2} \le 2$. If $u_{n-2}/2 \le u < 1/u_k$ then

$$
u_{n-2} u \ge 2^{-1} u_{n-3} u \ge \ldots \ge 2^{k+2-n} u_{k+1} u > 2^{k+2-n},
$$

and so

$$
\prod_{1\le i\le n-2} \max(1, 1/u_i u) \ll (u_1 \cdots u_k)^{-1} u^{-k}.
$$

By the same reasoning, when $1/u_{m+1} \le u \le 2/u_m$, then

$$
\prod_{1\le i\le n-2} \max(1, 1/u_i u) \ll (u_1 \cdots u_m)^{-1} u^{-m}.
$$

Setting $J_k = (u_1 \cdots u_k)^{-1} \int_{u_{n-2}/2}^{1/u_k} u^{1-k}\, du$, and $J_m = (u_1 \cdots u_m)^{-1} \int_{1/u_{m+1}}^{2/u_m} u^{1-m}\, du$, for $1 \le m < k$, it follows that $J \ll \sum_{m=1}^k J_m$.

We have $J_1 = u_1^{-1} \int_{1/u_2}^{2/u_1} du < 2u_1^{-2} \ll u_1^{3-n} (u_2 \cdots u_{n-2})^{-1}$, since $n \ge 5$ and $u_2 \cdots u_{n-2} \le 1$. Next,

$$
J_2 = \left\{ \begin{array}{lcl} (u_1u_2)^{-1} \int_{u_2/2}^{1/u_2} u^{-1}\, du & : & k = 2, \\ (u_1u_2)^{-1} \int_{1/u_3}^{2/u_2} u^{-1}\, du & : & k > 2. \end{array} \right\} \ll u_1^{3-n} (\log_+ u_1)(u_2 \cdots u_{n-2})^{-1},
$$

since $\log_+ u_1 \gg \log_+ u_2 \gg \log_+ u_3$. When $k > 2$ then

$$
J_k = (u_1 \cdots u_k)^{-1} \int_{u_{n-2}/2}^{1/u_k} u^{1-k}\, du < 2^{k-2} (u_1 \cdots u_k)^{-1} u_{n-2}^{2-k}.
$$

But $u_{n-2} \le 2/u_1$ and $u_{k+1} \cdots u_{n-2} \ll u_1^{k+2-n}$, so $J_k \ll u_1^{3-n} (u_2 \cdots u_{n-2})^{-1}$. By the same argument $J_m \ll u_1^{3-n} (u_2 \cdots u_{n-2})^{-1}$, for $3 \le m \le k$, and so $J \ll u_1^{3-n} (\log_+ u_1)(u_2 \cdots u_{n-2})^{-1}$. $\square$

Returning to (37), and inserting the estimate for $J(u_1, \ldots, u_{n-2})$ from the preceding lemma, we have

$$\begin{aligned} I(n,n-1,\varepsilon) &\ll \int_\varepsilon^1 \cdots \int_{u_{n-3}}^{1/u_2} u_1^n (\log_+ u_1) u_2^{2n-6} \cdots u_{n-2}^2 \prod_{1 \le i \le j \le n-2} \max(1, \frac{1}{u_i u_j})\, du_{n-2} \ldots du_2 du_1 \\ &\ll \int_\varepsilon^1 \cdots \int_{u_{n-3}}^{1/u_2} u_1 (\log_+ u_1) u_2^{2n-7} \cdots u_{n-2} \prod_{1 \le i \le j \le n-2} \max(1, \frac{1}{u_i u_j})\, du_{n-2} \ldots du_2 du_1. \end{aligned}$$

To pass from the first line to the second line, we observe that $\max(1, 1/u_1 u_j) \ll 1/u_1 u_j$, for $1 \le j \le n-2$, and that the product of these contributes a factor of $u_1^{1-n}(u_2 \cdots u_{n-2})^{-1}$.

The integrand in the last integral differs from that in the definition of $I(n-2, n-3, u_1/2)$ by the inclusion of the factor $\max(1, 1/u_{n-2}^2)$, which is different from 1 only when $u_{n-2} < 1$. This then implies that $u_j < 2^{n-j-2}$, and that $\prod_{2 \le i \le j \le n-2} \max(1, 1/u_i u_j) \ll (u_2 \cdots u_{n-2})^{2-n}$. Thus

$$I(n,n-1,\varepsilon) \ll \int_\varepsilon^1 u_1 (\log_+ u_1)(I(n-2,n-3,u_1/2) + I_1(n-2,u_1/2))\, du_1, \tag{38}$$

where for $0 < \eta < 1$

$$I_1(m,\eta) = \int_\eta^1 \int_{v_1/2}^{2^{m-2}} \cdots \int_{v_{m-2}/2}^2 \int_{v_{m-1}/2}^1 v_1^{m-2} v_2^{m-4} \cdots v_m^{-m}\, dv_m \cdots dv_2 dv_1.$$

For $m = 2$, $I_1(m,\eta) = \int_\eta^1 \int_{v_1/2}^1 v_2^{-2}\, dv_2 dv_1 < 2|\log \eta|$. If $m \ge 3$ then

$$\int_{u_{m-1}/2}^1 u_m^{-m}\, du_m \ll u_{m-1}^{1-m} \ll (u_1 \cdots u_{m-1})^{-1}.$$

Thus, after performing the first integration we see that

$$I_1(m,\eta) \ll \int_\eta^1 \int_{v_1/2}^{2^{m-2}} \cdots \int_{v_{m-2}/2}^2 v_1^{m-3} v_2^{m-5} \cdots v_{m-1}^{1-m}\, dv_{m-1} \cdots dv_2 dv_1 \ll I_1(m-1,\eta) \ll |\log \eta|,$$

by induction.

Inserting this estimate in (38), and invoking our induction hypothesis for $I(n-2, n-3, u_1/2)$, we obtain finally

$$I(n,n-1,\varepsilon) \ll \int_\varepsilon^1 u_1 \log_+^2 u_1\, du_1 = O(1).$$

This completes the proof of Lemma 7.2 and of Proposition 7.1. □

Proposition 7.4. For $n \ge 4$, $\mathcal{E}_2(T, n-1, \delta) = O(\delta^{n-2} T^{n(n-1)/2} \log T)$.

Proof: From the definition of $\mathcal{E}_2$, (25), it suffices to show that

$$S(T,\delta) = \sum_{E^\star \setminus E_\delta} d(\Lambda)^{-n} = O\left(\delta^{n-2} \log T\right),$$

where, as in §5, $E^\star = E^\star(T, n, n-1)$ and $E_\delta = E_\delta(T, n-1)$. We first remark that for a reduced basis $\{l_1, \ldots, l_{n-1}\}$ of Λ we have $d(\Lambda)^{-n} \ll |l_1|^{-n} \cdots |l_{n-1}|^{-n}$, so

$$S(T,\delta) \ll \sum_{E^\star \setminus E_\delta} |l_1|^{-n} \cdots |l_{n-1}|^{-n},$$

summing over one reduced basis for each lattice.

Now if $\Lambda \notin E_\delta$ then for some index $2 \le j \le n-1$, it follows that $|l_j| > \delta^{-1}|l_{j-1}|$. Fixing such a j let $A_j^j = \delta^{-1}|l_{j-1}|$, and for $k \ne j$ let $A_k^j = |l_{k-1}|$, with $A_1^j = 1$ always. We also define $B_1 = T^{1/2}$, and $B_k = T^{(n-1)/2}$ for $2 \le k \le n-1$. The B_k, $k \ge 2$, are larger than need be, but this will be seen to be irrelevant. (In fact we could take $B_k = (T^{(n-1)/2}(|l_1|\cdots|l_{k-1}|)^{-1})^{1/(n-k)}$, for $2 \le k \le n-1$, since $\Lambda \in E^\star$.)

With these notations it follows that if $\Lambda \in E^\star \setminus E_\delta$ then a reduced basis of Λ, $\{l_1, \ldots, l_{n-1}\}$, satisfies

$$l_1 \in H_0(A_1^j, B_1),\ l_2 \in H_1(A_2^j, B_2), \ldots,\ l_{n-1} \in H_{n-2}(A_{n-1}^j, B_{n-1}),$$

for some j between 2 and $n-1$. Hence

$$S(T,\delta) = O\left(\max_{2 \le j \le n-1} S_j(T,\delta)\right), \tag{39}$$

where

$$S_j(T,\delta) = \sum_{l_1 \in H_0(A_1^j,B_1)} |l_1|^{-n} \sum_{l_2 \in H_1(A_2^j,B_2)} |l_2|^{-n} \cdots \sum_{l_{n-1} \in H_{n-2}(A_{n-1}^j,B_{n-1})} |l_{n-1}|^{-n} \quad .$$

We now replace the multiple sum above by a multiple integral, beginning with the innermost sum. We have

$$\sum_{l_{n-1} \in H_{n-2}(A_{n-1}^j,B_{n-1})} |l_{n-1}|^{-n} \ll |l_1|\cdots|l_{n-2}| \int_{A_{n-1}^j/2}^{B_{n-1}} t_{n-1}^{1-n}\,dt_{n-1},$$

by Lemma 6.7 and Proposition 6.1. We have used, once again, the fact that $A_{n-1}^j \ge 1$, and hence $A_{n-1}^j - 1/2 \ge A_{n-1}^j/2$. Denoting by f_{n-2} the function of $|l_{n-2}|$ that appears in the next sum, we have

$$f_{n-2}(|l_{n-2}|) = \int_{A_{n-1}^j/2}^{B_{n-1}} |l_{n-2}|^{1-n} t_{n-1}^{1-n}\,dt_{n-1}. \tag{40}$$

By Proposition 6.2, $f_{n-2} \in \mathcal{P}$ since $A_{n-1}^{j/2}$ is increasing as a function of $|l_{n-2}|$, B_{n-1} is constant and the integrand and the integrand is in $\mathcal{P}$ both as a function of $|l_{n-2}|$ and as a function of t_{n-1}. Applying Lemma 6.7 to the next sum we have

$$\begin{aligned}\sum_{l_{n-2} \in H_{n-3}(A_{n-2}^j,B_{n-2})} f_{n-2}(|l_{n-2}) &\ll |l_1|\cdots|l_{n-3}| \int_{A_{n-2}^j/2}^{B_{n-2}} f_{n-2}(t_{n-2}) t_{n-2}^2\,dt_{n-2}\\ &= |l_1|\cdots|l_{n-3}| \int_{A_{n-2}^j/2}^{B_{n-2}} \int_{A_{n-1}^j/2}^{B_{n-1}} t_{n-2}^{3-n} t_{n-1}^{1-n}\,dt_{n-1}dt_{n-2}.\end{aligned}$$

Note that the inner integral above is in $\mathcal{P}$ as a function of t_{n-2}, by the same argument we used for f_{n-2}. Writing f_{n-3} for the function of $|l_{n-3}|$ that now appears in the sum over H_{n-4}, we have

$$f_{n-3}(|l_{n-3}|) = \int_{A_{n-2}^j/2}^{B_{n-2}} \int_{A_{n-1}^j/2}^{B_{n-1}} |l_{n-3}|^{2-n} t_{n-2}^{3-n} t_{n-1}^{1-n}\,dt_{n-1}dt_{n-2}.$$

The integrand of the outer integral above is in $\mathcal{P}$, both as a function of $|l_{n-3}|$ and of t_{n-2}. Also, $A_{n-2}^j/2$ is an increasing function of $|l_{n-3}|$ and B_{n-2} is constant, and so it follows, as before, that we may apply Lemma 6.7 to this sum as well. Proceeding in this manner, we replace all the sums by integrals, obtaining the estimate

$$S_j(T,\delta) \ll \int_{1/2}^{\sqrt{T}} \int_{A_2^j/2}^{B_2} \cdots \int_{A_{n-1}^j/2}^{B_{n-1}} t_1^{n-3} t_2^{n-5} \cdots t_{n-1}^{1-n} dt_{n-1} \ldots dt_2\, dt_1, \tag{41}$$

where the exponent of t_k is $n-2k-1$. Note that $|l_{k-1}|$ is replaced by t_{k-1} in the definition of A_k^j.

Let $e_{n-1} = 1-n$, and for $1 \le k \le n-2$ let $e_k = n-2k+e_{k+1}$, then e_k is precisely the exponent of t_k after performing the first $n-k-1$ integrations (from the inside out). It is easy to see that $e_k = (n-k)(1-k)-1$, and in particular, $e_k \le 1-n \le -3$ for $k > 1$, (whence the B_k's for these k's are irrelevant), and $e_1 = -1$. With these observations, (41) yields $S_j(T,\delta) = O(\delta^{-(e_j+1)} \log T)$. Since $1 > \delta$, this expression is maximized when $j = 2$ or $j = n-1$, for which $e_j = 1-n$. Inserting this in (39) completes the proof. □

Proposition 7.5. For $n \ge 4$

$$\mathcal{E}_3(T, n-1, \delta) = O\left(\delta^{(3-n)(n^2+2)/3} T^{n(n-1)/2}\right).$$

Proof: Recall that by Lemma 4.3, for $r = n-1$,

$$N_2(T, S(\Lambda)) = O\left(\frac{T^{(n-1)(n-2)/2}}{|l_1|^{(n-1)(n-2)}} \log_+(T/|l_1|^2)\right),$$

where l_1 is the shortest element in a fixed reduced basis for Λ. ¿From the definition, (26), of $\mathcal{E}_3$ it then follows that

$$\begin{aligned}
\mathcal{E}_3 &\ll T^{(n-1)(n-2)/2} \sum_{E_\delta} \frac{\log_+(T/|l_1|^2)}{|l_1|^{(n-1)(n-2)}} \\
&\ll T^{(n-1)(n-2)/2} \sum_{H_0(1,T^{1/2})} \sum_{H_1(A_2,B_2)} \cdots \sum_{H_{n-2}(A_{n-1},B_{n-1})} \frac{\log_+(T/|l_1|^2)}{|l_1|^{(n-1)(n-2)}}.
\end{aligned}$$

Here, $A_j = |l_{j-1}|$ and $B_j = \delta^{-1}|l_{j-1}|$, for $j = 2, \ldots, n-1$. Once again the sum over lattices has been replaced by a sum over reduced bases, and the definition of B_j expresses the fact that we are summing over lattices in E_δ.

For the innermost sum we have

$$\sum_{H_{n-2}(A_{n-1},B_{n-1})} \frac{\log_+(T/|l_1|^2)}{|l_1|^{(n-1)(n-2)}} \ll \frac{\log_+(T/|l_1|^2)}{|l_1|^{(n-1)(n-2)}} |l_1| \cdots |l_{n-2}| \int_{A_{n-1}/2}^{B_{n-1}} t_{n-1}\, dt_{n-1},$$

by Lemma 6.7 and Proposition 6.1 (a). As a function of $|l_{n-2}|$ the integral above is in $\mathcal{P}$, since the integrand satisfies the conditions of Proposition 6.2, and $B_{n-1} > A_{n-1}$, so $B_{n-1} > 2A_{n-1}/2$. Thus we may apply lemma 6.7 to the second sum as well. We continue in this manner to replace sums by integrals, (using arguments similar to those leading up to (41)), until we reach the first sum where we have

$$\mathcal{E}_3 \ll T^{(n-1)(n-2)/2} \sum_{H_0(1,T^{1/2})} \log_+(T/|l_1|^2)|l_1|^{-n^2+4n-4} \int_{A_2/2}^{B_2} \cdots \int_{A_{n-1}/2}^{B_{n-1}} t_2^{2n-5} \cdots t_{n-1}\, dt_{n-1} \ldots dt_2.$$

As a function of $|l_1|$, the multiple integral above is in $\mathcal{P}$ by the inductive argument invoking Proposition 6.2. Also, $\log_+(T/x^2)x^{-n^2+4n-4} \in \mathcal{P}$ by Proposition 6.1. Thus Lemma 6.7 may be applied to this last sum as well, giving

(42)
$$\mathcal{E}_3 \ll T^{(n-1)(n-2)/2} \int_{1/2}^{\sqrt{T}} \int_{t_1/2}^{\delta^{-1}t_1} \cdots \int_{t_{n-2}/2}^{\delta^{-1}t_{n-2}} t_1^{-n^2+5n-5} \log_+(T/t_1^2) t_2^{2n-5} \cdots t_{n-1}\, dt_{n-1} \ldots dt_2\, dt_1,$$

where the exponent of t_j is $2(n-j)-1$ for $1 < j \le n-1$. As before, the definitions of A_j and B_j have been changed by replacing $|l_{j-1}|$ with t_{j-1}.

We make the substitutions $t_j = \delta^{1-j}T^{1/2}u_j$ in (42), for $1 \le j \le n-1$. Thus

$$t_j^{2(n-j)-1}\,dt_j = \delta^{2(1-j)(n-j)}T^{n-j}u_j^{2(n-j)-1}\,du_j,$$

for $j = 2\ldots n-1$, and

$$t_1^{-n^2+5n-5}\log_+(T/t_1^2)\,dt_1 = T^{(-n^2+5n-4)/2}u_1^{-n^2+5n-5}\log_+ u_1\,du_1.$$

The *integral* is transformed to

$$T^{\alpha(n)}\delta^{-\beta(n)}\int_{T^{-1/2}/2}^{1}\int_{a_2}^{u_1}\cdots\int_{a_{n-1}}^{u_{n-2}} 2u_1^{-n^2+5n-5}\log_+ u_1 u_2^{2n-5}\cdots u_{n-1}\,du_{n-1}\ldots du_2\,du_1,$$

where the a_j's are all positive,

$$\alpha(n) = (-n^2+5n-4)/2 + \sum_{j=2}^{n-1}(n-j) = n-1,$$

and

$$\beta(n) = 2\sum_{j=2}^{n-1}(n-j)(j-1) = (n-3)(n^2+2)/3.$$

Performing the first $n-2$ integrations reduces this to

$$T^{n-1}\delta^{-\beta(n)}\int_{T^{-1/2}/2}^{1} 2u_1^{n-3}\log_+ u_1\,du_1 \ll T^{n-1}\delta^{-\beta(n)},$$

since the integral converges for $n \ge 4$. Inserting this back in (42) completes the proof. $\square$

Corollary 7.6. Choosing $\delta = (\log T)^{-\frac{1}{\beta(n)+n-2}}$, we have for $n \ge 4$

$$\mathcal{E}_2(T, n-1, \delta) + \mathcal{E}_3(T, n-1, \delta) = O\left(T^{n(n-1)/2}(\log T)^{1-\frac{n-2}{\beta(n)+n-2}}\right).$$

Proof: This follows from Proposition 7.4 and Proposition 7.5, where $\beta(n)$ is the cubic polynomial of n defined in the latter. $\square$

The authors thank the referee for the substantial time that he put into this paper, and for the many suggestions and corrections that he made.

References

1. J.W.S. Cassels, An Introduction to the Geometry of Numbers, Springer Verlag, (1971).
2. W. Duke, Z. Rudnick, P. Sarnak, *The density of integer points on affine homogeneous varieties*, Duke Math J. **71** (1993), 143–180.
3. A. Eskin, C. McMullen, *Mixing, counting and equidistribution in Lie groups*, Duke Math. J. **71** (1993), 181–210.
4. A. Eskin, S. Mozes, N. Shah, *Unipotent flows and counting lattice points on homogeneous varieties*, preprint.
5. P.M. Gruber and C.G. Lekkerkerker, Geometry of Numbers, 2^{nd} Ed., North–Holland (1987).
6. G.H. Hardy, E.M. Wright, An Introduction to the Theory of Numbers, 4^{th} ed. Oxford (1960).
7. Y.R. Katznelson, *Singular matrices and a uniform bound for congruence groups of $SL_n(\mathbb{Z})$*, Duke Math. J. **69** (1993), 121–136.
8. Y.R. Katznelson, *Matrices of fixed rank*, to appear Proc. of the A.M.S.
9. K. Mahler, *Ein Übertragunzprinzip für konvexe Körper*, Časopis Pěst. Mat. Fyz. **68** (1939), 93–102.
10. W. Schmidt, *Asymptotic formulae for point lattices of bounded determinant and subspaces of bounded height,* Duke Math J. **35** (1968), 327–339.
11. W. Schmidt, *Northcott's theorem on heights, II The quadratic case*, Preprint.
12. C.L. Siegel, Lectures on the Geometry of Numbers, Springer Verlag, 1988.

Alex Eskin: Institute for Advanced Studies, Princeton, NJ, USA
E-mail address: eskin@math.ias.edu

Yonatan Katznelson: Institut des Hautes Études Scientifiques, Bûres sur Yvette, France
E-mail address: yorick@cats.ucsc.edu