A METRICAL THEOREM IN GEOMETRY OF NUMBERS

BY

WOLFGANG SCHMIDT

Introduction. If S is a pointset in R_n, $n>1$, then we write

$$L(S)$$

for the number of lattice-points in S. Here, and throughout this paper, a lattice-point is a point with integral coordinates. If S is a Borel set of finite volume $V(S)$, one would expect that $L(S)$ is of about the same order of magnitude as $V(S)$. Hence we define the "discrepancy" $D(S)$ by

$$D(S) = |L(S)V(S)^{-1} - 1|. \tag{1}$$

As a companion for $L(S)$, we introduce

$$P(S),$$

the number of primitive lattice-points in S. (A lattice-point is primitive, if its coordinates are relatively prime.) We put

$$E(S) = |P(S)\zeta(n)V(S)^{-1} - 1|. \tag{2}$$

Next, let Φ be a family of Borel sets with finite volumes, such that

(i) If $S \in \Phi$, $T \in \Phi$, then either $S \subseteq T$ or $T \subseteq S$.

(ii) There exist $S \in \Phi$ with arbitrarily large $V(S)$.

Finally, throughout this paper, $\psi(s)$, $s \geqq 0$, should be a positive, nondecreasing function, such that $\int_0^\infty \psi(s)^{-1} ds$ exists.

THEOREM 1. *Suppose $n>2$. Then for almost every linear transformation A ("almost every" in the sense of the n^2-dimensional euclidean metric induced by matrix-representation for A),*

$$D(AS) = O(V^{-1/2} \log V \, \psi^{1/2}(\log V)), \tag{3}$$

$$E(AS) = O(V^{-1/2} \log V \, \psi^{1/2}(\log V)) \tag{4}$$

for $S \in \Phi$.

More explicitly, for almost every A there exist constants $c_1(A)$, $c_2(A)$, such that

$$D(AS) \leqq c_1(A) V^{-1/2} \log V \psi^{1/2}(\log V),$$

whenever $V(S) \geqq c_2(A)$ and $S \in \phi$.

In R_2 our results are a little weaker:

Received by the editors March 13, 1959.

THEOREM 2. *Let $n=2$. Then for almost every linear transformation A*

$$D(AS) = O(V^{-1/2} \log^2 V \, \psi^{1/2}(\log V)), \tag{5}$$

$$E(AS) = O(V^{-1/2} \log^2 V \, \psi^{1/2}(\log V)) \tag{6}$$

for $S \in \phi$.

Theorem 1 and Theorem 2 yield the

COROLLARY. *Suppose S is a Borel set with infinite volume in R_n, $n>1$. Then $P(AS)$ is infinite for almost all A.*

This corollary was first proved by C. A. Rogers [8] in case $n>2$. In fact, for $n>2$ our results will follow very easily from methods of Rogers [8] and Cassels [2]. The proof of Theorem 2 is more difficult. It was shown by Lekkerkerker [5], that the corollary does not hold in R_1. Therefore Theorem 1 has no analogon in R_1.

For the proof of Theorem 2 we shall need the 2-dimensional case of

THEOREM 3. *Let S be a Lebesgue-measurable set in R_n with characteristic function $\rho(X)$ and volume V. By $\|X_1, \cdots, X_n\|$ we denote the absolute value of the $n \times n$-determinant, whose row-vectors are $X_1, \cdots, X_n$. Then*

$$\int \cdots \int \rho(X_1) \cdots \rho(X_n) \chi(\|X_1, \cdots, X_n\|) dX_1 \cdots dX_n \leqq n2^n \int_0^\infty \chi(t) dt V^{n-1}$$

for every non-negative, nonincreasing function $\chi(t)$, defined for $t \geqq 0$, whose integral $\int_0^\infty \chi(t)dt$ converges.

Probably, it would be possible to prove Theorem 3, with $n2^n$ replaced by a smaller constant if $n>2$, using Steiner symmetrization and the methods of §§22-26 in [1]. Our proof will be shorter.

We shall first prove Theorem 1, then Theorem 3, finally Theorem 2.

1. We have to use the invariant measure $\mu(A)$ over the space of matrices A of determinant 1, defined by Siegel [10] and further developed in [6; 8] and [9]. $\mu(A)$ is normalized such that

$$\int_F d\mu(A) = 1,$$

where F is a fundamental region of matrices with respect to the subgroup of unimodular matrices.

We mention the following equations, proved in the papers cited above: If S is a Borel set not containing the origin 0, with characteristic function $\rho(X)$, volume V, then

$$\int_F \sum_g \rho(Ag) d\mu(A) = V, \tag{7}$$

$$\int_F \sum \begin{bmatrix} g \\ \text{primitive} \end{bmatrix} \rho(Ag)d\mu(A) = V\zeta(n)^{-1}. \tag{8}$$

The summation is over lattice-points g. Further, if $n>2$,

$$\int_F \sum \begin{bmatrix} g_1, g_2 \\ \text{lin. indep.} \end{bmatrix} \rho(Ag_1)\rho(Ag_2)d\mu(A) = V^2, \tag{9}$$

and

$$\int_F \sum \begin{bmatrix} g_1, g_2 \\ \text{lin. indep. primitive} \end{bmatrix} \rho(Ag_1)\rho(Ag_2)d\mu(A) = V^2\zeta(n)^{-2}. \tag{10}$$

We observe, again for $n>2$,

$$\begin{aligned} &\int_F \sum \begin{bmatrix} g_1, g_2 \\ \text{lin. dep.} \end{bmatrix} \rho(Ag_1)\rho(Ag_2)d\mu(A) \\ &= \frac{1}{2}\int_F \sum_{p,q} \sum \begin{bmatrix} g \\ \text{primitive} \end{bmatrix} \rho(Apg)\rho(Aqg)d\mu(A) \\ &\leqq \sum_{q\neq 0} \sum_{p;|p|\leq|q|} \int_F \sum \begin{bmatrix} g \\ \text{primitive} \end{bmatrix} \rho(Aqg)d\mu(A) \\ &< V\sum_{q\neq 0} \sum_{p;|p|\leq|q|} |q|^{-n} < c_3V, \end{aligned}$$

where c_3 is an absolute constant. Using this result, as well as (7), (8), (9) and (10), we obtain

$$\int_F \left(\sum_g \rho(Ag) - V \right)^2 d\mu(A) < c_3V$$

and

$$\int_F \left(\sum \begin{bmatrix} g \\ \text{primitive} \end{bmatrix} \rho(Ag) - V\zeta(n)^{-1} \right)^2 d\mu(A) < c_3V.$$

The second inequality was proved in [8] in the same way.

We also mention the following integral identity, established by Rogers [8]:

$$\int_C \sigma(A)dA = c_4(n) \int_0^1 \nu^{n-1} \left\{ \int_F \sigma(\nu^{1/n}A)d\mu(A) \right\} d\nu. \tag{11}$$

Here, C is the cone of matrices A which satisfy $\lambda A \in F$ for some $\lambda \geqq 1$. dA is the euclidean volume element in the n^2-dimensional space of matrices, and

$\sigma(A)$ a matrix-function integrable with respect to dA. $c_4(n)$ is a constant depending on n only. We obtain immediately

$$(12)\qquad \int_C \Big(\sum_g \rho(Ag) - V\|A\|^{-1}\Big)^2 dA < c_5(n)V \int_0^1 \nu^{n-2}d\nu < c_5(n)V,$$

$$(13)\qquad \int_C \Big(\sum \begin{bmatrix} g \\ \text{primitive} \end{bmatrix} \rho(Ag) - V\zeta(n)^{-1}\|A\|^{-1}\Big)^2 dA < c_5(n)V,$$

where $\|A\|$ is the absolute value of the determinant of A.

2. Lemma 1. *To every set* Φ *satisfying* (i), (ii), *there exists a set* $\Psi \supseteq \Phi$, *which satisfies* (i), (ii) *and*

(iii) *To every real number* $V \geqq 0$ *there exists a* $S \in \Psi$ *satisfying* $V(S) = V$.

Proof. Write $\alpha(\Phi)$ for the set of numbers $V \geqq 0$ such that there exists a $S \in \Phi$ with $V(S) = V$. Then (iii) states that $\alpha(\Psi)$ consists of every $V \geqq 0$. We first show: There exists a $X \supseteq \Phi$ satisfying (i), (ii) and

(iiia) $\alpha(X)$ is closed.

For let V be a limit point of $\alpha(\Phi)$. We first assume there exist $V_1 \geqq V_2 \cdots$, $V_j \in \alpha(\Phi)$, $\lim V_j = V$. Then there exist $S_1 \supseteq S_2 \supseteq \cdots$, $S_j \in \Phi$, $V(S_j) = V_j$. We take $S = S(V) = \bigcap_{j=1}^{\infty} S_j$. Then S is again a Borel set, and $V(S) = V$. If, however, there exists no sequence $V_1 \geqq V_2 \geqq \cdots$ of the required form, then there exists a sequence $V_1 \leqq V_2 \leqq, \cdots$, $V_j \in \alpha(\Phi)$, $\lim V_j = V$ and we can proceed similarly, taking $S = \bigcup_{j=1}^{\infty} S_j$. If we take χ as the union of Φ and the sets S just constructed, then X satisfies all required conditions.

It remains to show that there exists a $\Psi \supseteq X$, satisfying (i), (ii), and (iii). Suppose $V \notin \alpha(X)$. Then there exist(1) $V_1 \in \alpha(X)$, $V_2 \in \alpha(X)$, $V_1 < V < V_2$, such that no point in the open interval (V_1, V_2) belongs to $\alpha(X)$. There exist S_1, S_2 with $V(S_1) = V_1$, $V(S_2) = V_2$. We write $S_2 - S_1$ for the set of all $X \in S_2$, $X \notin S_1$ and $(S_2 - S_1)_t$ for the set of all $X \in S_2 - S_1$ with $|X| \leqq t$. Then $V(S_1 \cup (S_2 - S_1)_t)$ is a continuous function of t which equals $V(S_1)$ when $t = 0$ and approaches $V(S_2)$ when $t \to \infty$. Hence there exists a t_0 such that $V(S_1 \cup (S_2 - S_1)_{t_0}) = V$. We introduce

$$S(V) = S_1 \cup (S_2 - S_1)_{t_0}$$

and take Ψ to be the union of X and all sets $S(V)$.

3. According to Lemma 1, we may assume that Φ satisfies (i), (ii) and (iii). Hence for every positive integer N, we may pick a $S \in \Phi$ with $V(S) = N$ and denote it by $S(N)$. Write $\rho_N(X)$ for the characteristic function of $S(N)$ and

$$_{N_1}\rho_{N_2}(X)$$

for the characteristic function of $S(N_2) - S(N_1)$. Finally, we introduce, following Cassels [2],

(1) We may assume that the null set belongs to X, such that $0 \in \alpha(X)$.

$$R_N(A) = \sum_g \rho_N(Ag) - N\|A\|^{-1},$$

$$S_N(A) = \sum \left[\begin{matrix} g \\ \text{primitive}\end{matrix}\right] \rho_N(Ag) - N\zeta(n)^{-1}\|A\|^{-1},$$

$$_{N_1}R_{N_2}(A) = \sum_g {}_{N_1}\rho_{N_2}(Ag) - (N_2 - N_1)\|A\|^{-1},$$

$$_{N_1}S_{N_2}(A) = \sum \left[\begin{matrix} g \\ \text{primitive}\end{matrix}\right] {}_{N_1}\rho_{N_2}(Ag) - (N_2 - N_1)\zeta(n)^{-1}\|A\|^{-1}.$$

LEMMA 2. *Let T be a positive integer and K_T the set of all pairs of integers N_1, N_2 of the type $0 \leqq N_1 < N_2 \leqq 2^T$, $N_1 = u2^t$, $N_2 = (u+1)2^t$, for integers u and $t \geqq 0$. Then*

$$\sum_{(N_1,N_2)\in K_T} \int_C {}_{N_1}R^2_{N_2}(A)dA \leqq c_5(n)(T+1)2^T. \tag{14}$$

Proof. (12) yields

$$\int_C {}_{N_1}R^2_{N_2}(A)dA \leqq c_5(n)(N_2 - N_1).$$

Each value of $N_2 - N_1 = 2^t$ $(0 \leqq t \leqq T)$ occurs 2^{T-t} times. Hence

$$\sum_{(N_1,N_2)\in K_T} (N_2 - N_1) = (T+1)2^T.$$

LEMMA 3. *For all sufficiently large T there is a subset B_T of C of measure*

$$\int_{B_T} dA \leqq c_5(n)\psi^{-1}(T\log 2 - 1), \tag{15}$$

such that

$$R^2_N(A) \leqq T(T+1)2^T\psi(T\log 2 - 1) \tag{16}$$

for every $N \leqq 2^T$ and all $A \in C$, but $A \notin B_T$.

Proof. If we take B_T to be the set consisting of all $A \in C$ for which it is not true that

$$\sum_{(N_1,N_2)\in K_T} {}_{N_1}R^2_{N_2}(A) \leqq (T+1)2^T\psi(T\log 2 - 1), \tag{16a}$$

then (15) follows immediately from Lemma 2.

Assume $N \leqq 2^T$, $A \in C$ but not in B_T. The interval $[0, N)$ can be expressed as a union of at most T intervals of the type $[N_1, N_2)$, where $(N_1, N_2) \in K_T$. Thus

$$R_N(A) = \sum_{N_1} R_{N_2}(A),$$

where the sum is over at most T pairs $(N_1, N_2) \in K_T$. Using (16a) and Cauchy's inequality, we obtain

$$R_N^2(A) \leqq T(T+1)2^T \psi(T \log 2 - 1).$$

4. **Proof of Theorem** 1. The set of $A \in C$ belonging to B_T has measure at most $c_5(n)\psi^{-1}(T \log 2 - 1)$. Since

$$\sum_{T=1}^{\infty} \psi^{-1}(T \log 2 - 1)$$

is convergent, there is a $T_0 = T_0(A)$ for almost every A such that $A \notin B_T$ for $T \geqq T_0$. Assume $N \geqq N_0 = 2^{T_0}$ and choose T such that $2^{T-1} \leqq N < 2^T$. Then, by Lemma 3,

$$\begin{aligned} R_N^2(A) &< T(T+1)2^T \psi(T \log 2 - 1) \\ &< \left(\frac{\log N}{\log 2} + 1\right)\left(\frac{\log N}{\log 2} + 2\right) 2N\psi(\log N) \\ &< c_6 N \log^2 N \psi(\log N) \end{aligned}$$

for almost every A. Thus,

$$R_N(A) = O(N^{1/2} \log N \psi^{1/2}(\log N)) \tag{17}$$

for almost every $A \in C$. Since every $\|A\| \leqq 1$ is of the form $A = A'U$, where $A' \in C$, U unimodular, and since the unimodular matrices are enumerable, (17) holds for almost every $\|A\| \leqq 1$. Applying a linear transformation we see that (17) holds for almost every A satisfying $\|A\| \leqq c$, where c is an arbitrary constant. Hence (17) holds for almost every A generally.

Comparing the definition of $R_N(A)$ and $D(A^{-1}S(N))$, it is immediate that (17) implies

$$D(A^{-1}S(N)) = O(N^{-1/2} \log N \psi^{1/2}(\log N))$$

and therefore

$$D(AS(N)) = O(N^{-1/2} \log N \psi^{1/2}(\log N)) \tag{18}$$

for almost all A.

Now let S be a set in Φ and $N \leqq V(S) < N+1$. Then

$$L(AS(N)) - (N+1) \leqq L(AS) - V(S) \leqq L(AS(N+1)) - N$$

and

$$D(AS) \leqq \max\left\{ \left| L(AS(N+1)) - N \right| N^{-1}, \left| L(AS(N)) - (N+1) \right| N^{-1} \right\}.$$

Since, by (18), both terms on the right hand side of this relation are $O(N^{-1/2}\cdots)$ for almost all A, (3) follows. The proof of (4) is analogous. This completes the proof of Theorem 1.

5. For every point $X(x^{(1)}, \cdots, x^{(n)})$ in R_n we denote by $\overline{X}$ the point $\overline{X}(x^{(1)}, \cdots, x^{(n-1)})$ in R_{n-1}. We write

$$\|\overline{X}_1, \cdots, \overline{X}_{n-1}\|$$

for the absolute value of the $(n-1)\times(n-1)$-determinant with row vectors $\overline{X}_1, \cdots, \overline{X}_{n-1}$. We put

$$h_k(X_1, \cdots, X_n) = \|\overline{X}_1, \cdots, \overline{X}_{k-1}, \overline{X}_{k+1}, \cdots \overline{X}_n\|, \quad (1 \leqq k \leqq n)$$

and define H_k $(1\leqq k\leqq n)$ to be the subdomain of those $(X_1, \cdots, X_n)$ in the n^2-dimensional space of n-tuples of points, for which

$$h_k(X_1, \cdots X_n) \geqq h_i(X_1, \cdots, X_n), \qquad (1 \leqq i \leqq n).$$

By H or tH we denote the domain of n-tuples $(X_1, \cdots, X_n)$ satisfying

$$\|X_1, \cdots, X_n\| \leqq 1 \quad \text{or} \quad \|X_1, \cdots, X_n\| \leqq t$$

respectively. Finally, $H\cap H_1(X_2, \cdots, X_n)$ should, for given $X_2, \cdots, X_n$, be the set of all X_1 satisfying

$$(X_1, \cdots, X_n) \in H \cap H_1.$$

LEMMA 4. *$H\cap H_1(X_2, \cdots, X_n)$ is a parallelepipedon of volume 2^n, if* $\|\overline{X}_2, \cdots, \overline{X}_n\|>0$.

Proof. The lemma is true, if $\overline{X}_2, \cdots, \overline{X}_n$ are points on the $x^{(1)}, \cdots, x^{(n-1)}$-axis respectively. The general case follows after applying a linear transformation of determinant 1.

LEMMA 5.

$$\int\underset{tH}{\cdots}\int \rho(X_1)\cdots\rho(X_n)dX_1\cdots dX_n \leqq n2^n tV^{n-1}.$$

Proof. First, we observe

$$\int\underset{H}{\cdots}\int \rho(X_1)\cdots\rho(X_n)dX_1\cdots dX_n = \sum_{k=1}^{n}\int\underset{H\cap H_k}{\cdots}\int \rho(X_1)\cdots\rho(X_n)dX_1\cdots dX_n.$$

Next,

$$\int\underset{H\cap H_1}{\cdots}\int \rho(X_1)\cdots\rho(X_n)dX_1\cdots dX_n \leqq 2^n\int\cdots\int \rho(X_2)\cdots\rho(X_n)dX_2\cdots dX_n = 2^nV^{n-1}.$$

We obtain the same estimate for integrals over $H \cap H_k$. This proves Lemma 5 for $t=1$. The general case follows after a linear transformation.

Proof of Theorem 3. Using partial integration and Lemma 5, we obtain

$$\int \cdots \int \rho(X_1) \cdots \rho(X_n) \chi(\|X_1, \cdots, X_n\|) dX_1 \cdots dX_n$$

$$= -\int_0^\infty \left\{ \int \underset{tH}{\cdots} \int \rho(X_1) \cdots \rho(X_n) dX_1 \cdots dX_n \right\} d\chi(t)$$

$$\leqq -\int_0^\infty n 2^n V^{n-1} t d\chi(t) = n 2^n V^{n-1} \int_0^\infty \chi(t) dt.$$

6. From now on we restrict ourselves to 2 dimensions. $X, Y, \cdots$ are points in R_2, $g, h, \cdots$ are lattice-points in R_2. We are particularly interested in ordered pairs (g_1, g_2) of lattice-points. Two pairs (g_1, g_2) are called equivalent, if there exists a proper unimodular transformation U with $Ug_i = h_i$ $(i=1, 2)$. The determinant $|g_1, g_2|$ of a pair of lattice-points depends only on the class E of pairs to which it belongs. We denote this determinant by $d(E)$. Furthermore, if (g_1, g_2) and (h_1, h_2) belong to the same class, then either both g_1, h_1 are primitive, or neither one. The same holds for g_2, h_2. We call a class E primitive, if both g_1, g_2 are primitive for every $(g_1, g_2) \in E$. By $\sigma(k)$ we denote the sum of the (positive) divisors of k, by $\varphi(k)$ the Euler φ-function.

LEMMA 6. *Assume $k \neq 0$. There exist $\sigma(|k|)$ classes E with $d(E)=k$ and $\varphi(|k|)$ primitive classes with $d(E)=k$.*

Proof. To every E with $d(E)=k$ there belongs a pair $(g_1, g_2) \in E$ of the type

$$g_1(n, 0), \; g_2(m, k/n), \tag{19}$$

where $n>0$ is a divisor of k and $0 \leqq m < |k|/n$. No two pairs (19) are equivalent. Hence there exist

$$\sum_{n/|k|} |k|/n = \sum_{n/|k|} n = \sigma(|k|)$$

equivalence classes E with $d(E)=k$.

Of all the point pairs (19) only those are primitive where $n=1$ and m, k are relatively prime. This gives $\varphi(|k|)$ classes.

7. We define $\tau(t), \omega(t), t \geqq 0$ by

$$\tau(t) = t \sum_{k \geqq t} \sigma(k) k^{-3},$$

$$\omega(t) = t \sum_{k \geqq t} \varphi(k) k^{-3}.$$

LEMMA 7.

$$\tau(t) = \zeta(2) + O(t^{-1}\log t),$$
$$\omega(t) = \zeta(2)^{-1} + O(t^{-1}\log t).$$

Proof. According to [4, Theorem 324], we have

$$\sum_{t\leq k\leq s} \sigma(k) = (s^2 - t^2)\zeta(2)/2 + O(s\log s)$$

and therefore, using Theorem 421 of [4],

$$\sum_{t\leq k} \sigma(k)k^{-3} = \int_t^\infty (s^2 - t^2)3s^{-4}\zeta(2)/2ds + O\left(\int_t^\infty s\log s\, s^{-4}ds\right)$$
$$= t^{-1}\zeta(2) + O(t^{-2}\log t).$$

The second equation is proved in the same way, but with Theorem 324 of [4] replaced by Theorem 330.

REMARK. By using a result announced in [7], namely that

$$\sum_{1\leq k\leq s} \sigma(k) = s^2\zeta(2)/2 + O(s(\log s \log\log s)^{3/4})$$

instead of Theorem 324 of [4], it would be possible to improve the first equation of Lemma 7, and finally to improve equation (5) of Theorem 2 so that $\log^2 V$ is replaced by $\log^{15/8} V \log\log^{3/8} V$.

8. If $X(x^{(1)}, x^{(2)})$ is a point in R_2 and x real, then xX is the point with coordinates $xx^{(1)}, xx^{(2)}$. With every point $X\neq 0$ we associate a point $\tilde{X}$ satisfying $|X, \tilde{X}| = 1$.

As an analogue to (10) for the R_2 we mention

LEMMA 8. *If S is a Borel set with characteristic function $\rho(X)$ and E an equivalence class with $d(E)=k$, then*

$$\int_F \sum [(g_1, g_2) \in E]\rho(Ag_1)\rho(Ag_2)d\mu(A)$$
$$= |k|^{-1}\zeta(2)^{-1}\int\left\{\int \rho(X)\rho(k\tilde{X} + xX)dx\right\}dX.$$

Proof. Lemma 8 was proved as Satz 3 in [9]. The integral on the right hand side of Lemma 8 is 3-dimensional. Integrating dx first we note that it does not depend on a particular choice of $\tilde{X}$.

As an immediate consequence of Lemmas 6, 8 we obtain

$$\int_F \begin{bmatrix} g_1, g_2 \\ \|g_1, g_2\|\neq 0\end{bmatrix}\rho(Ag_1)\rho(Ag_2)d\mu(A)$$
$$(20)\qquad = \sum_{k\neq 0}\sigma(|k|)|k|^{-1}\zeta(2)^{-1}\int\left\{\int\rho(X)\rho(k\tilde{X} + xX)dx\right\}dX.$$

Instead of the volume element dA as used in §1, we now define the volume element $dm(A)$ by

$$\int_C \sigma(A)dm(A) = \int_0^1 \left\{ \int_F \sigma(\nu^{1/2}A)d\mu(A) \right\} d\nu. \tag{21}$$

Comparing this definition with (11), we see that $dm(A)$ and dA are equivalent [3, p. 126].

LEMMA 9.

$$I_1 = \int_C \|A\| \sum_g \rho(Ag)dm(A) = V,$$

$$I_1^* = \int_C \|A\| \sum \begin{bmatrix} g \\ \text{primitive} \end{bmatrix} \rho(Ag)dm(A) = V\zeta(2)^{-1}.$$

Proof. Using (7), which is also true for $n=2$, we have

$$I_1 = \int_0^1 \nu(V\nu^{-1})d\nu = V.$$

The second equation follows similarly.

9. Now we come to the crucial lemma of the proof of Theorem 2:

LEMMA 10. *Defining* I_2, I_2^* *by*

$$I_2 = \int_C \|A\|^2 \sum \begin{bmatrix} g_1, g_2 \\ \|g_1, g_2\| \neq 0 \end{bmatrix} \rho(Ag_1)\rho(Ag_2)dm(A),$$

$$I_2^* = \int_C \|A\|^2 \sum \begin{bmatrix} g_1, g_2 \\ \|g_1, g_2\| \neq 0 \\ \text{primitive} \end{bmatrix} \rho(Ag_1)\rho(Ag_2)dm(A)$$

and writing lg V *as an abbreviation for* $\max(1, \log V)$, *we have*

$$|I_2 - V^2| \leqq c_7 V \lg^2 V, \tag{22}$$

$$|I_2^* - V^2\zeta(2)^{-1}| \leqq c_8 V \lg^2 V. \tag{23}$$

Proof. Combining (20) and (21) we obtain

$$I_2 = \zeta(2)^{-1} \int_0^1 \nu^2 \left\{ \sum_{k\neq 0} \sigma(|k|)|k|^{-1} \int \left\{ \int \rho(\nu^{1/2}X)\rho(k\nu^{1/2}\tilde{X} + x\nu^{1/2}X)dx \right\} dX \right\} d\nu.$$

Substituting $\nu^{1/2}X = Y$ we have $dX = \nu^{-1}dY$ and may assume $\nu^{-1/2}\tilde{X} = \tilde{Y}$. Next, we write $t = k\nu$ and note that if we integrate $\int_{-\infty}^{\infty} dt$, we have to replace the sum $\sum_{k\neq 0}$ by $\sum_{k \geq |t|}$, since $0 \leqq \nu \leqq 1$. Thus,

$$I_2 = \zeta(2)^{-1} \int_{-\infty}^{\infty} dx \int dY \int_{-\infty}^{\infty} dt\, t \sum_{k \geqq |t|} \sigma(k) k^{-3} \rho(Y) \rho(t\tilde{Y} + xY)$$

$$(24) \qquad = \zeta(2)^{-1} \int\int \rho(X_1)\rho(X_2)\tau(\|X_1, X_2\|) dX_1 dX_2$$

$$= V^2 + I_3,$$

where

$$I_3 = \zeta(2)^{-1} \int\int \rho(X_1)\rho(X_2)\{\tau(\|X_1, X_2\|) - \zeta(2)\} dX_1 dX_2.$$

According to Lemma 7,

$$|\tau(t) - \zeta(2)| \leqq \max(c_9, c_{10} t^{-1} \log t).$$

Therefore, if $V \leqq 10$ for instance,

$$(25) \qquad I_3 \leqq c_9 V^2 < c_{11} V.$$

If $V > 10$, then we put

$$\chi(t) = \begin{cases} c_9, & \text{when} \quad t < 10 \log^{-1} 10, \\ c_{10} t^{-1} \log t, & \text{when} \quad 10 \log^{-1} 10 \leqq t < V \log^{-1} V, \\ 0, & \text{if} \quad t \geqq V \log^{-1} V. \end{cases}$$

Observing

$$\int_0^\infty \chi(t) dt \leqq c_{12} \log^2 V$$

and using Theorem 3, we obtain

$$I_3 < \int\int \rho(X_1)\rho(X_2)\chi(\|X_1, X_2\|) dX_1 dX_2 + c_{13} V \log^2 V$$

$$\leqq c_{14} V \log^2 V.$$

The last equation, together with (24) and (25), proves (22). The second assertion of Lemma 10 is proved similarly.

10. Every lattice-point $g \neq 0$ can uniquely be written $g = l \cdot g^*$, where l is a positive integer and g^* is primitive. We write $l = l(g)$. Let S be a Borel set with characteristic function $\rho(X)$ and volume $V(S)$, not containing the origin 0, and let k be a positive integer.

LEMMA 11.

$$\int_C \left(\|A\| \sum \left[\begin{matrix} g \\ l(g) \leqq k \end{matrix} \right] \rho(Ag) - V \right)^2 dm(A) \leqq c_{15}(V^2/k + V \lg k + V \lg^2 V),$$

$$\int_C \left(\| A \| \sum \left[\begin{matrix} g \\ \text{primitive} \end{matrix} \right] \rho(Ag) - V\zeta(2)^{-1} \right)^2 dm(A) \leqq c_{16} V \lg^2 V.$$

Proof. The first equation is a consequence of the fact, following from Lemma 9, that

$$\int_C \| A \| \sum \left[\begin{matrix} g \\ l(g) \leqq k \end{matrix} \right] \rho(Ag) dm(A) = V\zeta(2)^{-1} \sum_{t=1}^{k} t^{-2},$$

where

$$\left| V\zeta(2)^{-1} \sum_{t=1}^{k} t^{-2} - V \right| \leqq c_{17} V k^{-1},$$

and the estimate

$$\begin{aligned} &\| A \|^2 \left(\sum \left[\begin{matrix} g \\ l(g) \leqq k \end{matrix} \right] \rho(Ag) \right)^2 \\ &\leqq \| A \|^2 \sum \left[\begin{matrix} g_1, g_2 \\ \| g_1, g_2 \| \neq 0 \end{matrix} \right] \rho(Ag_1)\rho(Ag_2) \\ &\quad + \| A \|^2 \sum \left[\begin{matrix} g_1, g_2 \\ \| g_1, g_2 \| = 0 \\ l(g_i) \leqq k \end{matrix} \right] \rho(Ag_1)\rho(Ag_2). \end{aligned}$$

The integral of the first expression on the right hand side is dealt with in Lemma 10, while for the second expression

$$\begin{aligned} &\int_F \sum \left[\begin{matrix} g_1, g_2 \\ \| g_1, g_2 \| = 0 \\ l(g_i) \leqq k \end{matrix} \right] \rho(Ag_1)\rho(Ag_2) d\mu(A) \\ &\leqq \int_F \sum_{p \neq 0, |p| \leqq k} \sum_{q \neq 0, |q| \leqq |p|} \sum \left[\begin{matrix} g \\ \text{primitive} \end{matrix} \right] \rho(Apg)\rho(Aqg) d\mu(A) \\ &< 4 \sum_{1 \leqq p \leqq k} \frac{p}{p^2} V \leqq c_{18} V \lg k \end{aligned}$$

and therefore

$$\begin{aligned} &\int_C \| A \|^2 \sum \left[\begin{matrix} g_1, g_2 \\ \| g_1, g_2 \| = 0 \\ l(g_i) \leqq k \end{matrix} \right] \rho(Ag_1)\rho(Ag_2)\, dm(A) \\ &\leqq c_{18} \int_0^1 \nu^2 (V\nu^{-1}) \lg k d\nu < c_{18} V \lg k. \end{aligned}$$

The proof of the second equation of Lemma 11 is even simpler.

11. Now, as in §3, we pick sets $S(N)$ of volume N and define $\rho_N(X)$, ${}_{N_1}\rho_{N_2}(X)$, $R_N(A)$, $S_N(A)$, ${}_{N_1}R_{N_2}(A)$, ${}_{N_1}S_{N_2}(A)$ as before. We also introduce

$$R_N(T, A) = \sum \begin{bmatrix} g \\ l(g) \leqq 2^T \end{bmatrix} \rho_N(Ag) - N\| A \|^{-1},$$

$$R_N^*(T, A) = \sum \begin{bmatrix} g \\ l(g) > 2^T \end{bmatrix} \rho_N(Ag)$$

and similarly ${}_{N_1}R_{N_2}(T, A)$.

LEMMA 12. *If K_T is defined as in Lemma 2, then for large enough T*

$$\sum_{(N_1,N_2)\in K_T} \int_C \|A\|^2 \quad {}_{N_1}R^2_{N_2}(T, A)\, dm(A) < c_{19}T^3 2^T,$$

$$\sum_{(N_1,N_2)\in K_T} \int_C \|A\|^2 \quad {}_{N_1}S^2_{N_2}(A)\, dm(A) < c_{20}T^3 2^T,$$

$$\int_C \|A\| \quad R^*_{2^T}(T, A)\, dm(A) < 1.$$

Proof. Similarly as for Lemma 2. We use Lemma 11 and

$$\sum_{(N_1,N_2)\in K_T} (N_2 - N_1) \log^2 (N_2 - N_1) = O(T^3 2^T),$$

$$\sum_{(N_1,N_2)\in K_T} (N_2 - N_1) \log 2^T = O(T^2 2^T),$$

$$\sum_{(N_1,N_2)\in K_T} \frac{(N_2 - N_1)^2}{2^T} < \sum_{(N_1,N_2)\in K_T} (N_2 - N_1) = O(T2^T).$$

The last assertion of Lemma 12 is a consequence of

$$\int_F R^*_{2^T} \quad (T, A)d\mu(A) = 2^T\zeta(2)^{-1}\left(\sum_{k>2^T} k^{-2}\right) < 1.$$

LEMMA 13. *For all sufficiently large T there is a subset B_T of C of measure*

$$\int_{B_T} dm(A) \leqq c_{21}\psi^{-1}(T \log 2 - 1) + 2^{-T/2}$$

such that

$$\|A\| \quad R_N(A) \leqq T^2 2^{T/2}\psi^{1/2}(T \log 2 - 1) + 2^{T/2},$$

$$\|A\| \quad S_N(A) \leqq T^2 2^{T/2}\psi^{1/2}(T \log 2 - 1)$$

for every $N \leqq 2^T$ and all $A \in C$, but $A \notin B_T$.

Proof. We take B_T to be the set consisting of all $A \in C$ for which it is not true that

$$\|A\|^2 \sum_{(N_1,N_2)\in K_T} {}_{N_1}R^2_{N_2}(T, A) \leqq T^3 2^T \psi(T\log 2 - 1),$$

$$\|A\|^2 \sum_{(N_1,N_2)\in K_T} {}_{N_1}S^2_{N_2}(A) \leqq T^3 2^T \psi(T\log 2 - 1),$$

$$\|A\| R^*_{2^T}(T, A) \leqq 2^{T/2}.$$

Then

$$\int_{B_T} dA \leqq (c_{19} + c_{20})\psi^{-1}(T\log 2 - 1) + 2^{-T/2}.$$

If, however, $N \leqq 2^T$, $A \in C$ but not in B_T, then we can show, as in Lemma 3, that

$$\|A\|^2 \; R^2_N(T, A) \leqq T^4 2^T \psi(T\log 2 - 1),$$
$$\|A\| R_N(T, A) \leqq T^2 2^{T/2}\psi^{1/2}(T\log 2 - 1),$$
$$\|A\| S_N(A) \leqq T^2 2^{T/2}\psi^{1/2}(T\log 2 - 1),$$
$$\|A\| R^*_N(T, A) \leqq 2^{T/2}$$

and since $R_N(A) = R_N(T, A) + R^*_N(T, A)$, the result follows.

12. **Proof of Theorem** 2. From here on, the proof proceeds exactly as the proof of Theorem 1. We only have to observe that dA and $dm(A)$ are equivalent measures.

BIBLIOGRAPHY

1. W. Blaschke, *Vorlesungen ueber Differentialgeometrie* II, 1st and 2nd ed., Berlin, 1923.

2. J. W. S. Cassels, *Some metrical theorems in diophantine approximation* III, Proc. Cambridge Philos. Soc. vol. 46 (1950) pp. 219–225.

3. P. R. Halmos, *Measure theory*, New York, Van Nostrand, 1950.

4. G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, 3rd. ed., Oxford, 1954.

5. C. G. Lekkerkerker, *Lattice points in unbounded point sets* I, Indag. Math. vol. 20 (1958) pp. 197–205.

6. A. M. Macbeath and C. A. Rogers, *Siegel's mean value theorem in the geometry of numbers*, Proc. Cambridge Philos. Soc. vol. 54 (1958) pp. 139–151.

7. C. T. Pan, *On $\sigma(n)$ and $\vartheta(n)$*, Bull. Acad. Polon. Sci. Cl. III vol. 4 (1956) pp. 637–638.

8. C. A. Rogers, *Mean values over the space of lattices*, Acta Math. vol. 94 (1955) pp. 249–287.

9. W. Schmidt, *Mittelwerte ueber Gitter* II, Monatsh. Math. vol. 62 (1958) pp. 250–258.

10. C. L. Siegel, *A mean value theorem in geometry of numbers*, Ann. of Math. vol. 46 (1945) pp. 340–347.

MONTANA STATE UNIVERSITY,
MISSOULA, MONTANA