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

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

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

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

Next, let $\Phi$ 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, $\psi(s)$, $s \geq 0$, 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 $A$),

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

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

for $S \in \Phi$.

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

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

whenever $V(S) \geq c_2(A)$ and $S \in \Phi$.

In $R_2$ our results are a little weaker:

Received by the editors March 13, 1959.

516
Theorem 2. Let \( n = 2 \). Then for almost every linear transformation \( A \)

\[
D(AS) = O(V^{-1/2} \log^2 V \psi^{1/2}(\log V)),
\]

\[
E(AS) = O(V^{-1/2} \log^2 V \psi^{1/2}(\log V)),
\]

for \( S \in \Phi \).

Theorem 1 and Theorem 2 yield the

**Corollary.** Suppose \( S \) is a Borel set with infinite volume in \( \mathbb{R}^n \), \( n > 1 \). 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 \( \mathbb{R}_1 \). Therefore Theorem 1 has no analogon in \( \mathbb{R}_1 \).

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

**Theorem 3.** Let \( S \) be a Lebesgue-measurable set in \( \mathbb{R}_n \) with characteristic function \( \rho(X) \) and volume \( V \). By \( \|X_1, \ldots, X_n\| \) we denote the absolute value of the \( n \times n \)-determinant, whose row-vectors are \( X_1, \ldots, X_n \). Then

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

for every non-negative, nonincreasing function \( \chi(t) \), defined for \( t \geq 0 \), 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

\[
\int_F \sum \rho(Ag) d\mu(A) = V,
\]
518 WOLFGANG SCHMIDT

\[ \int_F \sum_{\text{primitive}} \rho(Ag) d\mu(A) = V\xi(n)^{-1}. \]

The summation is over lattice-points \( g \). Further, if \( n > 2 \),

\[ \int_F \sum_{\text{lin. indep.}} \rho(Ag_1)\rho(Ag_2) d\mu(A) = V^2, \]

and

\[ \int_F \sum_{\text{lin. indep. primitive}} \rho(Ag_1)\rho(Ag_2) d\mu(A) = V^2\xi(n)^{-2}. \]

We observe, again for \( n > 2 \),

\[ \int_F \sum_{\text{lin. dep.}} \rho(Ag_1)\rho(Ag_2) d\mu(A) \]

\[ = \frac{1}{2} \int_F \sum_{p,q} \sum_{\text{primitive}} \rho(Apg)\rho(Aqg) d\mu(A) \]

\[ \leq \sum_{q \neq 0} \sum_{p, |p| \leq |q|} \int_F \sum_{\text{primitive}} \rho(Aqg) d\mu(A) \]

\[ < V \sum_{q \neq 0} \sum_{p, |p| \leq |q|} |q|^{-n} < c_3 V, \]

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

\[ \int_F \left( \sum_{g} \rho(Ag) - V \right)^2 d\mu(A) < c_3 V \]

and

\[ \int_F \left( \sum_{\text{primitive}} \rho(Ag) - V\xi(n)^{-1} \right)^2 d\mu(A) < c_3 V. \]

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

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

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

Here, \( C \) is the cone of matrices \( A \) which satisfy \( \lambda A \in F \) for some \( \lambda \geq 1 \). \( dA \) 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

\[
(12) \quad \int_c \left( \sum_\psi \rho(A_\psi) - \sqrt{\|A\|^{-1}} \right)^2 dA < c_6(n) V \int_0^1 \nu^{n-2} d\nu < c_6(n) V, \]

\[
(13) \quad \int_c \left( \sum \left[ \text{primitive} \right] \rho(A_\psi) - V \xi(n)^{-1} \|A\|^{-1} \right)^2 dA < c_6(n) V, \]

where \( \|A\| \) is the absolute value of the determinant of \( A \).

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

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

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

(iiiia) \( \alpha(X) \) is closed.

For let \( V \) be a limit point of \( \alpha(\Phi) \). We first assume there exist \( V_1 \geq V_2 \cdots, V_j \in \alpha(\Phi) \), \( \lim V_j = V \). Then there exist \( S_1 \supseteq S_2 \supseteq \cdots, S_j \in \Phi \), \( V(S_j) = V_j \). We take \( S = S(V) = \cap_1^\infty S_j \). Then \( S \) is again a Borel set, and \( V(S) = V \). If, however, there exists no sequence \( V_1 \geq V_2 \geq \cdots \) of the required form, then there exists a sequence \( V_1 \leq V_2 \leq \cdots, V_j \in \alpha(\Phi) \), \( \lim V_j = V \) and we can proceed similarly, taking \( S = \cup_1^\infty S_j \). If we take \( \chi \) as the union of \( \Phi \) 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 \( V \in \alpha(X) \). Then there exist \( V_1 \in \alpha(X), V_2 \in \alpha(X), V_1 < V < V_2 \), 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 \( X \in S_2, X \in S_1 \), and \( (S_2 - S_1) \), for the set of all \( X \in S_2 - S_1 \) with \( |X| \leq t \). Then \( V(S_1 \cup (S_2 - S_1)) \) 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 \( V(S_1 \cup (S_2 - S_1)) = V \). We introduce

\[ S(V) = S_1 \cup (S_2 - S_1) \]

and take \( \Psi \) to be the union of \( X \) and all sets \( S(V) \).

3. According to Lemma 1, we may assume that \( \Phi \) 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

\[ n_1 \rho_{N_1}(X) \]

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

\[ \text{(i)} \quad \text{We may assume that the null set belongs to } X, \text{ such that } 0 \in \alpha(X). \]
\[ R_N(A) = \sum_{\sigma} \rho_N(Ag) - N\|A\|^{-1}, \]
\[ S_N(A) = \sum_{\text{primitive}} g \rho_N(Ag) - N\xi(n)^{-1}\|A\|^{-1}, \]
\[ N_1 R_{N_1}(A) = \sum_{\sigma} N_1 \rho_N(Ag) - (N_2 - N_1)\|A\|^{-1}, \]
\[ N_2 S_{N_2}(A) = \sum_{\text{primitive}} g N_2 \rho_N(Ag) - (N_2 - N_1)\xi(n)^{-1}\|A\|^{-1}. \]

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

\[ \sum_{(N_1, N_2) \in K_T} \int_C N_1 R_{N_1}^2(A) dA \leq c_6(n)(T + 1)2^T. \]

**Proof.** (12) yields

\[ \int_C N_1 R_{N_1}^2(A) dA \leq c_6(n)(N_2 - N_1). \]

Each value of \( N_2 - N_1 = 2^t \) \((0 \leq t \leq T)\) 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

\[ \int_{B_T} dA \leq c_6(n)\psi^{-1}(T \log 2 - 1), \]

such that

\[ R_N(A) \leq T(T + 1)2^T\psi(T \log 2 - 1) \]

for every \( N \leq 2^T \) 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

\[ \sum_{(N_1, N_2) \in K_T} N_1 R_{N_1}^2(A) \leq (T + 1)2^T\psi(T \log 2 - 1), \]

then (15) follows immediately from Lemma 2.

Assume \( N \leq 2^T \), \( A \in C \) 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
$R_N(A) = \sum N_i R_{N_i}(A)$,

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

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

4. **Proof of Theorem 1.** The set of $A \in C$ belonging to $B_T$ has measure at most $c_6(n)\psi^{-1}(T \log 2 - 1)$. 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 $A \in B_T$ for $T \geq T_0$. Assume $N \geq N_0 = 2^{T_0}$ and choose $T$ such that $2^{T-1} \leq N < 2^T$. Then, by Lemma 3,

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

$$< \left(\frac{\log N}{\log 2} + 1\right)\left(\frac{\log N}{\log 2} + 2\right)2N\psi(\log N)$$

$$< c_6N \log^2 N\psi(\log N)$$

for almost every $A$. Thus,

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

for almost every $A \in C$. Since every $\|A\| \leq 1$ 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 $\|A\| \leq 1$. Applying a linear transformation we see that (17) holds for almost every $A$ satisfying $\|A\| \leq c$, 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

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

for almost all $A$.

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

$$L(A\cdot S(N)) - (N + 1) \leq L(A\cdot S) - V(S) \leq L(A\cdot S(N + 1)) - N$$

and

$$L(A\cdot S(N + 1)) - (N + 1) = L(A\cdot S(N)) - (N + 1) < N^{-1}.$$
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

\[
\|X_1, \cdots, X_{n-1}\|
\]

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

\[
h_k(X_1, \cdots, X_n) = \|X_1, \cdots, \overline{X}_{k-1}, \overline{X}_{k+1}, \cdots, \overline{X}_n\|, \quad (1 \leq k \leq n)
\]

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

\[
h_k(X_1, \cdots, X_n) \geq h_i(X_1, \cdots, X_n), \quad (1 \leq i \leq n).
\]

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

\[
\|X_1, \cdots, X_n\| \leq 1 \quad \text{or} \quad \|X_1, \cdots, X_n\| \leq t
\]

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.**

\[
\int_{\overline{H}} \cdots \int_{\overline{H}} \rho(X_1) \cdots \rho(X_n) dX_1 \cdots dX_n \leq n2^n V^{n-1}.
\]

**Proof.** First, we observe

\[
\begin{align*}
\int_{\overline{H}} \cdots \int_{\overline{H}} \rho(X_1) \cdots \rho(X_n) dX_1 \cdots dX_n &= \sum_{k=1}^n \int_{H \cap H_k} \cdots \int_{H \cap H_k} \rho(X_1) \cdots \rho(X_n) dX_1 \cdots dX_n.
\end{align*}
\]

Next,

\[
\begin{align*}
\int_{H \cap H_1} \rho(X_1) \cdots \rho(X_n) dX_1 \cdots dX_n &\leq 2^n \int \cdots \int \rho(X_2) \cdots \rho(X_n) dX_2 \cdots dX_n = 2^n V^{n-1}.
\end{align*}
\]
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

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

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

$$
\leq - \int_0^\infty n2^n V^{n-1} t dX(t) = n2^n V^{n-1} \int_0^\infty X(t) dt.
$$

6. From now on we restrict ourselves to 2 dimensions. $X, Y, \ldots$ are points in $R_2$, $g, h, \ldots$ 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 $\varphi$-function.

**Lemma 6.** Assume $k \neq 0$. There exist $\sigma(\lfloor k \rfloor)$ classes $E$ with $d(E) = k$ and $\varphi(\lfloor k \rfloor)$ 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

$$
(19) \quad g_1(n, 0), g_2(m, k/n),
$$

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

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

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 $\varphi(\lfloor k \rfloor)$ classes.

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

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

$$
\omega(t) = t \sum_{k \geq t} \varphi(k) k^{-3}.
$$
Lemma 7.

\[
\tau(t) = \xi(2) + O(t^{-1} \log t),
\]

\[
\omega(t) = \xi(2)^{-1} + O(t^{-1} \log t).
\]

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

\[
\sum_{t \leq k \leq s} \sigma(k) = (s^2 - t^2) \xi(2)/2 + O(s \log s)
\]

and therefore, using Theorem 421 of [4],

\[
\sum_{t \leq k \leq s} \sigma(k) k^{-3} = \int_{t}^{s} (s^2 - t^2) 3s^{-4} \xi(2)/2 ds + O\left( \int_{t}^{s} s \log s s^{-4} ds \right)
\]

\[
= t^{-1} \xi(2) + O(t^{-2} \log t).
\]

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

\[
\sum_{t \leq k \leq s} \sigma(k) = s^2 \xi(2)/2 + O(s(\log s \log \log s)^{3/4})
\]

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 \(\bar{X}\) satisfying \(|X, \bar{X}| = 1\).

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

\[
\int_{F} \sum \left[ (g_1, g_2) \in E \right] \rho(Ag_1) \rho(Ag_2) d\mu(A)
\]

\[
= \left| k \right|^{-\xi(2)^{-1}} \int \left\{ \int \rho(X) \rho(\bar{kX} + xX) dx \right\} dX.
\]

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 \(\bar{X}\).

As an immediate consequence of Lemmas 6, 8 we obtain

\[
\int_{F} \left[ \left| \begin{array}{c} g_1, g_2 \\ ||g_1, g_2|| \neq 0 \end{array} \right| \right] \rho(Ag_1) \rho(Ag_2) d\mu(A)
\]

\[
= \sum_{k \neq 0} \sigma(\left| k \right|) \left| k \right|^{-\xi(2)^{-1}} \int \left\{ \int \rho(X) \rho(\bar{kX} + xX) dx \right\} dX.
\]
Instead of the volume element $dA$ as used in §1, we now define the volume element $dm(A)$ by

\[(21) \quad \int_C \sigma(A) dm(A) = \int_0^1 \left\{ \int_F \sigma(v^{1/2}A) d\mu(A) \right\} dv.\]

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

**Lemma 9.**

\[I_1 = \int_C ||A|| \sum_v \rho(Ag) dm(A) = V,\]

\[I_1^* = \int_C ||A|| \sum_{\text{primitive}} \rho(Ag) dm(A) = V \xi(2)^{-1}.\]

**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

\[I_2 = \int_C ||A||^2 \sum_{g_1, g_2} \rho(Ag_1) \rho(Ag_2) dm(A),\]

\[I_2^* = \int_C ||A||^2 \sum_{\text{primitive}} \rho(Ag_1) \rho(Ag_2) dm(A)\]

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

\[(22) \quad |I_2 - V^2| \leq c_7 V \lg^2 V,\]

\[(23) \quad |I_2^* - V^2 \xi(2)^{-1}| \leq c_9 V \lg^2 V.\]

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

\[I_2 = \xi(2)^{-1} \int_0^1 \nu^2 \left\{ \sum_{k \neq 0} \sigma(|k|) |k|^{-1} \left\{ \int \rho(v^{1/2}X) \rho(kv^{1/2}X + xv^{1/2}X) dx \right\} dX \right\} dv.\]

Substituting $\nu^{1/2}X = Y$ we have $dX = \nu^{-1} dY$ and may assume $\nu^{-1/2}X = \tilde{Y}$. Next, we write $t = k\nu$ and note that if we integrate $f_x dt$, we have to replace the sum $\sum_{k \neq 0}$ by $\sum_{k \neq 0 \mid 0}$, since $0 \leq \nu \leq 1$. Thus,
\[ I_2 = \xi(2)^{-1} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dt \sum_{k \geq 1} \sigma(k)k^{-3}\rho(Y)\rho(tY + xY) \]
\[
= \xi(2)^{-1} \int \int \rho(X_1)\rho(X_2)\tau(\|X_1, X_2\|)dX_1dX_2 \\
= V^2 + I_3,
\]
where
\[ I_3 = \xi(2)^{-1} \int \int \rho(X_1)\rho(X_2)\{\tau(\|X_1, X_2\|) - \xi(2)\}dX_1dX_2. \]

According to Lemma 7, 
\[ |\tau(t) - \xi(2)| \leq \max(\varepsilon_9, c_{10}t^{-1}\log t). \]

Therefore, if \( V \leq 10 \) for instance, 
\[ I_3 \leq c_9V^2 < c_{11}V. \]

If \( V > 10 \), then we put
\[ \chi(t) = \begin{cases} 
    c_9, & \text{when } t < 10 \log^{-1} 10, \\
    c_{10}t^{-1}\log t, & \text{when } 10 \log^{-1} 10 \leq t < V \log^{-1} V, \\
    0, & \text{if } t \geq V \log^{-1} V.
\end{cases} \]

Observing
\[ \int_0^{\infty} \chi(t)dt \leq c_{12} \log^2 V \]
and using Theorem 3, we obtain
\[ I_3 < \int \int \rho(X_1)\rho(X_2)\chi(\|X_1, X_2\|)dX_1dX_2 + c_{13}V \log^2 V \]
\[ \leq c_{14}V \log^2 V. \]

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.**
\[ \int_c \left( \|A\| \sum_{l(g) \leq k} \frac{g}{l(g)} \right) \rho(Ag) - V \right)^2 dm(A) \leq c_{16}(V^2/k + V \log k + V \log^2 V), \]
\[
\int_0^1 \left( \| A \| \sum \left[ \sum_{l(g) \leq k} \right] \rho(Ag) - V(z(2))^{-1} \right)^2 \, dm(A) \leq c_{16} V \lg^2 V.
\]

**Proof.** The first equation is a consequence of the fact, following from Lemma 9, that
\[
\int_0^1 \| A \| \sum \left[ \sum_{l(g) \leq k} \right] \rho(Ag) \, dm(A) = V(z(2))^{-1} \sum_{t=1}^{k} t^{-2},
\]
where
\[
\left| V(z(2))^{-1} \sum_{t=1}^{k} t^{-2} - V \right| \leq c_{17} V k^{-1},
\]
and the estimate
\[
\| A \|^2 \left( \sum \left[ \sum_{l(g) \leq k} \right] \rho(Ag) \right)^2 \leq \| A \|^2 \sum \left[ \sum_{\| g_1, g_2 \| = 0} \right] \rho(Ag_1) \rho(Ag_2) + \| A \|^2 \sum \left[ \sum_{\| g_1, g_2 \| = 0 \, l(g_i) \leq k} \right] \rho(Ag_1) \rho(Ag_2).
\]
The integral of the first expression on the right hand side is dealt with in Lemma 10, while for the second expression
\[
\int_F \sum \left[ \sum_{\| g_1, g_2 \| = 0 \, l(g_i) \leq k} \right] \rho(Ag_1) \rho(Ag_2) \, d\mu(A)
\]
\[
\leq \int_F \sum_{p=0}^{\lfloor \frac{1}{2} \rfloor} \sum_{q=0}^{\lfloor \frac{1}{2} \rfloor} \sum \left[ \sum_{\text{primitive}} \right] \rho(Apg) \rho(Aqg) \, d\mu(A)
\]
\[
< 4 \sum_{1 \leq p \leq k} \frac{p}{p^2} V \leq c_{18} V \lg k
\]
and therefore
\[
\int_0^1 \| A \|^2 \sum \left[ \sum_{\| g_1, g_2 \| = 0 \, l(g_i) \leq k} \right] \rho(Ag_1) \rho(Ag_2) \, d\mu(A)
\]
\[
\leq c_{18} \int_0^1 V^2(V^{-1}) \lg k \, dv < c_{18} V \lg k.
\]
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 $\rho_N(X)$, $n_3\rho_N(X)$, $R_N(A)$, $S_N(A)$, $n_4R_N(A)$, $n_5S_N(A)$ as before. We also introduce 

$$R_N(T, A) = \sum_{l(g) \leq 2^T} \rho_N(Ag) - N\|A\|^{-1},$$

$$R_N^*(T, A) = \sum_{l(g) > 2^T} \rho_N(Ag)$$

and similarly $n_4R_N^2(T, A)$.

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

$$\sum_{(N_1, N_2) \in K_T} \int_C \|A\|^2 N_1R_{N_1}^2(T, A) \, dm(A) < c_{16}T^32^T,$$

$$\sum_{(N_1, N_2) \in K_T} \int_C \|A\|^2 N_3S_{N_2}^2(A) \, dm(A) < c_{20}T^32^T,$$

$$\int_C \|A\| R_N^*T(T, A) \, dm(A) < 1.$$

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

$$\sum_{(N_1, N_2) \in K_T} (N_2 - N_1) \log^2 (N_2 - N_1) = O(T^32^T),$$

$$\sum_{(N_1, N_2) \in K_T} (N_2 - N_1) \log 2^T = O(T^22^T),$$

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

The last assertion of Lemma 12 is a consequence of

$$\int_C R_N^*T(T, A) \, d\mu(A) = 2T^3(2)^{-1} \left( \sum_{k \geq 2^T} k^{-2} \right) < 1.$$

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

$$\int_{B_T} dm(A) \leq c_2\psi^{-1}(T \log 2 - 1) + 2^{-T/2}$$

such that

$$\|A\| R_N(A) \leq T^22^{T/2}\psi^{1/2}(T \log 2 - 1) + 2^{T/2},$$

$$\|A\| S_N(A) \leq T^22^{T/2}\psi^{1/2}(T \log 2 - 1)$$

for every $N \leq 2^T$ 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

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

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

$$\| A \|^2 R_{N}^*(T, A) \leq 2^{T/2}.$$

Then

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

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

$$\| A \|^2 R_N(T, A) \leq T^4 2^T \psi(T \log 2 - 1),$$

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

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

$$\| A \|^2 R_N^*(T, A) \leq 2^{T/2}$$

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


Montana State University,
Missoula, Montana