HOMOGENEOUS DYNAMICS AND NUMBER THEORY

Gregory Margulis

Homogeneous dynamics is another name for the theory of flows on homogeneous spaces, or homogeneous flows. The study of homogeneous flows has been attracting considerable attention for the last 40-50 years. During the last three decades, it has been realized that some problems in number theory and, in particular, in Diophantine approximation, can be solved using methods from the theory of homogeneous flows. The purpose of these lecture notes is to demonstrate this approach on several examples rather than to give a survey of the subject. In particular, recent very important work by M.Einsiedler, E.Lindenstrauss, P.Michel and A.Venkatesh on application of diagonalizable flows to number theory will not be discussed. (This work was the subject of two invited addresses at ICM 2006).

We will start with the Oppenheim conjecture proved in the mid 1980's and the Littlewood conjecture, still not settled. The next topic is quantitative generalizations of the Oppenheim conjecture and counting lattice points on homogeneous varieties. After that, we will go to Diophantine approximation on manifolds where, in particular, we will very briefly describe the proof of Baker-Sprindzuk conjectures. We will also discuss results on translates of submanifolds and unipotent flows which are used in applications to number theory.

But before all that, let us mention the following observation which was originally made by D.Zagier about 20 years ago. Let $G = SL(2, \mathbf{R})$, let $\Gamma = SL(2, \mathbf{Z})$ and let $U = \{u_t : t \in \mathbf{R}\}$ where $u_t = \begin{pmatrix} 1 & t \\ 0 & 0 \end{pmatrix}$. For each $t > 0$, there is a unique closed orbit of U in G/Γ of period t, call it $C_t(u_t x = x$ and $u_{t'}x \neq x$ for $x \in C_t$ and $0 < t' < t)$. The famous Riemann hypothesis about zeros of Riemann zeta-function is equivalent to the following statement. For any function $f \in C_0^\infty(G/\Gamma)$ (i.e. f is smooth and compactly supported) and any $\varepsilon > 0$

$$\frac{1}{t} \int_{C_t} f dm = \int_{G/\Gamma} f d\mu + O(t^{-(3/4)+\varepsilon})$$

as $t \to \infty$ where m is the Lebesgue measure on C_t and μ is the G-invariant probability Borel measure on G/Γ. It should be noted that (a) the power $-\frac{3}{4} + \varepsilon$ is different from the usual power $-\frac{1}{2}+\varepsilon$ which one could expect from probabilistic type arguments; (b) the equivalence remains true if one considers only K-invariant functions f where K is a maximal compact subgroup of G; (c) apparently the just mentioned equivalence is considered mostly as a curiosity rather than a serious approach to the proof of the Riemann hypothesis.

§1. Values of quadratic forms and of products of linear forms at integral points

We will say that a real quadratic form is *rational* if it is a multiple of a form with rational coefficients and *irrational* otherwise.

Theorem 1. *Let Q be a real irrational indefinite nondegenerate quadratic form in $n \geq 3$ variables. Then for any $\varepsilon > 0$, there exist integers $x_1, \dots, x_n$ not all equal to 0 such that $|Q(x_1, \dots, x_n)| < \varepsilon$.*

Theorem 1 was conjectured by A.Oppenheim in 1929 and proved by the author in 1986 (see [Marg1] and [Marg2] and references therein). Oppenheim was motivated by Meyer's theorem that if Q is a rational quadratic form in $n \geq 5$ variables, then Q represents zero over $\mathbf{Z}$ nontrivially, i.e. there exists $x \in \mathbf{Z}^n, x \neq 0$ such that $Q(x) = 0$. Because of that he originally stated the conjecture only for $n \geq 5$. Let us also note that the condition "$n \geq 3$" cannot be replaced by the condition "$n \geq 2$". To see this, consider the form

$$x_1^2 - \lambda x_2^2 = (x_1 - \sqrt{\lambda} x_2)(x_1 + \sqrt{\lambda} x_2)$$

where λ is an irrational positive number such that $\sqrt{\lambda}$ has a continued fraction development with bounded partial quotients; for example $\lambda = (1+\sqrt{3})^2 = 4+2\sqrt{3}$.

It is a standard simple fact that if Q is a real irrational indefinite nondegenerate quadratic form in n variables and $2 \leq m < n$, then $\mathbf{R}^n$ contains a rational subspace L of dimension m such that the restriction of B to L is irrational, indefinite and nondegenerate. Hence if Theorem 1 is proved for some n_0, then it is proved for all $n \geq n_0$. As a consequence of this remark, we see that it is enough to prove Theorem 1 for $n = 3$.

Before the Oppenheim conjecture was proved, it was extensively studied mostly using analytic number theory methods. In particular it had been proved by H.Davenport and his coauthors for diagonal forms in five or more variables and for $n \geq 21$ (during approximately a 15 year period between 1945 and 1960). About 10 years ago, V.Bentkus and F.Götze proved the Oppenheim conjecture for $n \geq 9$ using analytic number theory methods. They also proved, under the same assumption

$n \geq 9$, that if Q is a positive definite quadratic form on $\mathbf{R}^n$ and

$$E = \{x \in \mathbf{R}^n : Q(x) \geq 1\}$$

then for $r \to \infty$ the number of points in $\mathbf{Z}^r \cap rE$ differs from $\mathrm{vol}(rE)$ by an error of order $O(r^{n-2})$ in general and of order $o(r^{n-2})$ for irrational Q. In [G]. Götze proved this result under a weaker assumption $n \geq 5$. As a corollary of this, one gets the proof of the Davenport-Lewis conjecture about gaps between values of irrational positive definite quadratic forms at integral points. The proofs given in [G] and in the previous work by Bentkus and Götze are in a certain sense "effective" contrary to the author's proof of Theorem 1. In the joint work in progress by F.Götze and the author, we obtained an "effective" proof of the Oppenheim conjecture and its quantitative versions in the case $n \geq 5$; this proof is based on the combination of methods from [EMM1] and [G]. But it seems that the methods of analytic number theory are not sufficient to prove the Oppenheim conjecture in the case where n is 3 or 4.

Theorem 1 was proved by studying orbits of the orthogonal group $SO(2,1)$ on the space of unimodular matrices in $\mathbf{R}^3$. It turns out that this theorem is equivalent to the following:

Theorem 2. *Let $G = SL(3,\mathbf{R})$ and $\Gamma = SL(3,\mathbf{Z})$. Let us denote by H the group of elements of G preserving the form $2x_1x_3 - x_2^2$ and by $\Omega_3 = G/\Gamma$ the space of lattice in $\mathbf{R}^3$ having determinant 1. Let G_y denote the stabilizer $\{g \in G : gy = y\}$ of $y \in \Omega_3$ in G. If $z \in \Omega_3 = G/\Gamma$ and the orbit Hz is relatively compact in Ω_3, then the quotient space $H/H \cap G_z$ is compact.*

For any real quadratic form B in n variables, let us denote the special orthogonal group

$$SO(B) = \{g \in SL(n,\mathbf{R}) : gB = B\} \text{ by } H_B$$

and

$$\inf\{|B(x)| : x \in \mathbf{Z}^n, x \neq 0\} \text{ by } m(B).$$

As it was already noted, it is enough to prove Theorem 1 for $n = 3$. Then the equivalence of Theorem 1 and 2 is a consequence of the assertions (i) and (ii) below. The assertion (i) easily follows from Mahler's compactness criterion. As for (ii), it was essentially proved in the 1955 paper [CS] by Cassels and Swinnerton-Dyer by some rather elementary considerations; (ii) can be also deduced from Borel's density theorem.

(i) *Let B be a real quadratic form in n variables. Then the orbit $H_B\mathbf{Z}^n$ is relatively compact in the space $\Omega_n = SL(n,\mathbf{R})/SL(n,\mathbf{Z})$ of unimodular* *lattices in* $\mathbf{R}^n$ *if and only if* $m(B) > 0$.

(ii) *Let* B *be a real irrational indefinite nondegenerate quadratic form in* 3 *variables. Then* $H_B/H_B \cap SL(3, \mathbf{Z})$ *is compact if and only if the form* B *is rational and anisotropic over* $\mathbf{Q}$.

Remark. In an implicit form, the equivalence of Theorem 1 and 2 appears already in the abovementioned paper [CS] by Cassels and Swinnerton-Dyer. In the mid-1970's, M.S.Raghunathan rediscovered this equivalence and noticed that the Oppenheim conjecture would follow from a conjecture about closures of orbits of unipotent subgroups. This conjecture and some other related conjectures will be discussed in §5. Raghunathan's observations inspired the author's work on the homogeneous space approach to the Oppenheim conjecture.

Theorem 2 was also used to prove the following stronger version of Theorem 1:

Theorem 3. *If* Q *and* n *are the same as in Theorem 1, then* $Q(\mathbf{Z}^n)$ *is dense in* $\mathbf{R}$ *or, in other words, for any* $a < b$ *there exist integers* $x_1, \dots, x_n$ *such that*

$$a < Q(x_1, \dots, x_n) < b.$$

An integer vector $x \in \mathbf{Z}^n$ is called *primitive* if $x \neq ky$ for any $y \in \mathbf{Z}^n$ and $k \in \mathbf{Z}$ with $|k| \geq 2$. We denote by $\mathcal{P}(\mathbf{Z}^n)$ the set of all primitive integral vectors in $\mathbf{Z}^n$. The following strengthening of Theorem 3 was proved in 1988 by S.G.Dani and the author:

Theorem 4. *If* Q *and* n *are the same as in Theorem 1 and 3, then* $Q(\mathcal{P}(\mathbf{Z}^n))$ *is dense in* $\mathbf{R}$.

In the same way as Theorem 1 is deduced from Theorem 2, Theorem 4 is deduced from the following:

Theorem 5. *Let* Q *be a real indefinite nondegenerate quadratic form in* $n \geq 3$ *variables. Let us denote by* H *the special orthogonal group* $SO(Q)$. *Then any orbit of* H *in* $SL(n, \mathbf{R})/SL(n, \mathbf{Z})$ *either is closed and carries an* H*-invariant probability measure or is dense.*

Let us go now from quadratic forms to a different topic. Around 1930, Littlewood stated the following:

Conjecture 1. *Let* $\langle x \rangle$ *denote the distance between* $x \in \mathbf{R}$ *and the closest integer. Then*

$$\liminf_{n \to \infty} n \langle n\alpha \rangle \langle n\beta \rangle = 0$$

for any real numbers α and β.

We already mentioned the paper [CS] by Cassels and Swinnerton-Dyer in connection with earlier discussion about quadratic forms. But they also consider another type of form, namely products of three linear forms in 3 variables. In particular, they show that the Littlewood conjecture will be proved if the following conjecture is proved for $n = 3$.

Conjecture 2. *Let L be the product of n linearly independent linear forms on $\mathbf{R}^n$. Suppose that $n \geq 3$ and L is not a multiple of a form with rational coefficients. Then for every $\varepsilon > 0$, there exist integers $x_1, \dots, x_n$ not all equal to 0 such that $|L(x_1, \dots, x_n)| < \varepsilon$.*

Conjecture 2 can be considered as an analogue of Theorem 1. As in Theorem 1 and because of the same example, in this conjecture the condition $n \geq 3$ cannot be replaced by the condition $n \geq 2$. Conjecture 2 turns out to be equivalent to the following conjecture about orbits of the diagonal subgroup in $SL(n, \mathbf{R})/SL(n, \mathbf{Z})$. For $n = 3$, this equivalence was essentially noticed in [CS].

Conjecture 3. *Let $n \geq 3$, $G = SL(n, \mathbf{R})$, $\Gamma = SL(n, \mathbf{Z})$, and let A denote the group of all positive diagonal matrices in G. If $z \in G/\Gamma$ and the orbit Az is relatively compact in G/Γ, then Az is closed.*

In the paper [EKL] published in 2006 and submitted in 2003, M.Einsiedler, A.Katok and E.Lindenstrauss made dramatic progress on Littlewood's conjecture by showing that the set

$$\{(\alpha, \beta) \in \mathbf{R}^2 : \liminf_{n \to \infty} n \langle n\alpha \rangle \langle n\beta \rangle \neq 0\}$$

of exceptions to Littlewood's conjecture has Hausdorff dimension zero. They deduce this result from the classification of the measures on $SL(n, \mathbf{R})/SL(n, \mathbf{Z})$ which are invariant and ergodic under the action of the group A of positive diagonal matrices with positive entropy.

§2. A quantitative version of the Oppenheim conjecture.

In the previous section, we described some applications to number theory based on topological behavior of orbits of points in $SL(n, \mathbf{R})/SL(n, \mathbf{Z})$ under the action of a subgroup of $SL(n, \mathbf{R})$. In this and the following two sections, we will give examples of number theoretic applications based on results about the distribution of translates of submanifolds of $SL(n, \mathbf{R})/SL(n, \mathbf{Z})$.

Let ρ be a positive continuous function on the sphere $\{v \in \mathbf{R}^n : \|v\| = 1\}$, and let $\Omega = \{v \in \mathbf{R}^n : \|v\| < \rho(v/\|v\|)\}$. We denote by $T\Omega$ the dilate of Ω by T. For an indefinite quadratic form Q in $n \geq 3$ variables, let us denote by $N_{Q,\Omega}(a,b,T)$ the cardinality of the set

$$\{x \in \mathbf{Z}^n : x \in T\Omega \text{ and } a < Q(x) < b\}$$

and by $V_{Q,\Omega}(a,b,T)$ the volume of the set

$$\{x \in \mathbf{R}^n : x \in T\Omega \text{ and } a < Q(x) < b\}.$$

It is easy to verify that, as $T \to \infty$

$$V_{Q,\Omega}(a,b,T) \sim \lambda_{Q,\Omega}(b-a)T^{n-2},$$

where

$$\lambda_{Q,\Omega} = \int\limits_{L \cap \Omega} \frac{dA}{\|\nabla Q\|},$$

L is the light cone $Q = 0$ and dA is the area element on L.

Let $\mathcal{O}(p,q)$ denote the space of quadratic forms of signature (p,q) and discriminant ± 1, let (a,b) be an interval. In [DM], S.G.Dani and the author proved the following theorem which essentially gives the asymptotically exact lower bound for $N_{Q,\Omega}(a,b,T)$.

Theorem 6. *(I) Let $p \geq 2$ and $q \geq 1$. Then for any irrational form $Q \in \mathcal{O}(p,q)$ and any interval (a,b),*

$$\liminf_{T\to\infty} \frac{N_{Q,\Omega}(a,b,t)}{V_{Q,\Omega}(a,b,T)} \geq 1.$$

Moreover, this bound is uniform over compact sets of forms: if $\mathcal{K}$ is a compact subset of $\mathcal{O}(p,q)$ which consists of irrational forms, then

$$\liminf_{T\to\infty} \inf_{Q\in\mathcal{K}} \frac{N_{Q,\Omega}(a,b,t)}{V_{Q,\Omega}(a,b,T)} \geq 1.$$

(II) If $p > 0, q > 0$ and $n = p + q \geq 5$, then for any $\varepsilon > 0$ and any compact subset $\mathcal{K}$ of $\mathcal{O}(p,q)$ there exists $c = c(\varepsilon, \mathcal{K}) > 0$ such that for all $Q \in \mathcal{K}$ and $T > 0$,

$$N_{Q,\Omega}(-\varepsilon,\varepsilon,T) \geq cV_{Q,\Omega}(-\varepsilon,\varepsilon,T).$$

Remark. The condition $n \geq 5$ in statement (II), is related to the same condition in Meyer's theorem.

The situation with the asymptotics and upper bounds for $N_{Q,\Omega}(a,b,T)$ is more delicate. Rather surprisingly, here the answer depends on the signature of Q. In [EMM1], A.Eskin, S.Mozes and the author proved the following Theorems 7-10.

Theorem 7. *If $p \geq 3, q \geq 1$ and $n = p + q$, then as $T \to \infty$*

$$(*) \qquad N_{Q,\Omega}(a, b, T) \sim \lambda_{Q,\Omega}(b - a)T^{n-2}$$

for any irrational form $Q \in \mathcal{O}(p, q)$.

If the signature of Q is $(2, 1)$ or $(2, 2)$, then no universal formula like (*) in Theorem 7 holds. In fact we have the following:

Theorem 8. *Let Ω be the unit ball, and let $q = 1$ or 2. Then for every $\varepsilon > 0$ and every interval (a, b) there exists an irrational form $Q \in \mathcal{O}(2, q)$ and a constant $c > 0$ such that for an infinite sequence $T_j \to \infty$*

$$N_{Q,\Omega}(a, b, T) > cT_j^q(\log T_j)^{1-\varepsilon}.$$

The case $q = 1, b \leq 0$ of this theorem was noticed by Sarnak and worked out by Brennan in his undergraduate thesis. The quadratic forms constructed are of the type $x_1^2 + x_2^2 - \alpha x_3^2$, or $x_1^2 + x_2^2 - \alpha(x_1^2 + x_4^2)$ where α is extremely well approximated by squares of rational numbers. Another point is that in the statement of Theorem 8, $(\log T)^{1-\varepsilon}$ can be replaced by $\log T/\nu(T)$ where $\nu(T)$ is any unbounded increasing function.

However, in the $(2, 1)$ and $(2, 2)$ cases, there is an upper bound of the form $cT^q \log T$. This upper bound is uniform over compact subsets $\mathcal{K}$ of $\mathcal{O}(p, q)$, and it is effective in the sense that there is an effective way to compute the constant c. There is also an effective upper bound for the case $p \geq 3$.

Theorem 9. *Let $\mathcal{K}$ be a compact subset of $\mathcal{O}(p, q)$ and $n = p + q$. Then, if $p \geq 3$ and $q \geq 1$, there exists a constant $c = c(\mathcal{K}, a, b, \Omega)$ such that for any $Q \in \mathcal{K}$ and all $T > 1$*

$$N_{Q,\Omega}(a, b, T) < cT^{n-2}.$$

If $p = 3$ and $q = 1$ or $q = 2$, then there exists a constant $c = c(\mathcal{K}, a, b, \Omega)$ such that for any $Q \in \mathcal{K}$ and all $T > 2$

$$N_{Q,\Omega}(a, b, T) < cT^{n-2} \log T.$$

Also, for the $(2, 1)$ and $(2, 2)$ cases, the following "almost everywhere" result is true:

Theorem 10. *The asymptotic formula (*) from Theorem 7 holds for almost all quadratic forms of the signature $(2, 1)$ or $(2, 2)$.*

We will now briefly describe how Theorems 6-10 are proved. Let $G = SL(n, \mathbf{R}), \Gamma = SL(n, \mathbf{Z}), \Omega_n = G/\Gamma = \{$the space of unimodular lattices in $\mathbf{R}^n$ with determinant $1\}$.

One can associate to an integrable function f on $\mathbf{R}^n$, a function $\tilde{f}$ on Ω_n by setting

$$\tilde{f}(\Delta) = \sum_{v\in\Delta, v\neq 0} f(v), \Delta \in \Omega_n.$$

According to a theorem of Siegel,

$$\int_{\mathbf{R}^n} f dm^n = \int_{\Omega_n} \tilde{f} d\mu,$$

where m^n is the Lebesgue measure on $\mathbf{R}^n$ and μ is the G-invariant probability measure on $\Omega_n = G/\Gamma$. In [DM], the proof of Theorem 6 is based on the following identity which is immediate from the definitions:

$$(A) \qquad \int_T^{\ell T}\int_F \sum_{v\in g\mathbf{Z}^n} f(u_t k v) d\sigma(k) dt = \int_T^{\ell T}\int_F \tilde{f}(u_t k g\Gamma) d\sigma(k) dt,$$

where $\{u_t\}$ is a certain one-parameter unipotent subgroup of $SO(p,q)$, F is a Borel subset of the maximal compact subgroup K of $SO(p,q)$, σ is the normalized Haar measure on K, and f is a continuous compactly supported function on $\mathbf{R}^n\backslash\{0\}$. The number $N_{Q,\Omega}(a,b,T)$ can be approximated by the sum over m of the integrals on the left hand side of (A) for an appropriate choice of $g, f = f_i, F = F_i, 1 \le i \le m$. The right hand side of (A) can be estimated, uniformly when $g\Gamma$ belongs to certain compact subsets of G/Γ, using the just mentioned theorem of Siegel and some results on the equidistribution for unipotent flows (for details see §5). The function $\tilde{f}$ on Ω_n is unbounded for any nonnegative nonzero continuous function f on $\mathbf{R}^n$. But the abovementioned results on the equidistribution for unipotent flows are proved (and in general true) only for bounded continuous functions. On the other hand, as was done in [DM], one can get lower bounds by considering bounded continuous functions $\varphi \le f$ and applying these results to f.

Let $p \ge 2,\ p \ge q,\ q \ge 1$. We denote $p+q$ by n. Let $\{e_1,\dots,e_n\}$ be the standard basis of $\mathbf{R}^n$. Let Q_0 be the quadratic form of signature (p,q) defined by

$$Q_0(\sum_{i=1}^n v_i e_i) = 2v_1 v_n + \sum_{i=2}^p v_i^2 - \sum_{i=p+1}^{n-1} v_i^2$$

for $v_1,\dots,v_n \in \mathbf{R}$. Let $H = SO(Q_0)$. For $t \in \mathbf{R}$, let a_t be the linear transformation so that $a_t e_1 = e^{-t} e_1$, $a_t e_n = e^t e_n$, and $a_t e_i = e_i$, $2 \le i \le n-1$. Let $\hat{K}$ be the subgroup of G consisting of orthogonal matrices, and let $K = H \cap \hat{K}$. It is easy to check that K is a maximal compact subgroup of H, and consists of all $h \in H$ leaving invariant the subspace spanned by $\{e_1+e_n, e_2, \ldots, e_p\}$. For technical reasons, we consider in [EMM1] not the identity (A) but another identity:

$$(B) \qquad \int_K \sum_{v \in g\mathbf{Z}^n} f(a_t k v)\nu(k)d\sigma(k) = \int_K \tilde{f}(a_t k g\Gamma)\nu(k)d\sigma(k),$$

where ν is a bounded measurable function on K.

Let Δ be a lattice in $\mathbf{R}^n$. We say that a subspace L of $\mathbf{R}^n$ is Δ-*rational* if $L \cap \Delta$ is a lattice in L. For any Δ-rational subspace L, we denote by $d(L)$ the volume of $L/L\cap\Delta$. Let us note that $d(L)$ is equal to the norm of $e_1 \wedge \ldots \wedge e_\ell$ in the exterior power $\bigwedge^\ell(\mathbf{R}^n)$ where $\ell = \dim L$ and $(e_1 \ldots, e_\ell)$ is a basis over $\mathbf{Z}$ of $L \cap \Delta$. If $L = \{0\}$ we write $d(L) = 1$. Let

$$\alpha(\Delta) = \sup\{\frac{1}{d(L)} : L \text{ is a } \Delta\text{-rational subspace}\}.$$

According to a well known fact from the geometry of numbers, sometimes called "Lipschitz principle", for any bounded compactly supported function f on $\mathbf{R}^n$, there exists a positive constant $c = c(f)$ such that $\tilde{f}(\Delta) < c\alpha(\Delta)$ for every lattice Δ in $\mathbf{R}^n$.

Theorems 7-9 are proved in [EMM1] by combining the abovementioned results on the equidistribution for unipotent flows with the identity (B), the Lipschitz principle and the following integrability estimates:

Theorem 11. *(a) If $p \geq 3, q \geq 1$ and $0 < s < 2$, or if $p = 2, q \geq 1$ and $0 < s < 1$, then for any lattice Δ in $\mathbf{R}^n$*

$$\sup_{t>0} \int_K \alpha(a_t k\Delta)^s d\sigma(k) < \infty.$$

(b) If $p = 2$ and $q = 2$, or if $p = 2$ and $q = 1$, then for any lattice Δ in $\mathbf{R}^n$

$$\sup_{t>1} \frac{1}{t} \int_K \alpha(a_t k\Delta) d\sigma(k) < \infty.$$

These upper bounds are uniform as Δ varies over compact sets in the space of lattices.

Theorem 10 is deduced from the identity (B), the Lipschitz principle, Howe-Moore estimates for matrix coefficients of unitary representations, and from the fact that the function α on the space $\Omega_n = L(n, \mathbf{R})/SL(n, \mathbf{Z})$ of unimodular lattices in $\mathbf{R}^n$ belongs to every $L^\mu, 1 \leq \mu \leq n$.

As was noticed before, there are quadratic forms Q of the type $x_1^2+x_2^2-\alpha x_3^2$ or $x_1^2+x_2^2-\alpha(x_3^2+x_4^2)$, where α is extremely well approximated by squares of rational numbers, such that the asymptotic formula (*) in Theorem 7 does not hold. These examples can be generalized by considering irrational forms of the signature $(2,1)$ or $(2,2)$ which are extremely well approximated by split (over $\mathbf{Q}$) rational forms. As Theorem 12 below shows, this generalization is essentially the only method of constructing such forms in the $(2,2)$ case.

Fix a norm $\|\cdot\|$ on the space $\mathcal{O}(2,2)$ of quadratic forms of the signature $(2,2)$. We say that a quadratic form $Q \in \mathcal{O}(2,2)$ is *extremely well approximable by split rational forms*, to be abbreviated as EWAS, if for any $N > 0$, there exist a split rational form Q' and a real number $\lambda > 2$ such that $\|\lambda Q - Q'\| \leq \lambda^{-N}$. It is clear that if the ratio of two nonzero coefficients of Q is Diophantine, then Q is not EWAS; hence the set of EWAS forms has Hausdorff dimension zero as a subset of $\mathcal{O}(2,2)$. (A real number α is called *Diophantine* if there exist $N > 0$ such that $|qx-p| > q^{-N}$ for any integers p and q, $|q| > 2$. All algebraic numbers are Diophantine.) In [EMM2], A.Eskin, S.Mozes and the author proved the following:

Theorem 12. *The asymptotic formula (*) in Theorem 7 holds if $Q \in \mathcal{O}(2,2)$ is not EWAS and $0 \notin (a,b)$.*

Remarks. (i) If a form $Q \in \mathcal{O}(2,2)$ has a rational 2-dimensional isotropic subspace L, then $L \cap T\Omega$ contains of the order of T^2 integral points x for which $Q(x)=0$, hence $N_{Q,\Omega}(-\varepsilon,\varepsilon) \geq cT^2$, independently of the choice of ε. This is exactly the reason why we assumed in Theorem 12 that $0 \notin (a,b)$. Let us also note that an irrational quadratic form of the signature $(2,2)$ may have at most 4 rational isotropic subspaces and that if Q is such a form, then the number of points in the set $\{x \in \mathbf{Z}^n : Q(x)=0, \|x\| < T, x \text{ is not contained in an isotropic (with respect to } Q) \text{ subspace}\}$ grows not faster than a linear function of T.

(ii) Thought it seems that an analogue of Theorem 12 should be true for forms of the signature $(2,1)$, it is not clear how the method of the proof of Theorem 12 can be extended to the $(2,1)$ case.

It was noted by Sarnak that the quantitative version of the Oppenheim conjecture in the $(2,2)$ case is related to the study of eigenvalue spacings of flat 2-tori. Let Δ be a lattice in $\mathbf{R}^2$ and let $M = \mathbf{R}^2/\Delta$ denote the associated flat torus. The eigenvalues of the Laplacian on M are the values of the binary quadratic form $q(m,n) = 4\pi^2\|mv_1+nv_2\|^2$, where $\{v_1,v_2\}$ is a $\mathbf{Z}$-basis for the dual lattices Δ^*. We label these eigenvalues (with multiplicity) by

$$0 = \lambda_0(M) < \lambda_1(M) \leq \lambda_2(M) \cdots$$

It is easy to see that Weyl's law holds, i.e.

$$|\{j : \lambda_j(M) \leq T\}| \sim c_M T$$

where $c_M = (\text{area} M)/4\pi$. We are interested in the so-called *pair correlation*

$$R_M(a, b, T) = \frac{|\{j \neq k : \lambda_j(M) \leq T, \lambda_k(M) \leq T, a \leq \lambda_j(M) - \lambda_k(M) \leq b\}|}{T}$$

Theorem 13. *Let M be a flat 2-torus rescaled so that one of the coefficients in the associated binary quadratic form q is 1. Let A_1, A_2 denote the two other coefficients of q. Suppose that there exists $N > 0$ such that for all triples of integers (p_1, p_2, r) with $r \geq 2$,*

$$\max_{i=1,2} |A_i - \frac{p_i}{q}| > \frac{1}{q^N}.$$

Then, for any interval (a, b) which does not contain 0,

$$(*\ .) \qquad \lim_{T\to\infty} R_M(a, b, T) = c_M^2(b - a)$$

In particular, if one of the A_i is Diophantine, then (*) holds, and therefore the set of $(A_1, A_2) \subset \mathbf{R}^2$ for which (*) does not hold has zero Hausdorff dimension.

This theorem is proved by applying Theorem 12 to the form $Q(m_1, n_1, m_2, n_2) = q(m_1, n_1) - q(m_2, n_2)$. It is not difficult to give the asymptotics of $R_M(a, b, T)$ also in the case when $0 \in (a, b)$ (under the same conditions on q). For this we have to study the multiplicity of eigenvalues $\lambda_i(M)$. This can be easily done if q is irrational, but it requires consideration of several cases. Note also that, for all $i > 0$, this multiplicity is at least 2.

The equality (*) in Theorem 13 is exactly what is predicted by the random number (Poisson) model. Sarnak showed that this equality holds on a set of full measure in the space of tori. But his method does not give any explicit example of such a torus. Let us also note that Theorem 13 is related to the Berry-Tabor conjecture that the distribution of the local spacings between eigenvalues of a completely integrable Hamiltonian is Poisson. Another result related to the Berry-Tabor conjecture is the following theorem, due to J.Marklof (see [Mark]), about the pair correlation of values of inhomogeneous quadratic forms.

Theorem 14. *Let $0 \leq \lambda_1 \leq \lambda_2 \leq \ldots$ be the infinite sequence given by the values of*

$$(m - \alpha)^2 + (n - \beta)^2$$

at lattice points $(m, n) \in \mathbf{Z}^2$ for fixed $\alpha, \beta \in [0, 1]$. For a given interval $[a, b] \subset \mathbf{R}$, let

$$R_2[a, b](\lambda) = \frac{|\{j \neq k : \lambda_j \leq \lambda, \lambda_k \leq \lambda, a \leq \lambda_j - \lambda_k \leq b\}|}{\pi\lambda}$$

denote the pair correlation function of this sequence. Suppose $\alpha, \beta, 1$ *are linearly independent over* $\mathbf{Q}$*, and assume* α *is Diophantine. Then*

$$\lim_{\lambda \to \infty} R_2[a,b](\lambda) = \pi(b-a).$$

The proof of Theorem 14 in [Mark] uses theta sums and some results on the equidistribution of translates of unipotent orbits.

§3. Counting lattice points on homogeneous varieties.

In this section we will mostly give an overview of the paper [EMS] by A.Eskin, S.Mozes and N.Shah published in 1996. About more recent results, we refer to a comprehensive survey [O] by Hee Oh.

Let V be a real algebraic subvariety of $\mathbf{R}^n$ defined over $\mathbf{Q}$, and let G be a reductive real algebraic subgroup of $GL(n, \mathbf{R})$ also defined over $\mathbf{Q}$. Suppose that V is invariant under G and that G acts transitively on V (or, more precisely, the complexification of G acts transitively on the complexification of V). Let $\|\cdot\|$ denote a Euclidean norm $\mathbf{R}^n$. Let B_T denote the ball of radius T in $\mathbf{R}^n$ around the origin, and define

$$N(T,V) = |V \cap B_T \cap \mathbf{Z}^n|,$$

the number of integral points in V with norm less than T. We are interested in the asymptotics of $N(T,V)$ as $T \to \infty$.

Let Γ denote $G(\mathbf{Z}) \stackrel{def}{=} \{g \in G : g\mathbf{Z}^n = \mathbf{Z}^n\}$. By a theorem of Borel and Harish-Chandra, $V(\mathbf{Z})$ is a union of finitely many Γ-orbits. Therefore to compute the asymptotics of $N(T,V)$, it is enough to consider each Γ-orbit, say $\mathcal{O}$, separately, and compute the asymptotics of

$$N(T,V,\mathcal{O}) = |\mathcal{O} \cap B_T|.$$

Of course, after that there is the problem, often non-trivial, of the summation over the set of Γ-orbits. This is essentially a problem from the theory of algebraic and arithmetic groups and is of completely different type than the computation of the asymptotics of $N(T,V,\mathcal{O})$.

Suppose that $\mathcal{O} = \Gamma v_0$ for some $v_0 \in V(\mathbf{Z})$. Then the stabilizer $H = \{g \in G : gv_0 = v_0\}$ is a reductive $\mathbf{Q}$-subgroup, and $V \cong G/H$. Define

$$R_T = \{gH \in G/H : gv_0 \in B_T\},$$

the pullback of the ball B_T to G/H.

Assume that H^0 and H^0 does not admit nontrivial $\mathbf{Q}$-characters, where G^0 (resp. H^0) denotes the connected component of identity in G (resp. in H). Then by a theorem of Borel and Harish-Chandra, G/Γ admits a G-invariant (Borel) probability measure, say μ_G, and $H/(\Gamma \cap H)$ admits an H-invariant probability measure, say μ_H. We can consider $H(\Gamma \cap H)$ as a (closed) subset of G/Γ; then μ_H can be treated as a measure on G/Γ supported on $H/(\Gamma \cap H)$. Let $\lambda_{G/H}$ denote the (unique) G-invariant measure on G/H induced by the normalization of the Haar measures on G and H. One of the main results in [EMS] is the following:

Theorem 15. *Suppose that H^0 is a maximal proper connected $\mathbf{Q}$-subgroup of G. Then asymptotically as $T \to \infty$,*

$$N(T, V, \mathcal{O}) \sim \lambda_{G/H}(R_T). \qquad (*)$$

Example. Let $p(\lambda)$ be a monic polynomial of degree $n \geq 2$ with integer coefficients and irreducible over $\mathbf{Q}$. Let $M_n(\mathbf{Z})$ denote the set of $n \times n$ integer matrices, and put

$$V_p(\mathbf{Z}) = \{A \in M_n(\mathbf{Z}) : \det(\lambda I - A) = p(\lambda)\}.$$

Thus $V_p(\mathbf{Z})$ is the set of integer matrices with characteristic polynomial $p(\lambda)$. Consider the norm on $n \times n$ matrices given by $\|(x_{ij})\| = \sqrt{\sum_{ij} x_{ij}^2}$.

Theorem 16. *Let $N(T, V_p)$ denote the number of points on $V_p(\mathbf{Z})$ with norm less than T. Then asymptotically as $T \to \infty$,*

$$N(T, V_p) \sim \; c_p T^{n(n-1)/2},$$

where $c_p > 0$ is an explicitly computable constant.

In the above example, the group G is $SL(n, \mathbf{R})$ which acts on the space $M(n, \mathbf{R})$ of $n \times n$ matrices by conjugation. In the case when $p(\lambda)$ splits over $\mathbf{R}$, and for a root α of $p(\lambda)$ the ring of algebraic integers in $\mathbf{Q}(\alpha)$ is $\mathbf{Z}[\alpha]$, the following formula for c_p is given in [EMS]:

$$c_p = \frac{2^{n-1} h R \omega_n}{\sqrt{D} \cdot \prod_{k=2}^{n} \Lambda(k/2)},$$

where h is the class number of $\mathbf{Z}[\alpha]$, R is the regulator of $\mathbf{Q}(\alpha)$, D is the discriminant of $p(\lambda)$, ω_n is the volume of the unit ball in $\mathbf{R}^{n(n-1)/2}$, and $\Lambda(s) = \pi^{-s}\Gamma(s)\zeta(2s)$.

For the case when V is affine symmetric, the asymptotic formula (*) in Theorem 15 had been earlier proved by Duke, Rudnik and Sarnak using harmonic analysis; subsequently a simpler proof using the mixing property of (generalizations of) the geodesic flow was given by Eskin and McMullen - a similar 'mixing property' approach had been also used in 1970 in the author's thesis to obtain asymptotic formulas for the number of closed geodesics in M^n and for the number of points from $\pi^{-1}(x)$ in balls of large radius in $\tilde{M}^n$, where M^n is a compact manifold of a negative curvature, $x \in M^n$, $\tilde{M}^n$ is the universal covering space of M^n and $\pi : \tilde{M}^n \to M^n$ is the natural projection.

The proof of Theorem 15 in [EMS] is based on the equidistribution properties of the translates of the measure μ_H. It turns out that if $\{g_i\} \subset G$ and the sequence $\{g_i H\}$ is divergent (that is, it has no divergent subsequences) in G/H, then the sequence $\{g_i \mu_H\}$ gets equidistributed with respect to μ_G as $i \to \infty$ (that is $g_i \mu_H \to \mu_G$ weakly).

§4. Diophantine approximation on manifolds

In this section we present an approach to metric Diophantine approximation on manifolds which uses the correspondence between approximation properties of numbers and orbit properties of certain homogeneous flows. We start by recalling several basic facts from the theory of simultaneous Diophantine approximation. For $\mathbf{x}, \mathbf{y} \in \mathbf{R}^n$, we let

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{n} x_i y_i, \ \ \|\mathbf{x}\| = \max_{1 \le i \le n} |x_i|,$$
$$\prod(\mathbf{x}) = \prod_{i=1}^{n} |x_i| \text{ and } \prod(\mathbf{x})_+ = \prod_{i=1}^{n} |x_i|_+,$$

where $|x|_+$ stands for $\max(|x|, 1)$. One says that a vector $\mathbf{y} \in \mathbf{R}^n$ is *very well approximable*, to be abbreviated as VWA, if the following two equivalent conditions are satisfied:

(V1) for some $\varepsilon > 0$ there are infinitely many $\mathbf{q} \in \mathbf{Z}^n$ such that

$$|\mathbf{q} \cdot \mathbf{y} + p| \cdot \|\mathbf{q}\|^n \le \|\mathbf{q}\|^{-n\varepsilon}$$

for some $p \in \mathbf{Z}$.

(V2) for some $\varepsilon > 0$ there are infinitely many $q \in \mathbf{Z}$ such that

$$\|q\mathbf{y} + \mathbf{p}\|^n \cdot |q| \le |q|^{-\varepsilon}$$

for some $\mathbf{p} \in \mathbf{Z}^n$.

A vector $\mathbf{y} \in \mathbf{R}^n$ is called *very well multiplicatively approximable*, to be abbreviated as VWMA, if the following two equivalent conditions are satisfied:

(VM1) for some $\varepsilon > 0$ there are infinitely many $\mathbf{q} \in \mathbf{Z}^n$ such that

$$|\mathbf{q} \cdot \mathbf{y} + p| \cdot \prod(\mathbf{q})_+ \leq \prod(\mathbf{q})_+^{-\varepsilon}$$

for some $p \in \mathbf{Z}$;

(VM2) for some $\varepsilon > 0$ there are infinitely many $q \in \mathbf{Z}$ such that

$$\prod(q\mathbf{y} + \mathbf{p}) \cdot |q| \leq |q|^{-\varepsilon}$$

for some $\mathbf{p} \in \mathbf{Z}^n$.

Remark. It is clear that if a vector is VWA, then it is also VWMA. The equivalence of (V1) and (V2) (resp. the equivalence of (VM1) and (VM2)) follows from (resp. from a modification of) Khintchine's transference principle.

It easily follows from (the simple part of) the Borel-Cantelli lemma that almost every $\mathbf{y} \in \mathbf{R}^n$ is not VWA. A more difficult question arises if one considers almost all points $\mathbf{y}$ on a submanifold M of $\mathbf{R}^n$ (in the sense of the natural measure class on M). In 1932, K.Mahler conjectured that almost all points on the curve

$$(*) \qquad M = \{(x, x^2, \ldots, x^n)\} : x \in \mathbf{R}\}$$

are not very well approximable. In 1964, V.Sprindžuk proved this conjecture using what is later become known as "the method of essential and inessential domains". According to Sprindžuk's terminology, a submanifold $M \subset \mathbf{R}^n$ is called *extremal* (resp. *strongly extremal*) if almost all $\mathbf{y} \in M$ are not VWA (resp. are not VWMA). Clearly any strongly extremal manifold is extremal. In his book on transcendental number theory published in 1975, A.Baker stated a conjecture about the strong extremality of the curve (*) from the Mahler conjecture. Later it was generalized by Sprindžuk in a survey paper published in 1980:

Conjecture. *Let $\mathbf{f} = (f_1, \ldots, f_n)$ be an n-tuple of real analytic functions on a domain V in $\mathbf{R}^d$ which together with 1 are linearly independent over $\mathbf{R}$. Then for almost all $\mathbf{x} \in V$ the vector $\mathbf{f}(\mathbf{x})$ is not VWMA.*

Remark.. In the same survey, Sprindžuk stated also a weaker version of this conjecture where VWMA is replaced by VWA.

The just stated conjecture was proved in 1997 by D.Kleinbock and the author (see [KM]) not only for analytic but also for smooth functions. To state our main result, we have to introduce the following definition: if V is an open subset of $\mathbf{R}^d$ and $\ell \leq k$, an n-tuple $\mathbf{f} = (f_1, \ldots, f_n)$ of C^k functions $V \to \mathbf{R}^d$ is called ℓ-*nondegenerate* at $\mathbf{x} \in V$ if the space $\mathbf{R}^n$ is spanned by partial derivatives of $\mathbf{f}$ at

$\mathbf{x}$ of order up to ℓ. The n-tuple $\mathbf{f}$ is *nondegenerate* at $\mathbf{x}$ if it is ℓ-nondegenerate at $\mathbf{x}$ for some ℓ. We say that $\mathbf{f} : V \to \mathbf{R}^n$ is *nondegenerate* if it is nondegenerate at almost every point of V. Note that if the functions $f_1, \dots, f_n$ are analytic and V is connected, the nondegeneracy of $\mathbf{f}$ is equivalent to the linear independence of $1, f_1, \dots, f_n$ over $\mathbf{R}$. The main result in [KM] is the following:

Theorem 17. *Let $\mathbf{f} : V \to \mathbf{R}^n$ be a nondegenerate C^k map of an open subset V of $\mathbf{R}^d$ into $\mathbf{R}^n$. Then $\mathbf{f}(\mathbf{x})$ is not VWMA (hence not VWA either) for almost every point $\mathbf{x} \in \mathbf{R}^n$.*

If $M \subset \mathbf{R}^n$ is a d-dimensional C^k submanifold, we say that M is *nondegenerate* at $\mathbf{y} \in M$ if any (equivalently some) C^k diffeomorphism $\mathbf{f}$ between an open subset V of $\mathbf{R}^d$ and a neighborhood of $\mathbf{y}$ in M is nondegenerate at $\mathbf{f}^{-1}(\mathbf{y})$. We say that M is *nondegenerate* if it is nondegenerate at almost every point of M (in the sense of the natural measure class on M). A connected analytic submanifold $M \subset \mathbf{R}^n$ is nondegenerate if and only if it is not contained in any hyperplane in $\mathbf{R}^n$. Now we can reformulate Theorem 17.

Theorem 17′. *Let M be a nondegenerate C^k submanifold of $\mathbf{R}^n$. Then almost all points of M are not VWMA (hence not VWA either).*

The proof of Theorem 17 (or Theorem 17′) is based in [KM] on a method which uses the correspondence, originally introduced by S.G.Dani in mid-1980's, between approximation properties of vectors $\mathbf{y} = (y_1, \dots, y_n) \in \mathbf{R}^n$ and the behavior of certain orbits in the space of unimodular lattices in $\mathbf{R}^n$. More precisely, let

$$U_{\mathbf{y}} = \begin{pmatrix} 1 & y_1 & y_2 & \dots & y_n \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & & 1 \end{pmatrix} \in SL(n+1, \mathbf{R}).$$

Thus $U_{\mathbf{y}}$ is a unipotent matrix with all rows, except the first one, the same as in the identity matrix. Note

$$U_{\mathbf{y}} \begin{pmatrix} p \\ \mathbf{q} \end{pmatrix} = \begin{pmatrix} \mathbf{q} \cdot \mathbf{y} + p \\ \mathbf{q} \end{pmatrix}, p \in \mathbf{Z}, q \in \mathbf{Z}^n.$$

We also have to introduce some diagonal matrices. Let

$$g_s = \begin{pmatrix} e^{ns} & 0 & \dots & 0 \\ 0 & e^{-s} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & e^{-s} \end{pmatrix} \in SL(n+1, \mathbf{R}), s \geq 0$$

and

$$g_{\mathbf{t}} = \begin{pmatrix} e^t & 0 & \dots & 0 \\ 0 & e^{-t_1} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & e^{-t_n} \end{pmatrix} \in SL(n+1, \mathbf{R}),$$

$$\mathbf{t} = (t_1, \dots, t_n),\ t_i \geq 0,\ t = \sum_{i=1}^{n} t_i;$$

(it is clear that $g_s = g_{(s,\dots s)}$). Next define a function δ on the space of lattices by

$$\delta(\Delta) \stackrel{def}{=} \inf_{\mathbf{v}\in\Delta\setminus\{0\}} \|\mathbf{v}\|,\ \text{where } \|(v_1, \dots, v_{n+1})\| = \max\{|v_i| : 1 \leq i \leq n+1\}.$$

Note the ratio of $1 + \log(1/\delta(\Delta))$ and $1 + \operatorname{dist}(\Delta, \mathbf{Z}^{n+1})$ is bounded between two positive constants for any metric "dist" on the space $SL(n+1, \mathbf{R})/SL(n+1, \mathbf{Z})$ of unimodular lattices Δ in $\mathbf{R}^{n+1}$ induced by a right invariant Riemannian metric on $SL(n+1, \mathbf{R})$. It is easy to check that

$$\delta(g_{\mathbf{t}} U_{\mathbf{y}} \mathbf{Z}^{n+1}) =$$
$$= \inf_{(p,\mathbf{q})\in\mathbf{Z}^{n+1}\setminus\{0\}} \min\{e^t(\mathbf{q}\cdot\mathbf{y} + p),\ e^{-t_1}q_1, \dots, e^{-t_n}q_n\},$$

where $\mathbf{q} = (q_1, \dots, q_n)$. From this, one an easily get that a vector $\mathbf{y} \in \mathbf{R}^n$ is VWA (resp. VWMA) if and only if there exists $\gamma > 0$ and infinitely many $t \in \mathbf{Z}_+$ (resp. infinitely many $\mathbf{t} \in \mathbf{Z}_+^n$) such that

$$(**) \qquad \delta(g_t U_{\mathbf{y}} \mathbf{Z}^{n+1}) \leq e^{-\gamma t} (\text{resp. } \delta(g_{\mathbf{t}} U_{\mathbf{y}} \mathbf{Z}^{n+1}) \leq e^{-\gamma\|\mathbf{t}\|})$$

in other words, $\mathbf{y}$ is not VWA (resp. not VWMA) if and only if

$$\operatorname{dist}(g_t U_{\mathbf{y}} \mathbf{Z}^{n+1}, \mathbf{Z}^{n+1})\ (\text{resp. } \operatorname{dist}(g_{\mathbf{t}} U_{\mathbf{y}} \mathbf{Z}^{n+1}, \mathbf{Z}^{n+1}))$$

as a function of $t \in \mathbf{Z}^+$ (resp. as a function of $\mathbf{t} \in \mathbf{Z}_+^n$), grows slower than any linear function. Thus Theorem 17 is equivalent to the statement that for almost all $\mathbf{x} \in V$ and any $\gamma > 0$, there are at most finitely many $\mathbf{t} \in \mathbf{Z}_+^n$ such that (**) holds for $\mathbf{y} = \mathbf{f}(\mathbf{x})$. In view of the Borel-Cantelli lemma, this statement can be proved by estimating the measure of the sets

$$E_{\mathbf{t}} \stackrel{def}{=} \{\mathbf{x} \in V : \delta(g_{\mathbf{t}} U_{\mathbf{f}(\mathbf{x})} \mathbf{Z}^{n+1}) \leq e^{-\gamma\|t\|}\}$$

for any given $\mathbf{t} \in \mathbf{Z}_+^n$, so that

$$\sum_{\mathbf{t}\in\mathbf{Z}_+^n} |E_{\mathbf{t}}| < \infty,$$

where $|\cdot|$ stands for the Lebesgue measure. Such estimates are easily deduced in [KM] from the following:

Theorem 18. *Let* $\mathbf{f} : V \to \mathbf{R}^n$ *be a* C^k *map of an open subset* V *of* $\mathbf{R}^d$ *into* $\mathbf{R}^n$*, and let* $\mathbf{x}_0 \in V$ *be such that* $\mathbf{R}^n$ *is spanned by partial derivatives of* $\mathbf{f}$ *at* $\mathbf{x}_0$ *of order up to* k*. Then there exists a ball* $B \subset U$ *centered at* $\mathbf{x}_0$*, and positive constants* D *and* ρ *such that for any* $\mathbf{t} \in \mathbf{R}^n_+$ *and* $0 \le \varepsilon \le \rho$ *one has*

$$|\{\mathbf{x} \in B : \delta(g_{\mathbf{t}} U_{\mathbf{f}(\mathbf{x})} \mathbf{Z}^{n+1}) \le \varepsilon\}| \le D(\frac{\varepsilon}{\rho})^{1/dk} |B|.$$

Theorem 18 is deduced in [KM] from a more general theorem (Theorem 22 in §5).

There are also other applications of the homogeneous space approach to Diophantine approximation on manifolds, in particular to Khintchine-type theorems on manifolds (Bernik, Kleinbock and the author) and to Diophantine properties of almost all points with respect to so called "friendly measures" (D.Kleinbock, E.Lindenstrauss and Barak Weiss). Let us describe in more details one more application.

There is the following multiplicative form of Dirichlet's theorem on simultaneous Diophantine approximation as formulated by Minkowski:

Given $\mathbf{y} = (y_1, \dots, y_n) \in \mathbf{R}^n$ *and positive integers* $N_1, \dots N_n$*, there exist integers* $q_1, \dots, q_n$ *and* p*, not all zero, such that*

$$|q_1 y_1 + \dots + q_n y_n - p| \le (N_1, \dots, N_n)^{-1}, |q_i| < N \ (1 \le i \le n).$$

Following the terminology introduced by Davenport and Schmidt, we say that given any infinite set $\mathcal{N} \subset \mathbf{Z}^n_+$, the Dirichlet's theorem (DT) cannot be improved along $\mathcal{N}$ for $\mathbf{y} = (y_1, \dots, y_n) \in \mathbf{R}^n$ if, for every $0 < \mu < 1$, there are infinitely many $(N_1, \dots, N_n) \in \mathcal{N}$ such that the following system of inequalities insoluble for integers $q_1, \dots, q_n$ and p, not all zero:

$$|q_1 y_1 + \dots + q_n y_n - p| \le \mu(N_1 \dots N_n)^{-1},$$
$$|q_i| < \mu N_i \ (1 \le i \le n).$$

In the late sixties, Davenport and Schmidt showed that for $\mathcal{N} = \{(N, \dots, N) \in \mathbf{Z}^n : N \in \mathbf{Z}_+\}$, the Dirichlet's theorem cannot be improved along $\mathcal{N}$ for almost all points of $\mathbf{R}^n$. The same conclusion was obtained recently by D.Kleinbock and Barak Weiss for sets $\mathcal{N} \subset \mathbf{Z}^n$ with infinite projection on each coordinate. In the preprint [S], N.Shah proves the following:

Theorem 19. *Let* $\mathcal{N}$ *be an infinite subset of* $\mathbf{Z}^n_+$*. Then for any analytic curve* $\mathbf{f} : [a, b] \to \mathbf{R}^n$ *whose image is not contained in a proper affine subspace, the Dirichlet's theorem cannot be improved along* $\mathcal{N}$ *for* $\mathbf{f}(x)$ *for almost every* $x \in [a, b]$*.*

The proof of Theorem 19 in [S] is based on the asymptotic equidistribution of the curves

$$\{g_{\mathbf{t}} U_{\mathbf{f}(x)} \mathbf{Z}^{n+1} : a \leq x \leq b, \mathbf{t} \in \mathcal{N}\}$$

in the space Ω_{n+1} of unimodular lattices in $\mathbf{R}^{n+1}$; these curves are obviously the translations by $g_{\mathbf{t}}, \mathbf{t} \in \mathcal{N}$, of the curve

$$\{U_{\mathbf{f}(x)} \mathbf{Z}^{n+1} : a \leq x \leq b\}.$$

§5. Translates of submanifolds and unipotent flows

Let G be a connected Lie group, Γ a discrete subgroup of G and Y a (smooth) submanifold of G/Γ. In previous sections, in connection with problems in Diophantine approximation and number theory, we essentially tried to answer in some cases the following general question:

(Q) What is the distribution of gY in G/Γ when g tends to infinity in G?

This question can be divided into two subquestions:

(Q1) What is the behavior of gY "near infinity" of G/Γ?

(Q2) What is the distribution of gY in the "bounded part" of G/Γ?

In §2 and implicitly in the part of §1 related to quadratic forms, G, Γ, g and Y were, respectively, $SL(n, \mathbf{R}), SL(n, \mathbf{Z}), a_t$ and $K\Delta$, where $\Delta \in \Omega_n \cong SL(n, \mathbf{R})/SL(n, \mathbf{Z})$ and K is a maximal compact subgroup of $SO(p, q)$. In §3, G, Γ and Y were, respectively a real reductive $\mathbf{Q}$-subgroup of $GL(n, \mathbf{R})$, $G(\mathbf{Z})$ and $H/(\Gamma \cap H)$, where H is a reductive $\mathbf{Q}$-subgroup of G. In §4 we had the following:

$$G = SL(n+1, \mathbf{R}), \Gamma = SL(n+1, \mathbf{Z}), g = g_{\mathbf{t}}, \text{ and}$$
$$Y = \{U_{\mathbf{f(x)}} \mathbf{Z}^{n+1} : x \in V\}.$$

Theorems 1-3 are related only to the question (Q2), and Theorem 9 is related only to the question (Q1). But other statements about quadratic forms and orbits of orthogonal groups (Theorem 4-7, 10 and 12) are related to both questions (Q1) and (Q2). Theorems 15 and 16 in §3 are also related to both questions (Q1) and (Q2). In §4, Theorems 17 and 18 are related only to the question (Q1), and Theorem 19 is related to both questions (Q1) and (Q2).

Let $y \in Y$ and let W be a 'small" neighborhood of y in Y. Then for any $w \in W$ there exists an element $h \in G$ that is "close" to 1, such that $w = hy$. It is clear that $gw = (ghg^{-1})gy$ and that $\mathrm{Ad}(ghg^{-1})$ has the same eigenvalues as Ad h where Ad denotes the adjoint representation of G. Hence gW consists of translates of gy by "almost" Ad-unipotent elements ($x \in G$ is called Ad-*unipotent* if Adx is unipotent; that is, after all eigenvalues of Adx are equal to 1). This is exactly the reason why results and methods from the theory of unipotent flows on homogeneous spaces play such an important role in the study of the distribution properties of translates of submanifolds. We will now state some of these results.

Theorem 20. *Let G be a connected Lie group, Γ a lattice (i.e. a discrete subgroup with finite covolume) in G, F a compact subset of G/Γ and $\varepsilon > 0$. Then there exists a compact subset K of G/Γ such that for any Ad-unipotent one-parameter subgroup $\{u(t)\}$ of G, any $x \in F$, and $T \geq 0$,*

$$|\{t \in [0, T] : u(t)x \in K\}| \geq (1 - \varepsilon)T.$$

This theorem is essentially due to S.G.Dani. He proved it in the mid-eighties, separately for semisimple groups G of $\mathbf{R}$-rank 1 and for arithmetic lattices. The general case can be easily reduced to these two cases using the arithmeticity theorem. In the case of arithmetic lattices, Theorem 20 can be considered as the quantitative version of the following:

Theorem 21. *Let $\Omega_n \cong SL(n, \mathbf{R})/SL(n, \mathbf{Z})$ denote the space of unimodular lattices in $\mathbf{R}^n$, and let $\{u(t)\}$ be a unipotent one-parameter subgroup of $SL(n, \mathbf{R})$. Then, for any lattice $\Delta \in \Omega_n$, $u(t)\Delta$ does not tend to infinity in Ω_n as $t \to +\infty$ or, equivalently,*

$$\limsup_{t \to +\infty} \delta(u(t)\Delta) \to 0,$$

where the function δ on Ω_n is defined in §4.

Theorem 21 in the case $n = 2$ and Theorem 20 in the case $G/\Gamma = \Omega_2$ can be easily proved. The proof is based on the following two facts:

(i) For every $r > 1$ one can find $\varepsilon = \varepsilon(r, n) > 0$ such that if $v \in \mathbf{R}^n$, $v \neq 0, a > 0$ and

$$\|u(t)v\| < \varepsilon \text{ for all } t \in [-a, a]$$

then

$$\|u(t)v\| < 1 \text{ for all } t \in [-ra, ra].$$

(ii) If $n = 2$, $\Delta \in \Omega_2, v_1, v_2 \in \Lambda - \{0\}$ and $\|v_i\| < \frac{1}{2}, i = 1, 2$ then v_1 and v_2 are proportional.

But for $n > 2$, the proof of Theorem 21 is much more complicated because an analog of (ii) is no longer true for $n > 2$. The main idea of the proof for $n > 2$ is to study $\{u(t)\}$-orbits not only of (primitive) vectors $v \in \Delta$ but also of primitive subgroups of Δ. (We say that a subgroup Λ of Δ is *primitive* (in Δ) if $\Lambda = \Lambda_{\mathbf{R}} \cap \Delta$ where $\Lambda_{\mathbf{R}}$ denotes the linear span of Δ.) Let us also note that the statement (i) is related to the *polynomial behavior* or *polynomial divergence* of unipotent flows.

The proof of Theorem 20 for arithmetic lattices, given by S.G.Dani, is similar to the proof of Theorem 21. Further modifications and refinements of these proofs eventually lead to a general Theorem 22 below. Before stating this theorem, we have to introduce some notation and terminology.

Let V be a subset of $\mathbf{R}^d$ and f a continuous function on V. We write $\|f\|_B \stackrel{def}{=} \sup_{x\in B} |f(x)|$ for a subset B of V. For positive numbers C and α, say that f is (C,α)-*good on* V if for any open ball $B \subset V$ and any $\varepsilon > 0$ one has

$$|\{x \in B : |f(x)| < \varepsilon\}| \le C \cdot \left(\frac{\varepsilon}{\|f\|_B}\right)^{\alpha} \cdot |B|.$$

A model example of good functions are polynomials: for any $k \in \mathbf{Z}_+$, any polynomial $f \in \mathbf{R}[x]$ of degree not greater than k is $(2k(k+1)^{1/k}, 1/k)$-good on $\mathbf{R}$.

For a discrete subgroup Λ of $\mathbf{R}^n$, we let $j = \dim(\Lambda_{\mathbf{R}})$ and say that $w \in \bigwedge^j(\mathbf{R}^n)$ *represents* Λ if $j > 0$ and $w = v_1 \wedge \ldots \wedge v_j$ where $v_1, \ldots, v_j$ is a basis of Λ. Clearly the element representing Λ is defined up to a sign. Therefore it makes sense to define the norm $\|\Lambda\| \stackrel{def}{=} \|w\|$ where w represents Λ.

Let us denote by $P(\Delta)$ the set of all nonzero primitive subgroups of $\Delta \in \Omega_n$. Another piece of notation which we need is $B(\mathbf{x}, r)$ which will stand for the open ball of radius $r > 0$ centerer at $\mathbf{x}$.

Theorem 22. *(see* [KM]*) Let $d, n \in \mathbf{Z}_+, C, \alpha > 0, 0 < \rho < 1/n$ and let a ball $B = B(\mathbf{x}_0, r_0) \subset \mathbf{R}^d$ and a map $h : \tilde{B} \to GL(n, \mathbf{R})$ be given, where $\tilde{B}$ stands for $B(\mathbf{x}_0, 3^n r_0)$. For any $\Lambda \in P(\mathbf{Z}^n)$, denote by ψ_Λ the function $\psi_\Lambda(\mathbf{x}) \stackrel{def}{=} \|h(\mathbf{x})\Lambda\|$, $x \in \tilde{B}$. Assume that for any $\Lambda \in P(\mathbf{Z}^n)$,*

(i) ψ_Λ is (C,α)-good on $\tilde{B}$;

(ii) $\|\psi_\Lambda\|_B \ge \rho$.

Then for any positive $\varepsilon \le \rho$ one has

$$|\{\mathbf{x} \in B : \delta(h(\mathbf{x})\mathbf{Z}^n) < \varepsilon\| \le R \left(\frac{\varepsilon}{\rho}\right)^{\alpha} |B|,$$

where $R = nC(3^d N_d)^n$ and N_d is an integer (from Besicovitch's Covering Theorem) depending only on d.

Remarks. (a) In [KM], we consider norms on the exterior products $\Lambda^j(\mathbf{R}^n)$ which are different from Euclidean norms. Using those norms it is easier to check (i) in the formulation of Theorem 22.

(b) Theorem 18 is deduced in [KM] from Theorem 22 using some standard facts from the geometry of numbers and some standard estimates for differentiable functions.

In a 1981 paper, S.G.Dani formulated two conjectures. One is Raghunathan's conjecture which states that if G is a connected Lie group, Γ a lattice in G, and U an Ad-unipotent subgroup of G, then for any $x \in G/\Gamma$ there exists a closed connected subgroup $L = L(x)$ containing U such that the closure of the orbit Ux coincides with Lx. The second conjecture is due to Dani himself and may be stated (in a slightly stronger form) as follows. If G, Γ and U are as above and μ is a finite Borel U-ergodic U-invariant measure on G/Γ, then there exists a closed subgroup F of G such that μ is F-invariant and supp $\mu = Fx$ for some $x \in G/\Gamma$ (a measure for which this condition holds is called *algebraic*). In the same paper, Dani proved his conjecture in the case where G is reductive and U is a maximal horospherical subgroup of G. In another paper, published in 1986, Dani proved Raghunathan's conjecture in the case where G is reductive and U is an arbitrary horospherical subgroup of G. These results of S.G.Dani generalize earlier results by Hedlund, Furstenberg, Bowen, Veech, Ellis and Perrizo. In two papers, published in 1987 and 1989, A.N.Starkov proved Raghunathan's conjecture for solvable G. We remark that the proof given by Dani in the just mentioned 1986 paper is restricted to horospherical U and the proof given by Starkov cannot be applied if G is not solvable.

The first result on Raghunathan's conjecture for nonhorospherical subgroups of semisimple groups was obtained in 1989 by S.G. Dani and the author. We proved the conjecture in the case where $G = SL(3, \mathbf{R})$ and $U = \{u(t)\}$ is a one-parameter unipotent subgroup of G such that $u(t) - 1$ has rank 2 for all $t \neq 0$. Though this is only a very special case, our proof together with the methods developed to prove Theorems 2 and 5 suggests an approach for proving the Raghunathan conjecture in general. In particular, it was used by N.Shah about a year ago, to prove the S-arithmetic generalization of Raghunathan's conjecture for the case where G is the product $\prod_{j=1}^{m} SL(2, K_j), 1 \leq j \leq m$, where K_j is a local field of characteristic zero. This approach is topological and is based on the technique which involves finding orbits of larger subgroups inside closed sets invariant under unipotent subgroups by studying the minimal invariant sets, and the limits of orbits of sequences of points converging to a minimal invariant set.

A far reaching generalization of Theorems 2 and 5 and other abovementioned results on orbit closures was obtained by M.Ratner in 1990 in the paper [R2] published in 1991. She proved the following:

Theorem 23. *Let G be a connected Lie group, Γ a lattice in G, and H a connected subgroup of G that is generated by the Ad-unipotent one-parameter subgroups contained in it. Then for any $x \in G/\Gamma$, there exists a closed connected subgroup*

$L = L(x)$ containing H such that the closure of the orbit Hx coincides with Lx and there is an L-invariant probability measure supported on Lx.

Theorem 23 settles (a generalization of) the Raghunathan conjecture. In [R2], this theorem is deduced from a rather simple result about the countability of a certain (depending on Γ) set of subgroups of G and from the following equidistribution theorem.

Theorem 24. *If G and Γ are the same as in Theorem 23, $\{u(t)\}$ is a one-parameter Ad-unipotent subgroup of G, and $x \in G/\Gamma$, then the orbit $\{u(t)x\}$ is equidistributed with respect to an algebraic probability measure μ_x on G/Γ in the sense that for any bounded continuous function f on G/Γ,*

$$\frac{1}{T}\int_0^T f(u(t)x)dt \to \int_{G/\Gamma} f d\mu_x \text{ as } T \to \infty.$$

The proof of Theorem 24 in [R2] uses Theorem 25 below together with Theorem 20 and the just mentioned countability result.

Theorem 25. *Let G and H be the same as in Theorem 23, and let Γ a discrete subgroup of G (not necessarily a lattice). Then any finite H-ergodic H-invariant measure μ on G/Γ is algebraic in the sense that there exists a closed subgroup F of G such that μ is F-invariant and supp $\mu = Fx$ for some $x \in G/\Gamma$.*

Theorem 25 settles (a generalization of) the Dani conjecture. It is a fundamental result with numerous applications. Theorem 25 was proved by M.Ratner in a series of three papers, the last of which is [R1]. The total length of Ratner's proof is more than 150 pages. We refer also to a shorter proof given by Tomanov and the author in [MT] for the crucial case where G is algebraic. The proof in [MT] bears a strong influence of Ratner's argument but is substantially different in approach and methods.

In order to prove Theorem 6, S.G.Dani and the author obtain in [DM] a refined version of the equidistribution Theorem 24. The proof of this version in [DM] uses, as in [R2], Theorems 20 and 25. The reduction to these theorems is based in [DM] on a countability result and on a variation of the following assertion: Let H be a connected closed subgroup of G such that $H \cap \Gamma$ is a lattice in H, and let $X(H, U) = \{g \in G : Ug \subset gH\}$, where $U = \{u(t)\}$ and G, Γ and $\{u(t)\}$ are the same as in Theorem 24. Then the subset $X(H, U)/\Gamma$ of G/Γ is *avoidable with respect to* $\{u(t)\}$ in the following sense: for any compact subset C of $(G/\Gamma)\backslash X(H, U)$ and any $\varepsilon > 0$ there exists a neighborhood Ψ of $X(H, U)$ in G/Γ such that for any

$x \in C$ and $t_1 \leq 0 \leq t_2$, we have

$$|\{t \in [t_1, t_2] : T^t x \in \Psi\}| \leq \varepsilon(t_2 - t_1).$$

The proof of this assertion in [DM] is based on what is now called the "linearization" technique.

Remarks. (i) The polynomial divergence of unipotent flows, mentioned in this section (after Theorem 21), was one of the main motivations for M.S.Raghunathan when he formulated his conjecture. He hoped that unipotent flows are likely to have "manageable behavior" because of the slow divergence of orbits of unipotent one-parameter subgroups (in contrast to the exponential divergence of orbits of diagonalizable subgroups). It should also be noted that the polynomial divergence plays a basic role in the theory of unipotent flows.

(ii) About S-arithmetic generalizations of theorems on unipotent flows, stated in this section, see [MT] and [R3]. It should be noted that the study of these S-arithmetic generalizations was inspired by a work of A.Borel and G.Prasad on values of isotropic quadratic forms at S-integral points.

References

[CS] J.W.S.Cassels and H.P.F.Swinnerton-Dyer, *On the product of three homogeneous forms and indefinite ternary quadratic forms*, Philos. Trans. Roy. Soc. London **248** Ser. A (1955), 73-96.

[DM] S.G.Dani and G.Margulis, *Limit distribution of orbits of unipotent flows and values of quadratic forms*, Adv. Soviet Math., **16**, 91-137.

[EKL] M.Einsiedler, A.Katok and E.Lindenstrauss, *Invariant measures and the set of exceptions to Littlewood's conjecture*, Ann. of Math. (2) **164** (2006), 513-560.

[EMM1] A.Eskin, G.Margulis and S.Mozes, *Upper bounds and asymptotics in a quantitative version of the Oppenheim conjecture*, Ann. of Math. (2)**147** (1998), 93-141.

[EMM2] ———, *Quadratic forms of signature* $(2,2)$ *and eigenvalue spacings on rectangular* 2-*tori*, Ann. of Math. (2) **161** (2005), 679-725.

[EMS] A.Eskin, S.Mozes and N.Shah, *Unipotent flows and counting lattice points on homogeneous varieties*, Ann. of Math. (2), **143** (1996), 253-299.

[G] F.Götze, *Lattice point problems and values of quadratic forms*, Invent. Math., **157** (2004), 195-226.

[GT] G.Margulis nad G.Tomanov, *Invariant measures for actions of unipotent groups over local fields on homogeneous spaces*, Invent. Math., **116** (1994), 347-392.

[KM] D.Kleinbock and G.Margulis, *Flows on homogeneous spaces and Diophantine approximation on manifolds*, Ann. of Math., (2), **148** (1988), 339-360.

[Marg1] G. Margulis, *Oppenheim conjecture*, In Fields Medalists' Lectures, World Scientific (1997), 272-327.

[Marg2] ———, *Diophantine approximation, lattices and flows on homogeneous spaces*, A panorama of number theory, Cambridge University Press (2002), 280-310.

[Mark] J.Markof, *Pair correlation densities of inhomogeneous quadratic forms*, Ann. of Math., (2) **158** (2003), 419-471.

[O] H. Oh, *Orbital counting via mixing and unipotent flows*, Preprint.

[R1] M.Ratner, *On Raghunathan's measure conjecture*, Ann. of Math., (2) **134** (1991), 545-607.

[R2] ———, *Raghunathan's topological conjecture and distribution of unipotent flows*, Duke Math. J., **63** (1991), 235-280.

[R3] ———, *Raghunathan's conjectures for Cartesian products of real and p-adic groups*, Duke Math. J., **77** (1995), 275-382.

[S] N.Shah, *Expanding translates of curves and Dirichlet-Minkowski theorem on linear forms.*