

QUANTITATIVE ERGODIC THEOREMS AND THEIR NUMBER-THEORETIC APPLICATIONS

ALEXANDER GORODNIK AND AMOS NEVO

ABSTRACT. We present an account of some recent applications of ergodic theorems for actions of algebraic and arithmetic groups to the solution of natural problems in Diophantine approximation and number theory. Our approach is based on spectral methods utilizing the unitary representation theory of the groups involved. This allows the derivation of ergodic theorems with a rate of convergence, an important phenomenon which does not arise in classical ergodic theory. Combining spectral and dynamical methods, quantitative ergodic theorems give rise to new and previously inaccessible applications. We demonstrate the remarkable diversity of such applications by deriving general uniform error estimates in non-Euclidean lattice points counting problems, explicit estimates in the sifting problem for almost-prime points on symmetric varieties, best-possible bounds for exponents of intrinsic Diophantine approximation on homogeneous algebraic varieties, and quantitative results on fast distribution of dense orbits on compact and non-compact homogeneous spaces.

CONTENTS

1. Introduction	65
2. Averaging in ergodic theory	70
3. From amenable to non-amenable groups	74
4. Spectral gaps and ergodic theorems	77
5. Counting lattice points	80
6. Almost prime points on homogeneous algebraic varieties	84
7. Intrinsic Diophantine approximation on homogeneous varieties	89
8. Ergodic theorems for lattice subgroups	94
9. Distribution of orbits on algebraic varieties	97
10. Recent developments: best possible Diophantine exponents	103
Acknowledgments	105
About the authors	105
References	106

1. INTRODUCTION

The two great classical ergodic theorems, namely von Neumann's mean ergodic theorem and Birkhoff's pointwise ergodic theorem, were motivated in large part by

Received by the editors April 25, 2013.

2010 *Mathematics Subject Classification*. Primary 37A15, 37P55, 22E46, 11J83, 11F70.

The first author was supported in part by EPSRC, ERC, and RCUK.

The second author was supported by an ISF grant.

Boltzmann's fundamental *ergodic hypothesis*, asserting that the time evolution of a physical dynamical system in phase space distributes it evenly in the space. The practical meaning of the ergodic hypothesis is the following remarkable property. Under the sole assumption of ergodicity, namely the absence of invariant sets, when one samples the values of an observable quantity on phase space at regular time intervals along the trajectory of the dynamical system, the average of these samples over time will converge to the space average of the observed quantity over the entire phase space. Equivalently, in the long run the trajectories will spend the right fraction of time in every subset of phase space, given by the measure of the subset.

The importance of these fundamental results was immediately recognized, and the study of the long term statistical properties of trajectories of general flows and measure-preserving transformations has been vigorously pursued ever since. The classical ergodic theorems have been applied and utilized in a remarkably wide array of different contexts, a consequence of the fact that measure-preserving flows and measure-preserving transformations are truly ubiquitous, from statistical physics to dynamics and geometry, and from probability theory to Diophantine approximation and number theory.

As we will note in more detail below, one of the most pleasant and elegant features of the study of dynamical systems defined by a measure-preserving flow or transformation is the possibility of applying arguments based on the existence of asymptotically invariant subsets in the real line or the integers. Asymptotic invariance arguments are crucial in the proofs of many of the main foundational results in ergodic theory. The list includes not only proofs of Birkhoff's pointwise ergodic theorem and von Neumann's mean ergodic theorem, where a particularly short and transparent proof was given by Riesz. It includes also the Krylov–Bogolyubov theorem on the existence of probability measures invariant under a continuous transformation of a compact space, as well as Furstenberg's correspondence principle. It also includes Wiener's covering argument in Euclidean space generalizing the Hardy–Littlewood maximal inequality on the real line, as well as Calderon's transference principle reducing aspects of analysis of the orbits of the flow on phase space to analysis on the group defining the flow, namely the real line or the integers.

These important and useful results pertaining to the dynamics of measure-preserving flows and transformations served as a very compelling motivation for developing the theory further, to encompass general groups of measure-preserving transformations acting on compact manifolds, or more generally arbitrary probability spaces. A major line of development in ergodic theory has been to generalize the arguments mentioned above to the class of all groups which possess an asymptotically invariant sequence of sets. We will give a brief exposition of some of the relevant arguments and indicate the appropriate references below, but let us highlight here two interesting aspects of this story, which will serve as our point of departure.

The first aspect is obvious: Clearly, the reliance of the classical methods of ergodic theory on asymptotic invariance arguments necessarily constrains their range of validity to the class of amenable groups, namely those groups possessing an asymptotically invariant sequence. While this class is extensive, it certainly excludes some very natural classes of groups, for example the groups $\mathrm{SL}_n(\mathbb{R})$ and $\mathrm{SL}_n(\mathbb{Z})$. How then does one approach the challenge of extending ergodic theory beyond amenable groups? The second aspect is yet another consequence of asymptotic

invariance arguments: they are inextricably tied to the impossibility of establishing a rate of convergence in the ergodic theorems, as we will explain in more detail below. Is the absence of rate an intrinsic feature of ergodic theorems for general group actions?

The present survey is devoted to an exposition of some developments aimed precisely at transcending these two fundamental constraints. We will describe results which develop and extend ergodic theorems beyond the class of amenable groups, and concentrate particularly on results which establish ergodic theorems with a rate of convergence. The key to the method we describe is, in a sense, a return to von Neumann's original approach to the mean ergodic theorem. Namely, we base our arguments on spectral theory, and thus the main challenge is to develop mean and pointwise ergodic theorems using just spectral methods alone, without any asymptotic invariance or transference arguments to lean on. Such methods have been developed by the authors for an extensive class of groups, including simple algebraic groups and their lattice subgroups. They combine a range of techniques, including aspects of the unitary representation theory of simple algebraic groups, the theory of matrix coefficients and spectral bounds in the automorphic representation, spectral techniques originating in classical singular integrals theory such as square functions and analytic interpolation, as well as geometric analysis on homogeneous spaces and elements of the structure theory of simple algebraic groups. However, here we will not delve into the technical aspects of these arguments, and the reader is referred instead to [GN10] for full details.

Rather, we will briefly describe some mean and pointwise ergodic theorems for actions of algebraic groups and their lattice subgroups, but we devote most of our exposition to discussing some of the many applications that such ergodic theorems have. We will focus below solely on results which have been recently developed by the authors, and which constitute an expansion of the range of applications of spectral methods in ergodic theory, particularly in the context of dynamics on homogeneous space. The survey is therefore limited in scope, describing only the authors' specific spectral approach, and it does not discuss other approaches within the very large and rapidly expanding field of homogeneous dynamics and its application in number theory. In particular, we will not discuss actions of diagonalizable or unipotent subgroups nor techniques associated with mixing or entropy.

To demonstrate the utility and diversity of applications of ergodic theorems for actions of algebraic groups and lattice subgroups, let us briefly describe a selection of counting, equidistribution, and Diophantine approximation problems that may at first sight appear completely unrelated. One of the main rationales of our exposition in the present survey is the fact, which will be explained in detail below, that solutions to all of the following problems depend on applying a suitable ergodic theorem in a suitable action that is appropriate to the case under consideration.

1. *The uniform lattice point counting problem.* Let Γ be a lattice, namely a discrete subgroup of a locally compact group G such that the space G/Γ carries a finite G -invariant measure. The lattice point counting problem is to count the number of elements of Γ in an increasing family of bounded domains $B_t \subset G$, and ideally to establish an asymptotic formula

$$|\Gamma \cap B_t| = \frac{|B_t|}{|G/\Gamma|} + O(|B_t|^\kappa)$$

with $\kappa < 1$ as small as possible. One classical instance of this question is the Gauss circle problem of counting integral points in the Euclidean plane, and below we will consider a variety of natural non-Euclidean lattice point counting problems. For example, what is the asymptotics of the number of unimodular integral matrices in a norm ball as the radius tends to infinity, or in more general increasing domains in matrix space? What is the asymptotics for the number of such matrices A satisfying a congruence condition $A = A_0 \pmod{q}$? More generally, what is the asymptotics of the number of unimodular matrices with entries in the ring of integers of an algebraic number field which are of bounded height?

Counting lattice points is in fact the most basic of all mean ergodic theorems that we consider. This point of view will be explained in full in Section 5, where we will describe an explicit solution to the general lattice point counting problem. The argument gives the best currently known bound for lattices in simple algebraic groups of split rank at least two.

2. *Almost prime points on algebraic varieties.* As we shall see below, establishing robustness and uniformity properties of the solution to the lattice point counting problem is crucial in several important applications which depend on answering more refined counting questions. Of those, let us mention the problem of existence of *almost prime* integral points on homogeneous varieties, raised in [NS10] and [BGS10]. Consider for example the set of symmetric integral positive-definite matrices of fixed determinant, and call such a matrix prime if all its entries are prime numbers. How large is the set of prime points in the set of integral points, for example, does it satisfy an analog of the prime number theorem? Less ambitiously, how large is the set of integral points all of whose entries are r -almost primes (i.e., products of at most r prime factors)? For example, is the set of r -almost prime points Zariski dense for some r ? and if so, can a meaningful bound be given for the saturation number, namely the least r with the latter property?

This question has been the subject of intensive activity in recent years, and we refer to [BGS10] and the recent surveys [Lu12], [K2], [K3], and [K] for further discussion. In Section 6 below we will utilize aspects of uniformity in the solution of the lattice point counting problem to establish an asymptotic lower bound of the right magnitude for the number of r -prime points lying on symmetric varieties and to provide an effective explicit bound on the saturation number.

3. *Diophantine approximation on algebraic varieties.* Consider the rational ellipsoid $\{(x, y, z); ax^2 + by^2 + cz^2 = d\}$, $a, b, c, d \in \mathbb{N}_+$. Assuming that the rational points on the ellipsoid are dense, is it possible to establish a rate of approximation of general points on the ellipsoid by rational points? Is it possible to establish a rate of approximation by rational points satisfying additional integrality constraints?

These problems, in the much more general context of homogeneous algebraic varieties, were raised already half a century ago in S. Lang's 1965 *Report on Diophantine approximation* [L65]. Except for the case of Abelian groups, very little progress has been obtained on them since then. We will state this problem precisely in Section 7 and describe a general solution, obtained jointly with A. Ghosh, in the case where the variety is homogeneous under an action of a simple algebraic group. For the rational ellipsoids we establish a rate of approximation by rational points satisfying integrality constraints that is in fact best possible in a natural sense.

In Section 10 we will describe best possible results for rates of Diophantine approximation by general lattice orbits on homogeneous varieties, generalizing the case of lattices of S -integral or rational points.

4. *Distribution of typical orbits on homogeneous spaces.* Consider the one-sheeted rational hyperboloid $\{(x, y, z); ax^2 + by^2 - cz^2 = d\}$, $a, b, c, d \in \mathbb{N}_+$, also known as de-Sitter space. The group of integral matrices preserving the corresponding quadratic form naturally acts on its level sets, and thus on de-Sitter space. It is a basic fact that typical orbits are dense, but does a typical orbit have a limiting distribution in a suitable sense?

This problem, for general lattices in the isometry group of a Lorentz form acting on de-Sitter space of arbitrary dimension, was raised by V. Arnold in 1985 [Ar]. A precise quantitative solution for general homogeneous varieties will be described in Section 9, and it offers a number of surprises. As we shall see, in the present context, the limiting distribution can be non-invariant under the group action, and in fact may depend on the orbit under consideration. This is in sharp contrast to the phenomena usually encountered in the classical context of amenable groups.

5. *Fast equidistribution in algebraic number fields.* The group given by $\Gamma = \mathrm{SL}_2(\mathbb{Z}[\sqrt{d}])$, where $d > 0$ is not a square, is a dense subgroup of $G = \mathrm{SL}_2(\mathbb{R})$. Consider elements of Γ of bounded norm: $\Gamma_t = \{\gamma \in \Gamma : \log(\|\gamma\|^2 + \|\bar{\gamma}\|^2) \leq t\}$, where $\bar{\gamma}$ denotes matrix whose entries are the Galois conjugates of the entries of γ . Do these finite sets become equidistributed in G in a suitable sense, and if so, do they equidistribute at a uniform rate?

We note that the action of Γ on G by translations is isometric, and this fact naturally raises the much more ambitious problem of proving fast equidistribution for *every single orbit*, i.e., every coset of Γ . One possible such statement is a ratio equidistribution theorem, establishing that for nice functions ϕ and ψ of compact support, with $\psi \geq 0$ and $\int_G \psi(g)dg > 0$, and for *every* $x \in G$

$$\lim_{t \rightarrow \infty} \frac{\sum_{\gamma \in \Gamma_t} \phi(x\gamma)}{\sum_{\gamma \in \Gamma_t} \psi(x\gamma)} = \frac{\int_G \phi(g) dg}{\int_G \psi(g) dg},$$

with convergence taking place at an explicit and effective uniform rate. Such a result will be stated precisely and explained in Section 9.

We remark that the problem of establishing a ratio equidistribution theorem for dense subgroups acting by isometries on the group or its homogeneous spaces was raised originally half a century ago by Kazhdan [K65], in the case of a dense group of Euclidean isometries acting on the Euclidean plane. We refer to [V] for recent progress in the case of the isometry groups of Euclidean spaces, but we note that in the non-amenable case no precedents seem to exist.

6. *Ergodic theorems for lattice subgroups.* Finally, let us mention some natural questions which constitute direct generalizations of the fundamental convergence results of classical ergodic theory. Taking the group $\Gamma = \mathrm{SL}_n(\mathbb{Z})$ as an example, consider its linear action on the torus $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$. Does the mean ergodic theorem hold for the action of Γ on \mathbb{T}^n ? Specifically, let $\Gamma_t = \{\gamma \in \Gamma; \log \|\gamma\| \leq t\}$, and for $f \in L^p(\mathbb{T}^n)$ and $x \in \mathbb{T}^n$, consider the averages

$$\frac{1}{|\Gamma_t|} \sum_{\gamma \in \Gamma_t} f(\gamma^{-1}x).$$

Do these averages converge to the space average $\int_{\mathbb{T}^n} f(x) dx$ in L^p ? Does convergence hold almost everywhere? If so, is there a rate of convergence to the space average, in L^p and almost everywhere?

These questions also arise for other interesting actions of $\mathrm{SL}_n(\mathbb{Z})$, for instance, for the action on the space of unimodular lattices or for the action on its profinite completion. They are all resolved by a general pointwise and mean ergodic theorem for the averages on norm balls in $\mathrm{SL}_n(\mathbb{Z})$, a result that will be formulated more generally and precisely in Section 8 and explained there.

To conclude the introduction, let us emphasize that it is ultimately the existence of an explicit effective rate of convergence in the ergodic theorems that is responsible for an explicit quantitative solution of the counting, equidistribution and Diophantine problems described above. Therefore, considerable attention will be given below to the problem of establishing an effective rate of convergence of the averages we consider to the space average in measure-preserving dynamical systems, and estimating it as accurately as possible. This effort is amply motivated by the fact that in a number of instances, including the intrinsic Diophantine approximation problem discussed in Section 7 and more generally the approximation problem by dense orbits on homogeneous spaces discussed in Section 10, the exponents we obtain are in many cases best possible.

In order to demonstrate both the power and the limitations of asymptotic invariance arguments alluded to above and to situate our spectral approach in its proper context, we now turn to describe in more detail some aspects of the development of the mean ergodic theorem and of asymptotic invariance arguments in classical ergodic theory that were mentioned above. We will also briefly describe along the way some of the milestones reached using applications of spectral methods in ergodic theory.

2. AVERAGING IN ERGODIC THEORY

2.1. The classical mean ergodic theorem. The basic objects of study in classical Hamiltonian mechanics consist of a compact Riemannian manifold M , which is called the phase space, together with a divergence-free vector field on M . We denote by μ the Riemannian volume on M , normalized to be a probability measure. The integral curves of the vector field give rise to a one-parameter group of volume-preserving transformations $a_s : M \rightarrow M$, $s \in \mathbb{R}$, which describes the time evolution in phase space. For a function $f : M \rightarrow \mathbb{R}$, the time averages along the orbits are defined by

$$\mathcal{A}_t f(x) = \frac{1}{t} \int_0^t f(a_s x) ds.$$

One of the most fundamental questions in classical dynamics is the elucidation of their long term behavior. Two important questions that arise are as follows.

- Do the time averages $\mathcal{A}_t f$ converge at all, and if so, what is their limit?
- How does convergence to the limit depend on the initial state x and the function f ?

The solution proposed by von Neumann to this problem was to consider for every measure-preserving map $a_s : M \rightarrow M$ an associated unitary transformation, the Koopman operator $U_s : L^2(M) \rightarrow L^2(M)$, given by $U_s f(x) = f(a_s x)$. Thus

the measure-preserving flow gives rise to a one-parameter group of unitary transformations. This made it possible for von Neumann to apply his recently developed spectral theory to prove the mean ergodic theorem [vN32]:

Theorem 2.1 (von Neumann). *Let (X, μ) be a (standard) probability space with a one-parameter flow of measure-preserving transformations $a_s : X \rightarrow X$, $s \in \mathbb{R}$. Then for any $f \in L^2(X)$,*

$$\mathcal{A}_t f \rightarrow \mathcal{P}f \quad \text{as } t \rightarrow \infty$$

in L^2 -norm, where \mathcal{P} denotes the orthogonal projection on the space of functions invariant under the flow.

In particular, consider the case where the space of invariant functions in $L^2(M)$ consists of constant functions only, or equivalently, every invariant measurable set has measure zero or one. Flows satisfying this property are called *ergodic*. Then the orthogonal projection of f to the space of constant functions is given by the *space average* $\int_X f d\mu$. Thus, for ergodic flows the time averages converge in L^2 -norm to the space average

$$\mathcal{A}_t f \rightarrow \int_X f d\mu \quad \text{as } t \rightarrow \infty,$$

verifying, in particular, Boltzmann's ergodic hypothesis for phase space.

Several years later, F. Riesz [R38] gave an elegant elementary proof of the mean ergodic theorem, which became very influential in the subsequent development of ergodic theory. We will indicate it here in full, in order to highlight its main idea. Let $U_s : \mathcal{H} \rightarrow \mathcal{H}$, $s \in \mathbb{R}$, be a one-parameter group of unitary operators on a Hilbert space \mathcal{H} , and let \mathcal{P} denote the projection on the space of vectors invariant under the group. We will show that the averaging operators $\mathcal{A}_t = \frac{1}{t} \int_0^t U_s ds$ satisfy $\mathcal{A}_t f \rightarrow \mathcal{P}f$ for every $f \in \mathcal{H}$ by breaking \mathcal{H} up to two orthogonal complementary subspaces and arguing separately in each of them. First, we apply \mathcal{A}_t to vectors f of the form $f = U_{t_0} h - h$. Then

$$\begin{aligned} \mathcal{A}_t f &= \mathcal{A}_t(U_{t_0} h - h) = \frac{1}{t} \left(\int_{t_0}^{t+t_0} U_s h ds - \int_0^t U_s h ds \right) \\ &= \frac{1}{t} \left(\int_t^{t+t_0} - \int_0^{t_0} \right) U_s h ds, \end{aligned}$$

and therefore

$$\|\mathcal{A}_t f\| \leq \frac{2t_0}{t} \|h\| \rightarrow 0 \quad \text{as } t \rightarrow \infty.$$

Since f is orthogonal to every invariant vector, namely $\mathcal{P}f = 0$, the mean ergodic theorem holds for f as stated, hence also for the closed linear span of such f 's. Obviously, the mean ergodic theorem holds in the space of invariant vectors. The proof is therefore complete upon observing that the span of $U_{t_0} h - h$ with $t_0 \in \mathbb{R}$ and $h \in \mathcal{H}$ is dense in the orthogonal complement of the space of invariant vectors.

Notice that the main ingredient of the above proof is the property of asymptotic invariance (under translation) of the intervals $[0, t]$, namely the fact that the measure of the symmetric difference of $[0, t]$ and $[0, t] + t_0$ divided by the measure of $[0, t]$ converges to zero at $t \rightarrow \infty$. This crucial property of intervals on the real line admits a straightforward generalization, as follows. Let G be a locally compact second countable group G equipped with a left-invariant Haar measure. Define a

one-parameter family of sets $F_t \subset G$ of positive finite measure to be *asymptotically invariant* if it satisfies

$$(2.1) \quad \frac{|F_t g \Delta F_t|}{|F_t|} \rightarrow 0 \quad \text{as } t \rightarrow \infty$$

uniformly for g in compact sets of G . A group admitting such an asymptotically invariant family (also called a *Følner family*) is called *amenable*.

Given a measure-preserving action of the group G on a probability space (X, μ) , define for $g \in G$ the unitary Koopman operators

$$\pi_X(g) : L^2(X) \rightarrow L^2(X) : f \mapsto f(g^{-1}x).$$

It is straightforward to adapt Riesz's argument and conclude that amenable groups satisfy the mean ergodic theorem. Namely, denoting by β_t the uniform probability measures supported on an asymptotically invariant (Følner) family F_t and by $\pi_X(\beta_t)$ the corresponding averaging operators in $L^2(X)$, we conclude that

$$\pi_X(\beta_t)f = \frac{1}{|F_t|} \int_{F_t} \pi_X(g)f dg \rightarrow \mathcal{P}f \quad \text{as } t \rightarrow \infty$$

in L^2 -norm.

2.2. Equidistribution and pointwise convergence. So far we have discussed the convergence of the averages $\pi_X(\beta_t)f$ in L^2 -norm. The question as to the convergence $\pi_X(\beta_t)f(x)$ for individual points $x \in X$ is of fundamental importance as well.

We begin by describing the best-case scenario, namely equidistribution, and its characterization in amenable dynamics. Let X be a compact metric space on which G acts continuously, and let $F_t \subset G$ be a family of Følner sets that define the corresponding averaging operators $\pi_X(\beta_t)$. An extreme possibility is that for every continuous function f and for every point $x \in X$, the *time averages* $\pi_X(\beta_t)f(x)$ converge uniformly to a limit $m(f)$, which is independent of x . In that case, the functional $f \mapsto m(f)$ defines a probability measure m on X , which is invariant under G . Furthermore m is then the unique G -invariant probability measure on X , and the system is called *uniquely ergodic*. Conversely, when there exists a unique G -invariant probability measure on X , the time averages of f converge uniformly to the space average $\int_X f dm$.

In general the situation is far more complicated, and the limit of the averages $\pi_X(\beta_t)f(x)$ depends very sensitively on the initial point x . Typically, there exist two (or even uncountably many) mutually singular invariant probability measures μ_1 and μ_2 on X , such that the following holds. The time averages $\pi_X(\beta_t)f(x)$ of a continuous function f converge on a set X_1 of full μ_1 -measure to $\int_X f d\mu_1$, and on a set X_2 of full μ_2 -measure to $\int_X f d\mu_2$, with the sets X_1 and X_2 disjoint and $\int_X f d\mu_1 \neq \int_X f d\mu_2$. The question of pointwise convergence for one-parameter flows is addressed by Birkhoff's pointwise ergodic theorem [B31]:

Theorem 2.2 (Birkhoff). *Let (X, μ) be a (standard) probability space with a one-parameter flow of measure-preserving transformations $a_s : X \rightarrow X$, $s \in \mathbb{R}$. Then for any $f \in L^1(X)$, the time averages $\frac{1}{t} \int_0^t f(a_s x) ds$ converge for μ -almost every $x \in X$ to $\mathcal{P}f(x)$.*

An elementary proof of Birkhoff's theorem which relies on asymptotic invariance of intervals as well as their order structure was given by [KW82]. So far, this

proof has not been generalized to other groups, including \mathbb{Z}^2 . The most successful approach to the proof of pointwise ergodic theorems relies on controlling the maximum deviation between the time averages and the space average throughout the entire sampling process. Thus we consider the following fundamental quantity, called the *maximal function* associated to the flow:

$$\mathcal{M}f = \sup_{t>0} |\pi_X(\beta_t)f - \mathcal{P}f|.$$

In order to establish pointwise convergence, we seek to prove the following two properties of the flow:

- I. Existence of a dense set of functions $f \in L^1(X)$ for which $\pi_X(\beta_t)f(x)$ converges for almost every $x \in X$.
- II. The maximum error in the measurement performed by time averages is controlled by the *strong maximal inequality* in $L^2(X)$

$$\|\mathcal{M}f\|_2 \leq \text{const} \cdot \|f\|_2$$

or by the *weak maximal inequality* in $L^1(X)$

$$\mu(\{\mathcal{M}f > r\}) \leq \text{const} \cdot r^{-1} \|f\|_1.$$

Once these properties are established, it follows by a classical approximation argument known as the Banach principle that the time averages converge almost everywhere for every $f \in L^1(X)$.

The proof of the first property proceeds by using asymptotic invariance of a Følner family, which implies that pointwise convergence holds for the family of functions $\pi_X(g)f - f$ with $f \in L^\infty(X)$ and $g \in G$. This is a straightforward adaptation of Riesz's argument described above.

The proof of the maximal inequality utilizes two separate ideas: the *transference principle* and *covering arguments*. The transference principle asserts that the maximal inequality for any family of operators $\pi_X(\beta_t)$ acting on a general space (X, μ) can be reduced to the maximal inequality for the convolution operators defined by β_t on the group G . This principle was initiated by Wiener [W39], and generalized by Calderon [C53] and Coifman and Weiss [CW76]. Finally, the investigation of the convolution operators exploits *geometric covering properties* of the translates of sets in a Følner family F_t . This approach originated with Wiener [W39], who has generalized the Hardy–Littlewood maximal inequality from the real line to Euclidean space by introducing an elegant and influential covering argument based on the *doubling property*

$$|B_{2t}| \leq \text{const} \cdot |B_t|$$

for volumes of the Euclidean balls B_t .

The pointwise ergodic theorem was subsequently established, in turn, for groups with polynomial volume growth, exponential solvable Lie groups, and connected amenable Lie groups. This was carried out by Templeman [T67], Emerson and Greenleaf [EG74], Coifman and Weiss [CW76], and others. It should be pointed out, however, that a Følner sequence may fail to satisfy the pointwise ergodic theorem, even for the group \mathbb{Z} . The pointwise ergodic theorem for general amenable groups requires a regularity assumption stronger than asymptotic invariance, namely temperedness, and was established by Lindenstrauss in [L01]. We refer to [N06] for a recent survey and to [AAB] for a detailed account of the ergodic theory of amenable groups.

3. FROM AMENABLE TO NON-AMENABLE GROUPS

The proofs of the ergodic theorems indicated in the previous section were founded upon the asymptotic invariance property (2.1) of the Følner families. How then can one proceed when the group is not amenable so that there are no asymptotically invariant families of sets at all?

Before describing our approach, let us note one more important appearance of the notion of asymptotic invariance, this time in probability measure preserving actions of the group. For any properly ergodic (i.e., ergodic but not transitive) action of G , a sequence of measurable sets $A_n \subset X$ is called a non-trivial asymptotically invariant sequence in X if

$$0 < \liminf_{n \rightarrow \infty} \mu(A_n) \leq \limsup_{n \rightarrow \infty} \mu(A_n) < 1 \quad \text{and} \quad \mu(gA_n \Delta A_n) \rightarrow 0$$

uniformly for g in compact subsets of G . It then follows easily that there exists an *asymptotically invariant sequence* of functions $f_n \in L^2(X)$ with zero integral such that for every $g \in G$

$$(3.1) \quad \|f_n\| = 1 \quad \text{and} \quad \|\pi_X(g)f_n - f_n\| \rightarrow 0$$

uniformly for g in compact subsets of G .

When G is countable and amenable, a sequence of asymptotically invariant subsets of the group can be used to construct a non-trivial asymptotically invariant sequence of sets in every properly ergodic action of G . Moreover, the existence of an asymptotically invariant sequence in every properly ergodic action characterizes amenable groups in this case (see [S81, JR79]).

We can therefore conclude that a countable group G is amenable if and only if in every properly ergodic action, the averaging operators satisfy

$$(3.2) \quad \left\| \pi_X(\beta)|_{L_0^2(X)} \right\| = 1$$

for every probability measure β on G , where $L_0^2(X)$ denotes the orthogonal complement to the space of constant functions. In particular, it follows that for a properly ergodic action of a countable amenable group, *no rate of convergence* in the mean ergodic theorem can be established; namely, the norm of the averaging operators in $L_0^2(X)$ does not decay at all.

3.1. Spectral gap property. Our discussion so far has lead us to the following observation: When G is a countable non-amenable group, at least some actions X have the property that the operators $\pi_X(\beta)$ are strict contractions on $L_0^2(X)$, namely $\|\pi_X(\beta)\|_{L_0^2(X)} < 1$, for some generating probability measures β on G . We can therefore conclude that a non-amenable countable group G necessarily has actions which admit a spectral gap, in the sense of the following general definition.

Definition 3.1. A measure-preserving ergodic action of a group G on a probability space (X, μ) has a *spectral gap* if one of the following two equivalent conditions holds:

- There are no asymptotically invariant sequences of functions as in (3.1).
- There exists an absolutely continuous probability measure β with the support of $\beta^* * \beta$ generating G such that

$$(3.3) \quad \left\| \pi_X(\beta)|_{L_0^2(X)} \right\| < 1.$$

We remark that it follows that for actions with spectral gap the estimate (3.3) holds for every absolutely continuous probability measure β on G with the support of $\beta * \beta$ generating G as a semigroup.

Kazhdan [K67] has made the truly fundamental discovery that many groups satisfy an extremely general form of the spectral gap condition. It is satisfied for all ergodic probability preserving actions of the group, and furthermore an analogous property holds in fact for all unitary representations without invariant unit vectors.

Definition 3.2. A group G has *Kazhdan property (T)* if one of the following equivalent conditions holds:

- Every strongly continuous unitary representation π of G which has no invariant unit vectors also has no sequences of asymptotically invariant unit vectors as in (3.1).
- For some absolutely continuous probability measure β on G with the support of $\beta * \beta$ generating G

$$(3.4) \quad \sup_{\pi} \|\pi(\beta)\| < 1,$$

where the supremum is taken over all unitary representations of G that have no invariant unit vectors.

The class of groups satisfying property (T) is very rich and extensive. For instance, the Lie groups $SL_n(\mathbb{R})$, $SL_n(\mathbb{C})$, and $Sp_{2n}(\mathbb{R})$ with $n \geq 3$ have this property, as do many other linear algebraic simple groups. In fact, among simple Lie groups, only the isometry groups of real or complex hyperbolic spaces do not have property (T) (the case of rank-one groups is dealt with in [K69]). Similarly, simple algebraic groups over non-Archimedean local fields all have property (T), unless their split rank is one. Furthermore, every lattice subgroup of a group with property (T) also has that property (T); for example, $SL_n(\mathbb{Z})$, $SL_n(\mathbb{Z}[i])$, and $Sp_{2n}(\mathbb{Z})$ with $n \geq 3$. We refer to the monograph [BHV] for a comprehensive discussion of property (T).

The isometry groups of real and complex hyperbolic spaces (and algebraic group of split rank-one) do in fact have probability preserving actions admitting an asymptotically invariant sequence of unit vectors, namely actions which do not have a spectral gap. Nevertheless, many of the most important and natural actions of these groups do have a spectral gap, although its size may vary from action to action, and it does not obey the remarkable uniform estimate provided by property (T). The problem of establishing uniform spectral estimates for certain families of actions (or families of unitary representations) of these groups is a problem of foundational importance, which has been studied extensively. Let us now turn to describe in more detail some of the results and the challenges of establishing spectral gap and uniform spectral gap estimates in this context. Such estimates will be crucial to our later arguments proving ergodic theorems with rates of convergence for actions of semisimple algebraic groups and to establishing some of their number-theoretic applications.

3.2. Selberg’s property and property (τ). The phenomenon where a family of actions with increasing complexity satisfies the spectral gap property uniformly, namely the estimate in (3.3) is uniform, is termed *property (τ)*.

The first instance of property (τ) was discovered by A. Selberg in [S65] for congruence subgroups of $SL_2(\mathbb{Z}) \subset SL_2(\mathbb{R})$. Let us consider a sequence of covers

consisting of finite area hyperbolic surfaces

$$\mathbb{H}^2 \rightarrow \cdots \rightarrow S_n \rightarrow \cdots \rightarrow S_2 \rightarrow S_1 = \mathbb{H}^2/\mathrm{SL}_2(\mathbb{Z}).$$

We denote by $\lambda(S_n)$ the smallest non-zero element of the spectrum of the Laplace operator $\Delta = y^2 \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right)$ on S_n . One might expect that when the covers S_n grow in a regular fashion, they *approximate* the hyperbolic space \mathbb{H}^2 , and $\lambda(S_n)$ approaches the smallest non-zero element of the spectrum of the Laplace operator on the hyperbolic plane which is $\lambda(\mathbb{H}^2) = \frac{1}{4}$. In particular, $\lambda(S_n)$ is expected to stay bounded away from zero. Here the assumption that covers grow in a regular fashion is essential, and A. Selberg constructed an example of a sequence of covers S_n such that $\lambda(S_n) \rightarrow 0$ as $n \rightarrow \infty$. On the other hand, for the surfaces $S_\ell = \mathbb{H}^2/\Gamma(\ell)$, where $\Gamma(\ell) = \{\gamma \in \mathrm{SL}_2(\mathbb{Z}) : \gamma = \mathrm{I} \pmod{\ell}\}$ denotes the congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$ of level ℓ , he showed that

$$(3.5) \quad \lambda(S_\ell) \geq \frac{3}{16}.$$

Moreover, A. Selberg conjectured that for the congruence covers S_ℓ ,

$$(3.6) \quad \lambda(S_\ell) \geq \lambda(\mathbb{H}^2) = \frac{1}{4}.$$

Although estimate (3.5) has been improved in [GJ78, LRS95, I96, KS03], the Selberg conjecture (3.6) is still open. We refer to the surveys [S95, S03, BB13] for more detailed discussions. Nowadays this conjecture is understood as a special case of the generalized Ramanujan conjectures which are most conveniently stated in representation-theoretic terms. These conjectures also include the original Ramanujan conjecture concerning the bounds on Fourier coefficients of holomorphic modular forms, which was solved by P. Deligne.

Thus, more generally, let $G \subset \mathrm{GL}_d(\mathbb{C})$ be a linear algebraic semisimple group defined over \mathbb{Q} , $\Gamma = G(\mathbb{Z})$, and let

$$\Gamma(\ell) = \{\gamma \in \Gamma : \gamma = \mathrm{I} \pmod{\ell}\}$$

denote the congruence subgroups. Let $L_0^2(G(\mathbb{R})/\Gamma(\ell))$ be the space of square-integrable functions with zero integral on the space $G(\mathbb{R})/\Gamma(\ell)$. We consider the family of unitary representation of $G(\mathbb{R})$ on $L_0^2(G(\mathbb{R})/\Gamma(\ell))$ defined by

$$(3.7) \quad \pi_\ell(g)\phi(x) = \phi(g^{-1}x), \quad \phi \in L_0^2(G(\mathbb{R})/\Gamma(\ell)).$$

The generalized Ramanujan conjecture seeks to describe the decomposition of the representations π_ℓ (and more general automorphic representations) into irreducible constituents. We refer to [S05, BB13] for a comprehensive survey of this subject, but note that what is of crucial importance for our purposes is the uniform spectral gap property generalizing (3.5). In representation-theoretic language, this property can be reformulated in terms of isolation of the trivial representation of $G(\mathbb{R})$ from the other irreducible representations appearing in π_ℓ for all $\ell \geq 1$, as follows.

In general, if Π is a family of unitary representations, we say that Π has *property* (τ) if (3.4) holds with $\pi \in \Pi$, and naturally for number-theoretic applications the family Π is taken as the family of automorphic representations (3.7). The problem of establishing a uniform spectral gap for the automorphic family has a long history and is closely connected with work on Langlands functoriality conjectures. In particular, for other forms of the group SL_2 , property (τ) follows from Gelbart and Jacquet's correspondence [GJ78]. An important result due to M. Burger and

P. Sarnak [BS91] shows that it is sufficient to verify property (τ) for the automorphic family associated with certain proper algebraic subgroups of the ambient algebraic group. Combined with previous work, this result established property (τ) for the automorphic family for all semisimple \mathbb{Q} -simple simply connected algebraic groups except the unitary groups. Finally, the case of unitary groups was settled by L. Clozel in [C03].

In the case of semisimple algebraic groups, it is convenient to measure the size of the spectral gap in terms of integrability exponents that were introduced by Cowling [C79] and Howe and Moore [HM79], as follows. Let π be a unitary representation of a locally compact group G on a Hilbert space \mathcal{H} . For $v, w \in \mathcal{H}$, we denote by $c_{v,w}(g) = \langle \pi(g)v, w \rangle$ the corresponding matrix coefficient, which is a continuous bounded function on G . We define the *integrability exponent* of the representation π by

$$(3.8) \quad q(\pi) = \inf \{q > 0 : c_{v,w} \in L^q(G) \text{ for all } v, w \text{ in a dense subset of } \mathcal{H}\}.$$

It was shown in [BW80, C79] that when G is a connected simple Lie group with property (T) ,

$$(3.9) \quad \sup_{\pi} q(\pi) < \infty,$$

where the supremum is taken over all unitary representations of G that have no non-zero invariant vectors. The value of the supremum has been estimated or computed in [HT, L95, LZ96, O98, O02, N03]. The estimate (3.9) can be considered as a quantitative version of property (T) . We also note the important fact that when the group does not satisfy property (T) , still the automorphic family satisfies property (τ) , and the results on property (τ) for automorphic representations cited above provide explicit estimates specifically for $\sup_{\ell \geq 1} q(\pi_{\ell})$.

4. SPECTRAL GAPS AND ERGODIC THEOREMS

4.1. Quantitative mean and pointwise theorems for G -actions. We now turn to discuss the role of the spectral estimates described in the previous section in the proofs of ergodic theorems. Let (X, μ) be a (standard) probability space with a measure-preserving action of a locally compact second countable G equipped with left-invariant Haar measure. For a family of measurable sets B_t of G with finite positive measure, we denote by β_t the uniform probability measures supported on B_t . We aim to prove ergodic theorems for the corresponding averaging operators $\pi_X(\beta_t)$ defined by

$$(4.1) \quad \pi_X(\beta_t)f(x) = \frac{1}{|B_t|} \int_{B_t} f(g^{-1}x) dg, \quad f \in L^p(X).$$

In the rest of this section, we discuss quantitative ergodic theorems for connected simple Lie groups G (for instance, $G = \mathrm{SL}_n(\mathbb{R})$). In order to prove the quantitative mean ergodic theorem for the averages $\pi_X(\beta_t)$, we need to estimate the norms of the operators $\pi_X(\beta_t)|_{L^2_0(X)}$, and this is achieved using a method of *spectral transfer*, originating in [N98]. The method consists of two separate steps. First, it is possible to reduce the norm estimate of $\pi(\beta_t)$ in a general representation π to a norm estimate for the convolution operator $\lambda_G(\beta_t)$ for the regular representation λ_G of G acting on $L^2(G)$. This serves as an analogue, in this context, of the transference principle for amenable groups. Second, it is then necessary to estimate the norm of the convolution operators on G .

Put in quantitative form, the principle of spectral transfer is based on the following two fundamental facts:

- If the action of G on (X, μ) has spectral gap, then the representation $\pi_X|_{L_0^2(X)}$ has finite integrability exponent, and it follows that a sufficiently high tensor power representation $(\pi_X|_{L_0^2(X)})^{\otimes n}$ embeds as a subrepresentation of a multiple of the regular representation λ_G of G . It then follows from Jensen's inequality that for even n

$$\left\| \pi_X(\beta_t)|_{L_0^2(X)} \right\| \leq \|\lambda_G(\beta_t)\|^{1/n}.$$

- The convolution operator $\lambda_G(\beta_t)$ in the regular representation obey a remarkable convolution inequality, known as *the Kunze–Stein phenomenon*: for $\beta_t \in L^r(G)$ with $1 \leq r < 2$ and $f \in L^2(G)$,

$$\|\beta_t * f\|_2 \leq c_r \|\beta_t\|_r \|f\|_2$$

with uniform $c_r > 0$. This convolution inequality for $\mathrm{SL}_2(\mathbb{R})$ is due to Kunze and Stein [KS60] and in general to Cowling [C78]. The Kunze–Stein phenomenon therefore implies that

$$\|\lambda_G(\beta_t)\| \leq c_r \left\| \frac{\chi_{B_t}}{|B_t|} \right\|_r = c_r |B_t|^{-1+1/r}.$$

This leads to an estimate for $\left\| \pi_X(\beta_t)|_{L_0^2(X)} \right\|$, and using an interpolation argument, one extends this estimate for the action of G on $L^p(X)$, $1 < p < \infty$. Hence, we deduce the following quantitative mean ergodic theorem, stated in [N98].

Theorem 4.1. *Given a measure-preserving action with a spectral gap of a connected simple Lie group G on a (standard) probability space (X, μ) , for $1 < p < \infty$ there exist $c_p, \theta_p > 0$ such that for the averaging operators $\pi_X(\beta_t)$ as in (4.1),*

$$\left\| \pi_X(\beta_t)f - \int_X f d\mu \right\|_p \leq c_p |B_t|^{-\theta_p} \|f\|_p, \quad f \in L^p(X).$$

Moreover, the parameters c_p, θ_p depend only on G, p and the integrability exponent of the representation $\pi_X|_{L_0^2(X)}$.

It is a remarkable phenomenon that the rate of convergence stated in Theorem 4.1 depends only on the measure of the sets B_t . Thus, L^2 -convergence with a rate in a general ergodic action requires no geometric, regularity, or stability assumptions of the family of sets in question, so that the quantitative mean ergodic theorem is very robust.

The pointwise ergodic theorem, however, is considerably more delicate, and for its validity some stability and regularity assumptions on the family B_t are indispensable. We turn now to discuss these issues in greater detail.

Definition 4.2. An increasing family of bounded measurable subsets B_t , $t > 0$, of G is called *admissible* if there exists $c > 0$ such that for all $t \geq t_0$ and $\varepsilon \in (0, \varepsilon_0)$,

$$O_\varepsilon B_t O_\varepsilon \subset B_{t+c\varepsilon},$$

where O_ε denotes the ball of radius ε centered at the identity with respect to a right-invariant metric on G , and in addition

$$|B_{t+\varepsilon}| \leq (1 + c\varepsilon)|B_t|.$$

For instance, if a simple group G is embedded as a closed subgroup of $\mathrm{GL}_n(\mathbb{R})$ and $\|\cdot\|$ is a norm on $M_n(\mathbb{R})$, one can show that the family

$$B_t = \{g \in G : \log \|g\| < t\}$$

is admissible (see [EMS96, Appendix], [GN10, Ch. 7]).

Under this mild regularity property, we also establish a quantitative pointwise ergodic theorem [MNS00, GN10]:

Theorem 4.3. *Consider a measure-preserving action with a spectral gap of a connected simple Lie group G on a (standard) probability space, and let $\pi_X(\beta_t)$ be the averaging operators with respect to an admissible family of sets. Then there exist $c'_p, \theta'_p > 0$ such that for every $f \in L^p(X)$, $p > 1$, and μ -almost every $x \in X$,*

$$\left| \pi_X(\beta_t)f(x) - \int_X f d\mu \right| \leq C_p(f, x) |B_t|^{-\theta'_p},$$

where $\|C_p(f, \cdot)\|_p \leq c'_p \|f\|_p$.

As noted in the introduction, in the absence of asymptotic invariance arguments, the proof of Theorem 4.3 is based primarily on spectral methods, with one key ingredient being a quantitative strong maximal inequality established using analytic interpolation.

We remark that it is also possible to establish a pointwise ergodic theorem for general actions of simple Lie groups which do not satisfy the spectral gap assumption. In this case however there can be no estimate on the rate of convergence. Here one key ingredient is the use of derivative estimates of matrix coefficients and square function arguments to establish the necessary maximal inequality.

We will not elaborate further on the proofs and refer instead to [GN10] for a detailed account of ergodic theorems for semisimple algebraic groups.

4.2. On spectral methods in ergodic theory. Before proceeding with the description of the applications of the mean and pointwise quantitative ergodic theorems stated above, let us make several very brief comments on some of the uses that spectral methods have found in ergodic theory.

1. *Euclidean groups: Singular averages.* The methods of classical harmonic analysis on \mathbb{R}^d , including singular integrals, square function estimates, restriction theorems, fractional integrals, and analytic interpolation, have been utilized by several authors to prove groundbreaking pointwise ergodic theorems. Pioneering contributions here were Stein's ergodic theorem for the powers of a self-adjoint operator [St61] and the celebrated spherical maximal theorem in \mathbb{R}^d , proved by Stein and Wainger for $d \geq 3$ [SW78] and by Bourgain for $d = 2$ [Bo83]. The spherical maximal inequalities were used to prove a pointwise ergodic theorem for sphere averages, with the case $d \geq 3$ established by Jones [Jo93], and the case $d = 2$ by Lacey [La95]. Ergodic theorems for sphere averages on the Heisenberg group were developed in [NT97].

For actions of the integers, the methods of discrete harmonic analysis and especially estimates of exponential sums were developed to prove pointwise ergodic theorem for sparse sets of integers, such as squares and primes, in Bourgain's pioneering work on the subject [Bo89]. For actions of the integer lattices \mathbb{Z}^d , ergodic theorems were developed for discrete spheres and other averages by Magyar [Mag02]. We remark that all the ergodic theorems just mentioned depend on the

transference principle for the groups involved; namely, they reduce the proof of the requisite maximal inequalities to the case of the action by convolution.

2. *Free groups: Radial averages.* The problem of establishing ergodic theorems for general non-amenable groups was raised already half a century ago by Arnol'd and Krylov [AK63]. An important case that has figured prominently in the theory ever since is that of the free (non-Abelian) group. Arnol'd and Krylov proved an equidistribution theorem for radial averages on dense free subgroups of isometries of the unit sphere \mathbb{S}^2 via a spectral argument similar to Weyl's equidistribution theorem on the circle. Guivarc'h has established a mean ergodic theorem for radial averages on the free group, using von Neumann's original approach via the spectral theorem [G69]. The pointwise ergodic theorem for sphere averages in general actions of the free group was proved in [N94] for L^2 -functions and was extended to function in L^p , $p > 1$, in [NS94]. The distinctly non-amenable phenomenon of an ergodic theorem with a quantitative estimate on the rate of convergence was realized by the celebrated Lubotzky–Phillips–Sarnak construction [LPS86, LPS87] of a dense free group of isometries of \mathbb{S}^2 which has an optimal spectral gap.

3. *Semisimple groups: Ball and sphere averages.* An important and influential general spectral result on ergodic actions of semisimple Lie group has been the Howe–Moore mixing theorem [HM79], a consequence of the decay of matrix coefficients in unitary representations of these groups. Tempelman (see [T92]) has used the Howe–Moore theorem to prove mean ergodic theorems for averages on a semisimple Lie group.

Some of the methods of classical harmonic analysis alluded to above were combined with unitary representation theory of semisimple groups and used extensively in the development of pointwise ergodic theorems for radial averages on semisimple Lie groups and some of their lattices. For sphere and ball averages on simple groups of real rank-one, pointwise ergodic theorems were developed in [N94b, N97, NS97], and for complex groups in [CN01]. Pointwise ergodic theorems with a rate of convergence to the ergodic mean were developed for radial averages on general semisimple Lie groups in [MNS00] and for more general averages in [N98].

5. COUNTING LATTICE POINTS

Let G be a locally compact second countable group, and let Γ be a discrete subgroup with finite covolume, namely a lattice subgroup. We are interested in determining the asymptotic behavior of the number of elements of Γ contained in a family of increasing compact domains B_t in G . Problems of this type arise naturally in arithmetic geometry when $X \subset \mathbb{C}^n$ is an affine algebraic variety and one is interested in estimating the cardinality of the set of integral points $x \in X(\mathbb{Z})$ with bounded norm. The techniques discussed in this section can be applied in particular when $G \subset \mathrm{GL}_n(\mathbb{C})$ is an algebraic group defined over \mathbb{Q} , and $\Gamma = G(\mathbb{Z})$ is its arithmetic subgroup, and more generally to symmetric varieties of G .

The most classical instance of the lattice counting problem is to estimate the number of integral vectors contained in dilations $tB = B_t$ of a compact domain $B \subset \mathbb{R}^d$. In this case, it is easy to see that

$$|\mathbb{Z}^d \cap B_t| = |B_t| + O(|\partial B_t|),$$

and typically $\frac{|\partial B_t|}{|B_t|} \rightarrow 0$ as the volume $|B_t| \rightarrow \infty$, which leads to an asymptotic formula. The analogous problem for more general groups presents significant new challenges. In particular, for groups with exponential volume growth, $|B_t|$ is comparable to $|\partial B_t|$, so that the error term indicated above is comparable to the main term. Furthermore, when the group has exponential growth, the fundamental domain for the lattice Γ in G might be unbounded. Thus, in order to establish an asymptotic formula for $|\Gamma \cap B|$, one must use more sophisticated analytic techniques. A number of methods have been developed to address this problem, which include, in particular:

- direct spectral expansion of the automorphic kernel [D42, H56, S56, P76, G83, LP82, MW92, DRS93, BMW99, BGM11];
- mixing arguments based on decay of matrix coefficients [M04, B82, EM93, Ma07, BO12, O10];
- symbolic coding and transfer operator techniques when the lattice is a hyperbolic group [L89, P95];
- equidistribution using the theory of unipotent flows [EMS96].

We refer to [B02] for an extensive survey.

Our approach to the lattice point counting problem in the domains B_t is different than those just mentioned, and it is based on applying the ergodic theorem to the averaging operators defined by B_t in the action of G on the homogeneous space $X = G/\Gamma$ equipped with the invariant probability measure μ . Thus given a family of domains B_t in G , we consider

$$\pi_{G/\Gamma}(\beta_t)f(x) = \frac{1}{|B_t|} \int_{B_t} f(g^{-1}x) dg, \quad f \in L^2(G/\Gamma),$$

and we will demonstrate that a (quantitative) mean ergodic theorem for these operators implies a (quantitative) solution of the lattice point counting problem for the family of sets B_t , provided that it satisfies the regularity property of Definition 4.2.

Theorem 5.1. *Let G be a locally compact second countable group, let Γ be a discrete lattice subgroup, and assume that the family of sets B_t is admissible.*

- (i) *Assume the averages $\pi_{G/\Gamma}(\beta_t)$ satisfy the mean ergodic theorem*

$$\left\| \pi_{G/\Gamma}(\beta_t)f - \int_{G/\Gamma} f d\mu \right\|_2 \rightarrow 0, \quad f \in L^2(X).$$

Then

$$|\Gamma \cap B_t| \sim \frac{|B_t|}{|G/\Gamma|} \quad \text{as } t \rightarrow \infty.$$

- (ii) *Assume the averages $\pi_X(\beta_t)$ satisfy the quantitative mean ergodic theorem*

$$(5.1) \quad \left\| \pi_X(\beta_t)f - \int_X f d\mu \right\|_2 \leq C_2 |B_t|^{-\theta} \|f\|_2, \quad f \in L^2(X, \mu),$$

with $C_2, \theta > 0$. Then for all sufficiently large t ,

$$|\Gamma \cap B_t| = \frac{|B_t|}{|G/\Gamma|} + O\left(|B_t|^{1-\frac{\theta}{\dim(G)+1}}\right),$$

where the implied constants depends only on the parameters in (5.1) and Definition 4.2 and on the covolume of Γ .

Note that the formulation of Theorem 5.1 is very general, and by $\dim(G)$ in part (ii) we mean the upper local dimension, namely the infimum over $d > 0$ such that $|O_\varepsilon| \geq C(d)\varepsilon^d$ for all sufficiently small $\varepsilon > 0$, where O_ε denotes the ε -neighborhood of identity appearing in Definition 4.2. For connected real Lie groups this coincides of course with the usual notion of dimension.

Let us outline our proof of Theorem 5.1, which proceeds in two steps. The first step is to observe that the quantity $|\Gamma \cap B_t|$ can be approximated by suitable averages on space $X = G/\Gamma$. Recall that O_ε is the ball of radius ε centered at the identity with respect to a right-invariant metric on G , which clearly satisfies $O_\varepsilon^{-1} = O_\varepsilon$. We normalize the invariant measure on G so that Γ has covolume one. Let χ_ε be the characteristic function of O_ε normalized to have integral one and $f_\varepsilon(g\Gamma) = \sum_{\gamma \in \Gamma} \chi_\varepsilon(g\gamma)$. Then $\int_X f_\varepsilon d\mu = 1$. We claim that for $h \in O_\varepsilon$,

$$(5.2) \quad \int_{B_{t-c\varepsilon}} f_\varepsilon(g^{-1}h\Gamma) dg \leq |\Gamma \cap B_t| \leq \int_{B_{t+c\varepsilon}} f_\varepsilon(g^{-1}h\Gamma) dg.$$

Indeed, observe first that

$$\int_{B_t} f_\varepsilon(g^{-1}h\Gamma) dg = \sum_{\gamma \in \Gamma} \int_{B_t} \chi_\varepsilon(g^{-1}h\gamma) dg = \sum_{\gamma \in \Gamma} \frac{|B_t \cap h\gamma O_\varepsilon|}{|O_\varepsilon|}.$$

Now if $\gamma \in B_{t-c\varepsilon}$, then, by the admissibility property,

$$h\gamma O_\varepsilon \subset O_\varepsilon B_{t-c\varepsilon} O_\varepsilon \subset B_t$$

so that

$$|\Gamma \cap B_{t-c\varepsilon}| = \sum_{\gamma \in \Gamma \cap B_{t-c\varepsilon}} \frac{|B_t \cap h\gamma O_\varepsilon|}{|O_\varepsilon|} \leq \sum_{\gamma \in \Gamma} \frac{|B_t \cap h\gamma O_\varepsilon|}{|O_\varepsilon|}.$$

On the other hand, if $h\gamma O_\varepsilon \cap B_t \neq \emptyset$, then, by the admissibility property,

$$\gamma \in h^{-1}B_t O_\varepsilon \subset O_\varepsilon B_t O_\varepsilon \subset B_{t+c\varepsilon}.$$

Hence,

$$\sum_{\gamma \in \Gamma} \frac{|B_t \cap h\gamma O_\varepsilon|}{|O_\varepsilon|} = \sum_{\gamma \in \Gamma \cap B_{t+c\varepsilon}} \frac{|B_t \cap h\gamma O_\varepsilon|}{|O_\varepsilon|} \leq |\Gamma \cap B_{t+c\varepsilon}|,$$

and this proves (5.2).

The second step is to observe that from the mean ergodic theorem it follows that for every $\delta > 0$ and every $\varepsilon > 0$,

$$\mu(\{h\Gamma \in G/\Gamma : |\pi_{G/\Gamma}(\beta_t) f_\varepsilon(h\Gamma) - 1| > \delta\}) \rightarrow 0$$

as $t \rightarrow \infty$. In particular, it follows that given $\delta, \varepsilon > 0$, for all sufficiently large t , there exists $h_t \in O_\varepsilon$ such that

$$(5.3) \quad |\pi_{G/\Gamma}(\beta_t) f_\varepsilon(h_t\Gamma) - 1| \leq \delta,$$

or equivalently,

$$(1 - \delta) |B_t| \leq \int_{B_t} f_\varepsilon(g^{-1}h_t\Gamma) dg \leq (1 + \delta) |B_t|.$$

Combining this estimate with (5.2), we deduce that

$$(5.4) \quad |\Gamma \cap B_t| \leq (1 + \delta) |B_{t+c\varepsilon}| \leq (1 + \delta)(1 + c^2\varepsilon) |B_t|$$

and similarly,

$$(5.5) \quad |\Gamma \cap B_t| \geq (1 - \delta)(1 - c^2\varepsilon) |B_t|$$

for sufficiently large t . Since these estimates hold for arbitrary $\delta, \varepsilon > 0$, this proves part (i) of Theorem 5.1.

The proof of the quantitative estimate in part (ii) of Theorem 5.1 follows similar ideas and is equally straightforward. We assume that ε is sufficiently small, so that the neighborhoods $O_\varepsilon\gamma$, $\gamma \in \Gamma$, are disjoint. Then since $|O_\varepsilon| \geq C(n)\varepsilon^n$ when $n > \dim(G)$, we deduce that

$$\|\chi_\varepsilon\|_2^2 = |O_\varepsilon|^{-1} \leq \varepsilon^{-n}.$$

The validity of (5.1) now implies that for every $\delta > 0$, $\varepsilon > 0$, and $t > 0$,

$$\mu(\{h\Gamma \in G/\Gamma : |\pi_X(\beta_t)f_\varepsilon(h\Gamma) - 1| > \delta\}) \leq C_2^2 \delta^{-2} C(n)^{-1} \varepsilon^{-n} |B_t|^{-2\theta}.$$

Taking $\delta = 2C_2C(n)^{-1}\varepsilon^{-n}|B_t|^{-\theta}$, we observe that the above measure is bounded by $\frac{1}{4}C(n)\varepsilon^n \leq \frac{1}{4}|O_\varepsilon|$ and so is strictly less than $\mu(O_\varepsilon\Gamma) = |O_\varepsilon|$. Thus we can find $h_t \in O_\varepsilon$ satisfying (5.3) and deduce estimates (5.4)–(5.5) above. Finally, to optimize the estimate, we take $\varepsilon = |B_t|^{-\theta/(n+1)}$ with sufficiently large t , and this completes the proof of Theorem 5.1(ii).

Let us note that Theorem 5.1 can be generalized significantly in two important respects (for full details we refer to [GN12a]):

- (1) The quantitative lattice point count extends to families of sets B_t much more general than admissible ones. It suffices that the families satisfy the following weaker regularity property. There exist $c, a > 0$ such that for all $t \geq t_0$ and $\varepsilon \in (0, \varepsilon_0)$,

$$\left| \bigcup_{u,v \in O_\varepsilon} uB_tv \right| \leq (1 + c\varepsilon^a) \left| \bigcap_{u,v \in O_\varepsilon} uB_tv \right|,$$

which we call *Hölder well-roundedness*. This extends the solution of the lattice point counting problem, for example, to families of sector averages, families associated naturally with symmetric varieties, and many others.

- (2) The quantitative solution to the lattice point counting problem is uniform as the lattice varies in the space of lattice subgroups, provided only that the following holds. First, the averaging operators must satisfy a common norm decay estimate in the spaces $L^2(G/\Gamma)$ as Γ varies, and second, G/Γ must contain an injective copy of a fixed neighborhood of the identity in G as Γ varies. In particular, these uniformity properties hold when Γ varies over any set of finite index subgroups of a fixed lattice which satisfies property (τ) . Note that in the ergodic-theoretic approach we have taken, the fact that the uniformity properties imply a uniform solution to the lattice point counting problem is completely obvious from the proof, which treats all probability measure-preserving transitive actions of G on an equal footing.

In particular, combining the arguments of this section with Theorem 4.1 and property (τ) for congruence subgroups, we can deduce:

Corollary 5.2. *Let $G \subset \mathrm{GL}_n(\mathbb{C})$ be a simply connected simple algebraic group defined over \mathbb{Q} , and let $\Gamma(\ell) = \{\gamma \in G(\mathbb{Z}) : \gamma = \mathrm{id} \pmod{\ell}\}$ be the family of congruence subgroups. Then for any Hölder well-rounded family of sets B_t in $G(\mathbb{R})$, there exists $\delta > 0$ such that for sufficiently large t ,*

$$|x\Gamma(\ell)y \cap B_t| = \frac{|B_t|}{|G(\mathbb{R})/\Gamma(\ell)|} + O(|B_t|^{1-\delta}),$$

where the implied constant is independent of ℓ , and it is bounded uniformly as x and y vary in sets which are compact modulo Γ .

As we will see presently, this estimate plays an indispensable role in the sifting, lifting, and restricting problems on symmetric varieties to which we now turn.

6. ALMOST PRIME POINTS ON HOMOGENEOUS ALGEBRAIC VARIETIES

In the present section we will describe a solution to the problem of sifting the integral points on certain homogeneous algebraic varieties to obtain almost prime points, following [NS10, GN12c]. As will be explained in due course, the uniformity in the lattice point counting problem established in the previous section will play a crucial role. Furthermore, the quality of the error estimate obtained there directly determines the quality of the estimate of the saturation number. But let us begin with some remarks on the origins of this problem.

6.1. Sifting along finite-index subgroups. The prime number theorem describes the asymptotic distribution of the set $\mathbb{P} = \{2, 3, 5, 7, 11, \dots\}$ of prime numbers

$$(6.1) \quad |\{p \in \mathbb{P} : p \leq t\}| \sim \frac{t}{\log t} \quad \text{as } t \rightarrow \infty.$$

More generally, let us consider an integral irreducible polynomial f . We assume that the leading coefficient of f is positive and $\gcd(f(\mathbb{Z})) = 1$. Is it true that $f(x)$ is prime for infinitely many integers x ? The question is answered for linear polynomials by the Dirichlet theorem about primes in arithmetic progressions, but it turns out to be extremely difficult for polynomials of higher degrees. A natural approximation to this problem is the question whether for a fixed $r \geq 1$, the values $f(x)$ belongs to the set \mathbb{P}_r of r -almost primes infinitely often, where \mathbb{P}_r denotes the set of integers having at most r prime factors. This problem has been extensively studied using sieve methods. For example, it has been shown (see [HR, Sec. 9.5]) that when $r = \deg(f) + 1$, there is a lower estimate of correct order

$$(6.2) \quad |\{x \in \mathbb{Z} : f(x) \in \mathbb{P}_r, |x| \leq t\}| \geq \text{const} \cdot \frac{t}{\log t}.$$

Although it is widely believed that under the above necessary conditions on f , the same estimate should hold for the set of primes as well, this is currently out of reach. In particular, it is not known whether there are infinitely many primes of the form $n^2 + 1$ (see [I78] for the best result in this direction).

A far-reaching conjectural generalization of this problem was proposed by Bourgain, Gamburd, and Sarnak (see [S08] and [BGS10]) and was termed the saturation problem, as follows. Given

- a subgroup Γ of $\text{GL}_d(\mathbb{Z})$ whose Zariski closure G is a simply connected semisimple \mathbb{Q} -group,
- a linear representation $\sigma : \text{GL}_d(\mathbb{C}) \rightarrow \text{GL}_n(\mathbb{C})$ defined over \mathbb{Q} such that $\sigma(\Gamma) \subset \text{GL}_n(\mathbb{Z})$,
- a vector $v \in \mathbb{Z}^n$,
- a polynomial $f \in \mathbb{Z}[x_1, \dots, x_n]$,

consider the orbit $\sigma(\Gamma)v$ of an integral vector v . Is it true that $f(x)$, with x in the orbit $\sigma(\Gamma)v$, is r -prime infinitely often, and more generally, is the set of such points $x \in \rho(\Gamma)v$ Zariski dense in the Zariski closure of $\sigma(\Gamma)v$?

One can also ask for a quantitative version of this question. Let us omit the explicit mention of the representation σ in the notation and set

$$\mathcal{O}(t) = \{x \in \Gamma v : \|x\| \leq t\}.$$

Is it true that, for a suitable m ,

$$(6.3) \quad |\{x \in \mathcal{O}(t) : f(x) \in \mathbb{P}_r\}| \geq \text{const} \cdot \frac{|\mathcal{O}(t)|}{(\log t)^m} \quad \text{as } t \rightarrow \infty,$$

in analogy with (6.2)?

It was a profound idea of P. Sarnak that classical sieve methods can be applied in this non-commutative setting provided that one is able to establish an asymptotic estimate with an error bound for the cardinalities of the sets

$$(6.4) \quad \mathcal{O}(t, \ell) = \{x \in \Gamma v : \|x\| \leq t, f(x) = 0 \pmod{\ell}\}$$

as $t \rightarrow \infty$. It is crucial for the sieve method to succeed that the error bound is sufficiently uniform over ℓ . Letting

$$(6.5) \quad \Gamma(\ell) = \{\gamma \in \Gamma : \gamma = \mathbf{I} \pmod{\ell}\}$$

denote the congruence subgroup of level ℓ , it is clear that $f(\gamma\Gamma(\ell)v) = f(\gamma v) \pmod{\ell}$ for $\gamma \in \Gamma/\Gamma(\ell)$. Thus in order to estimate $|\mathcal{O}(t, \ell)|$, it is sufficient to give an asymptotics for the cardinalities of the sets

$$(6.6) \quad \{x \in \gamma\Gamma(\ell)v : \|x\| \leq t\} \quad \text{for } \gamma \in \Gamma/\Gamma(\ell).$$

Let us now consider the important special case of lattice actions on principal homogeneous spaces, where $\mathcal{O} = \Gamma v \subset Gv$, Γ is a lattice in G , and the stability group of v in G is finite. Then estimating (6.6) amounts to nothing other than the lattice point counting problem in the sets given by the intersection of norm balls with cosets of congruence subgroups of Γ . As already noted, it is crucial for the sieve method to succeed that the solution to the lattice point counting problem in these sets be uniform over the congruence subgroups and their cosets. The uniform counting property over cosets of finite-index subgroups does indeed hold provided that the finite-index subgroups in question satisfy property (τ) ; see [NS10]. Thus when Γ is an arithmetic subgroup of a semisimple algebraic group G (for instance, when $\Gamma = \text{SL}_d(\mathbb{Z})$), the required estimates for the sieve method ultimately follow from the uniform spectral gap property, namely property (τ) for congruence subgroups discussed in Section 3.2. Indeed, property (τ) gives a uniform spectral estimate for the ball averages acting on the spaces $L^2(G(\mathbb{R})/\Gamma(\ell))$, $\ell \geq 1$, which ultimately implies a uniform estimate in the lattice point counting problem, as we saw in Section 5.

Establishing a Zariski dense subset of almost prime points on more general homogeneous varieties requires the more general uniform lattice point counting results of Corollary 5.2. Before stating an estimate for the saturation number on symmetric varieties, let us illustrate how the estimates on $|\mathcal{O}(t, \ell)|$ comes into play on a simple sieve based on the Möbius function. Recall that the Möbius function $\mu(\ell)$ is equal to $(-1)^r$, if ℓ is a product of r distinct prime factors, and is 0 otherwise, and the Möbius inversion formula is given by

$$\sum_{\ell: \ell|n} \mu(\ell) = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{if } n > 1. \end{cases}$$

Let P_z be the product of all prime numbers up to z , and let $S(t, z)$ denote the cardinality of the set of $x \in \mathcal{O}(t)$ such that $f(x)$ is coprime to P_z . Then

$$\begin{aligned}
 (6.7) \quad S(t, z) &= \sum_{k: \gcd(k, P_z)=1} |\mathcal{O}(t) \cap \{f = k\}| \\
 &= \sum_k \left(\sum_{\ell: \ell | \gcd(k, P_z)} \mu(\ell) \right) |\mathcal{O}(t) \cap \{f = k\}| \\
 &= \sum_{\ell: \ell | P_z} \mu(\ell) \sum_{k: \ell | k} |\mathcal{O}(t) \cap \{f = k\}| \\
 &= \sum_{\ell: \ell | P_z} \mu(\ell) |\mathcal{O}(t, \ell)|.
 \end{aligned}$$

Now suppose that for some $\kappa, \delta > 0$, we have an estimate

$$(6.8) \quad |\mathcal{O}(t, \ell)| = \rho(\ell) |\mathcal{O}(t)| + O(\ell^\kappa |\mathcal{O}(t)|^{1-\delta})$$

with $\kappa, \delta > 0$ being uniform in ℓ , and where

$$\rho(\ell) = \frac{|\{\Gamma v \bmod \ell\} \cap \{f = 0 \bmod \ell\}|}{|\{\Gamma v \bmod \ell\}|}.$$

When z is sufficiently small compared with t , the above computation leads to favorable estimate for $S(t, z)$. However, to derive (6.3), one needs to estimate $S(t, z)$ when $z = t^\alpha$ with some $\alpha > 0$, and the above sieve with the Möbius weights is not sufficiently efficient for our purpose. Elaborate developments, starting with the work of V. Brunn in the 1920s, has produced more efficient sifting weights, and we refer to the recent surveys [F06, Mo08, Gr10, Mol10] and monographs [DH, FI, G, H, HR, IK] for a comprehensive discussion of sieve methods. These techniques allow us to deduce a favorable estimate on $S(t, z)$ for $z = t^\alpha$ with $\alpha > 0$ provided that one knows (6.8) and asymptotic estimates of the coefficients $\rho(\ell)$. The latter estimates can be deduced from the strong approximation property for algebraic groups and homogeneous varieties, and results on the number of points on varieties over finite fields. Ultimately the main difficulty lies in establishing (6.8), and the crucial initial input is the uniformity in the lattice point counting which follows from property (τ) .

As already mentioned, it is possible to apply this method to the case of orbits on general symmetric homogeneous spaces of G , not only principal ones. In particular, let us mention the following result [GN12c, Example 1.11], which shows that quadratic surfaces contain many integral points with almost prime coordinates, and it implies that the set of almost prime points is not contained in any proper algebraic subset. This example is a partial case of Theorem 6.1 below. Let $Q(x) = \sum_{i,j=1}^d a_{ij} x_i x_j$ be a non-degenerate indefinite integral quadratic form in d variables. We assume that $d \geq 4$ (see [LS10] for $d = 3$ case) and denote by Spin_Q the spinor group of Q . We fix $v \in \mathbb{Z}^d$ such that $n = Q(v) \neq 0$ and consider the orbit $\mathcal{O} = \text{Spin}_Q(\mathbb{Z})v$ which lies in the quadratic surface $X = \{Q(x) = n\}$. Assuming that the polynomial function f is absolutely irreducible on X and that $\gcd(f(x) : x \in \mathcal{O}) = 1$, we can establish that $r = 54d^4 \deg(f)$ is such that (6.3) holds, regardless of the signature of Q .

Turning to sifting on general symmetric varieties, recall that Corollary 5.2 establishes uniform asymptotics for the number of elements of $x\Gamma(\ell)y$ contained in

a family of Hölder well-rounded subsets B_t of $G(\mathbb{R})$. Thus the following estimate holds, for every $\gamma \in \Gamma/\Gamma(\ell)$,

$$(6.9) \quad |\gamma\Gamma(\ell) \cap B_t| = \frac{|B_t|}{|G(\mathbb{R})/\Gamma(\ell)|} + O(|B_t|^{1-\delta}),$$

where $\delta > 0$ depends only on the supremum of the integrability exponents in the representations $L_0^2(G/\Gamma(\ell))$, the dimension of G , the Hölder exponent of B_t 's, and where the implicit constant is independent of ℓ and $\gamma \in \Gamma/\Gamma(\ell)$. Estimate (6.9) can be used directly for sifting for almost primes on the arithmetic groups Γ itself, or more generally in its orbits in principal homogeneous spaces, and this was carried out in [NS10]. To sift for almost primes in more general orbits $\mathcal{O} = \Gamma v$, one may attempt to construct a Hölder well-rounded family $\{B_t\}$ of sets in $G(\mathbb{R})$ such that the map

$$\Gamma \cap B_t \rightarrow \mathcal{O}(t) : \gamma \mapsto \gamma v$$

is onto and has fibers with uniformly bounded cardinalities. Then using (6.9), one would show that there exists $r \geq 1$ and a suitable m such that

$$|\{\gamma \in \Gamma \cap B_t : f(\gamma v) \in \mathbb{P}_r\}| \geq \text{const} \cdot \frac{|\Gamma \cap B_t|}{(\log t)^m} \quad \text{as } t \rightarrow \infty,$$

which implies (6.3). This strategy has been realized in [GN12c] when the stabilizer of v in G is a symmetric subgroup, namely the set of fixed points of a non-trivial involution on G . It leads to the following effective result (see [GN12c, Theorem 1.10] for a more general formulation) on the saturation problem for symmetric varieties.

Theorem 6.1. *Let $G \subset \text{GL}_n(\mathbb{C})$ be a connected \mathbb{Q} -simple simply connected algebraic group defined over \mathbb{Q} , let $\Gamma = G(\mathbb{Z})$, and let $v \in \mathbb{Z}^n$ be such that $\text{Stab}_G(v)$ is connected and symmetric. Then given any polynomial $f : \mathbb{Z}^n \rightarrow \mathbb{Z}$ such that the function $g \mapsto f(gv)$ is absolutely irreducible on G and $\gcd(f(x) : x \in \mathcal{O}) = 1$, there exists an explicit $r = r(G, v, \deg(f)) \geq 1$ for which the estimate (6.3) holds.*

To conclude our discussion of sifting, let us note that major progress has been obtained very recently on the saturation problem for almost prime elements in orbits of *thin* linear groups, which are Zariski dense subgroups of infinite covolume. First, property (τ) for finite-index congruence subgroups of Zariski dense discrete linear groups has been established in great generality, and we refer to [SV] and the references therein for the definitive result. Second, using property (τ) for these Zariski dense subgroups, the construction of Zariski dense sets of r -almost prime points in their orbits (for some finite, but unspecified r) has been put in a definitive form in [SS], solving the saturation problem raised in [BGS06], [BGS10]. For accessible accounts of these remarkable developments, we refer to the two informative surveys [K2] and [K3], and for a detailed account of sieve theory in this context to [K1].

Other results in this rapidly developing area include the extensive development regarding Apollonian packings [S07, S09, KO11, BF11] and orbits of thin subgroups more generally [K09, BK10, KO12]. We refer to [K, O11, O14] for recent comprehensive surveys on this subject.

6.2. The lifting and non-concentration problems for lattice points. The uniformity properties of the solution to the lattice point counting problem, as stated in Corollary 5.2 have further useful consequences, and we proceed to describe results

on the *lifting problem* for Diophantine equations and the *non-concentration problem* of lattice points in subvarieties.

The lifting problem is motivated by the following fundamental question regarding the distribution of prime numbers studied by Linnik. By Dirichlet's theorem, for every coprime $b, \ell \in \mathbb{N}$ one can find a prime number p satisfying

$$p = b \pmod{\ell}.$$

Y. Linnik [L44a, L44b] raised the question of whether this congruence can be solved effectively and showed that there exist $\ell_0, \sigma > 0$ such that one can find a prime p satisfying $p \leq \ell^\sigma$ for all $\ell \geq \ell_0$. This problem admits a sweeping generalization to the non-commutative set-up, namely estimating the norm of a minimal prime solution to a system of polynomial equations satisfying a congruence condition. The latter problem is out of reach at this time, but it is possible to give a general solution to the problem of bounding the norm of an almost prime solution in orbits of linear groups, as follows.

Let $\mathcal{O} \subset \mathbb{Z}^d$ be an orbit of an arithmetic lattice subgroup, and let $f : \mathcal{O} \rightarrow \mathbb{Z}$ be a polynomial function. We would like to find a small solution of the congruence

$$(6.10) \quad f(x) = b \pmod{\ell}, \quad x \in \mathcal{O},$$

with $f(x)$ being almost prime. Using the standard sifting techniques combined with the estimate (6.9), we can prove the following (see [GN12c, Theorem 1.13] for a more general result):

Theorem 6.2. *Let $G \subset \mathrm{GL}_n(\mathbb{C})$ be a connected \mathbb{Q} -simple simply connected algebraic group defined over \mathbb{Q} , $\Gamma = G(\mathbb{Z})$, and $v \in \mathbb{Z}^n$. Then given a polynomial $f : \mathbb{Z}^n \rightarrow \mathbb{Z}$ such that the function $g \mapsto f(gv)$ is absolutely irreducible on G and $\gcd(f(x) : x \in \mathcal{O}) = 1$, there exist explicit $r, \ell_0, \sigma > 0$ such that for every coprime $b, \ell \in \mathbb{N}$ satisfying $\ell \geq \ell_0$ and $b \in f(\mathcal{O}) \pmod{\ell}$, one can find x with $\|x\| \leq \ell^\sigma$ satisfying (6.10) such that $f(x)$ is r -prime.*

Let us turn to consider the phenomenon of non-concentration of lattice points in a collection of proper subvarieties X of the group variety of a semisimple group G defined over \mathbb{Q} . Namely, we seek to establish that the number of integral points on X has strictly lower rate of growth than that of the integral points in G .

A general non-concentration result can be found in [GN12c]. Here we will demonstrate it in the case of integral points on subvarieties of the special linear group SL_n , $n \geq 2$. Let $N_T(D)$ denote the number of integral points of norm at most T which is in a set $D \subset \mathrm{SL}_n(\mathbb{R})$. Recall that $N_T(\mathrm{SL}_n(\mathbb{R})) \sim \mathrm{const} \cdot T^{n^2-n}$ as $T \rightarrow \infty$. We then have

Theorem 6.3. *For every $n \geq 2$ and every absolutely irreducible proper affine subvariety $X \subset \mathrm{SL}_n(\mathbb{R})$ defined over \mathbb{Q} , we have*

$$(6.11) \quad N_T(X) \ll_{n, \deg(X), \eta} T^{n^2-n-\frac{n^2-1-\dim(X)}{(n^2+n)2n_e}+\eta}, \quad \eta > 0,$$

as $T \rightarrow \infty$, where n_e is the least even integer $\geq n - 1$.

Now let X_i , $1 \leq i \leq k$, be a collection of irreducible hypersurfaces in G . Since the number of lattice points in each hypersurface has a lower rate of growth than the number of lattice points in G , the same holds for their union. Thus, the rate of growth of the number of lattice points in the complement of these hypersurfaces is the same as the rate of growth of all lattice points. This observation gives rise to

host of results asserting that the set of integral points which are generic (i.e., that avoid the union of the hypersurfaces) has maximal possible rate of growth. Let us illustrate this principle concretely by the following example.

Proposition 6.4. *Denote by N_T the number of unimodular integral $(n \times n)$ -matrices ($n \geq 3$) of norm bounded by T , and denote by N'_T the number of such matrices satisfying*

- all the matrix entries are non-zero,
- all the principal minors do not vanish,
- all the eigenvalues are distinct,
- all the singular values (eigenvalues of $A^t A$) are distinct.

Then

$$N'_T = N_T \cdot \left(1 + O_\epsilon \left(T^{-\frac{1}{2n_e n(n+1)} + \epsilon}\right)\right), \quad \epsilon > 0,$$

where n_e is the least even integer $\geq n - 1$.

7. INTRINSIC DIOPHANTINE APPROXIMATION ON HOMOGENEOUS VARIETIES

7.1. Diophantine approximation and ergodic theory. In the classical theory of Diophantine approximation one is interested in finding an efficient rational approximation for vectors $x \in \mathbb{R}^d$. More explicitly, one would like to know for what range of parameters ϵ and R , the system of inequalities

$$\|x - r\| \leq \epsilon \quad \text{and} \quad D(r) \leq R$$

has a solution $r \in \mathbb{Q}^d$, where $\|\cdot\|$ denotes the maximum norm, and $D(r)$ denotes the common denominator of the rational numbers in the vector r when written in reduced form. In this section we consider a more general question regarding *intrinsic Diophantine approximation* on algebraic varieties, namely approximation by rational points belonging to the variety itself. Let then

$$X = \{f_1(x) = \cdots = f_k(x) = 0\} \subset \mathbb{C}^n$$

be an algebraic variety defined by a family of polynomial with rational coefficients. For a real point x in $X(\mathbb{R})$, which is contained in the closure of the set $X(\mathbb{Q})$ of rational points, we would like to produce an estimate that quantifies density. This amounts to solving the system of inequalities

$$(7.1) \quad \|x - r\| \leq \epsilon \quad \text{and} \quad D(r) \leq \epsilon^{-\kappa}$$

with $r \in X(\mathbb{Q})$, namely with the rational approximation constrained to lie on the variety itself. This question cannot of course be addressed in full generality because it is not even known how to determine whether $X(\mathbb{Q})$ is finite or infinite.

Let us note however that the problem of Diophantine approximation on a homogeneous algebraic variety by rational points on it was raised already half a century ago by S. Lang, in his 1965 survey *Report on Diophantine approximation* [L65]. When X is an elliptic curve (or, more generally, an Abelian variety), the problem of Diophantine approximation on X has been studied by M. Waldschmidt in [W99]. We have considered the problem of Diophantine approximation on an algebraic variety X equipped with a transitive action of a semisimple algebraic group G defined over \mathbb{Q} . A key new feature of our approach is that it allows approximation by rational points satisfying an arbitrary set of integrality constraints, and not only by the set of all rational points. To simplify our notation for this exposition, we consider the problem of Diophantine approximation by the rational points in

$X(\mathbb{Z}[1/p])$ where p is prime. Here $\mathbb{Z}[1/p]$ is the subring of \mathbb{Q} generated by \mathbb{Z} and $1/p$, and $X(\mathbb{Z}[1/p])$ denotes the points on the variety X with entries in this ring. The problem of Diophantine approximation by points in $X(\mathbb{Z}[1/p_1, \dots, 1/p_s])$, for instance, or by $X(\mathbb{Q})$ can be handled similarly but uses analysis on adelic spaces which would require much more elaborate notation.

Let us illustrate our general result stated below by the following example, briefly mentioned in the introduction. Let X be a rational two-dimensional ellipsoid such that $X(\mathbb{Z}[1/p])$ is not discrete in $X(\mathbb{R})$. Then we prove (see below) that

- for *almost every* $x \in X(\mathbb{R})$, every $\kappa > 2$, and $\epsilon \in (0, \epsilon_0(x, \kappa))$, the system of inequalities

$$(7.2) \quad \|x - r\| \leq \epsilon \quad \text{and} \quad D(r) \leq \epsilon^{-\kappa}$$

has a solution $r \in X(\mathbb{Z}[1/p])$;

- for *every* $x \in X(\mathbb{R})$, every $\kappa > 4$, and $\epsilon \in (0, \epsilon_0(\kappa))$, the system of inequalities

$$(7.3) \quad \|x - r\| \leq \epsilon \quad \text{and} \quad D(r) \leq \epsilon^{-\kappa}$$

has a solution $r \in X(\mathbb{Z}[1/p])$.

We note that the estimate (7.2) is essentially the best possible. Indeed, the number of $r \in X(\mathbb{Z}[1/p])$ with $D(r) \leq R$ grows as $\text{const} \cdot R$ as $R \rightarrow \infty$ and $\dim(X) = 2$, so that one can deduce from the pigeon-hole principle that $\kappa \geq 2$ in (7.2).

Let us begin by noting that it is possible to give an upper bound for the quality of approximation for a general algebraic variety X , using geometric arguments. To that end, consider the following exponent bounding the empirical distribution of $X(\mathbb{Z}[1/p])$ -points in the variety, given by

$$a_p(X) = \sup_{\text{compact } \Omega \subset X(\mathbb{R})} \limsup_{R \rightarrow \infty} \frac{\log |\{r \in \Omega \cap X(\mathbb{Z}[1/p]) : D(r) \leq R\}|}{\log(R)}.$$

It is then possible to show, using pigeon-hole packing arguments, that the exponent $\kappa = \frac{\dim(X)}{a_p(X)}$ in (7.1) cannot be improved, namely no smaller κ will do in (7.2). It is a compelling challenge to give a matching upper bound that establishes $\kappa = \frac{\dim(X)}{a_p(X)}$ as the best possible result.

Before we proceed to describe our approach to establishing upper bounds on κ , let us note that the spheres constitute the only example for which we are aware of previous results in the literature, as follows. Diophantine approximation results by points in $X(\mathbb{Z}[\frac{1}{p}])$ for $X = S^2$ and $X = S^3$ can be deduced from the celebrated construction by Lubotzky, Phillips, and Sarnak of a dense subgroup of $\text{SO}(3, \mathbb{R})$ with entries in $\mathbb{Z}[\frac{1}{p}]$ possessing the optimal spectral gap [LPS86, LPS87]. A rate of uniform approximation by *all rational points* in the spheres S^d , $d \geq 2$, was established by Schmutz [Sch08] using elementary methods based on rational parametrizations of the spheres, the rate being given by $\kappa = 2 \log_2(d + 1)$. Duke [Du03] has established equidistribution of the set of all rational points on S^2 , from which a rate of uniform approximation can be derived, and the same method applies to spheres of any dimension. Finally, we mention the remarkable very recent results of Kleinbock and Merrill [KM], which give the best possible results for approximation by all rational points on spheres of any dimension, establishing complete analogs of Dirichlet's and Khinchin's theorems in this case.

Turning now to a general variety, in order to study the density of the set $X(\mathbb{Z}[1/p])$ in $X(\mathbb{R})$, we construct a suitable dynamical system and reduce our problem to the study of the distribution of certain orbits of a suitable group in it. For example, in the sphere case the underlying algebraic group we consider is $G = \text{SO}_3$, the group preserving the standard sum-of-squares quadratic form in three variables, and $\text{SO}_3(\mathbb{Z}[1/p])$ acts on the homogeneous space S^2 with dense orbits (when $p \cong 1$ modulo 4).

In general, the reduction of our Diophantine approximation problem on the variety $X = H \backslash G$ to a dynamical one is possible because the set $X(\mathbb{Z}[1/p])$ is parametrized by the orbits of the group $G(\mathbb{Z}[1/p])$. It is crucial for the classical theory of Diophantine approximation that \mathbb{Z}^d forms a lattice \mathbb{R}^d , and in order to have a similar framework in our setting, we consider the topological group $G(\mathbb{R}) \times G(\mathbb{Q}_p)$ (where \mathbb{Q}_p denotes the field of p -adic numbers) equipped with an invariant measure. Then the group $G(\mathbb{Z}[1/p])$ embeds diagonally in $G(\mathbb{R}) \times G(\mathbb{Q}_p)$ as a discrete subgroup with finite covolume. We consider the space

$$Y = (G(\mathbb{R}) \times G(\mathbb{Q}_p))/G(\mathbb{Z}[1/p]).$$

The group $G(\mathbb{Q}_p)$ acts naturally on this space preserving the finite measure. This is the dynamical system that plays a crucial role in our considerations. It turns out that from the dynamical systems point of view, the original problem of Diophantine approximation corresponds to a *shrinking target property* (see Figure 1). Suppose that we have an increasing sequence of compact subsets B_n of $G(\mathbb{Q}_p)$ and a shrinking sequence of neighborhoods O_ϵ in Y . The shrinking target problem asks for what range of parameters n and ϵ the partial orbits $B_n y$ with $y \in Y$ reach the neighborhoods O_ϵ . While in the classical theory of dynamical systems, this question is asked for a one-parameter flow, here we are required to answer a similar question for orbits of the semisimple group $G(\mathbb{Q}_p)$.

We remark that the connection between Diophantine approximation in \mathbb{R}^d and the shrinking target property of one-parameter flows has been previously exploited by S. G. Dani, D. Kleinbock, and G. Margulis [D85, D86, D89, K98a, K98b, KM98, K99, KM99]. These developments are surveyed in [K01, M02, K10]. Also in the context of negatively curved manifolds, this connection was explored in [HP01, HP02a, HP02b].

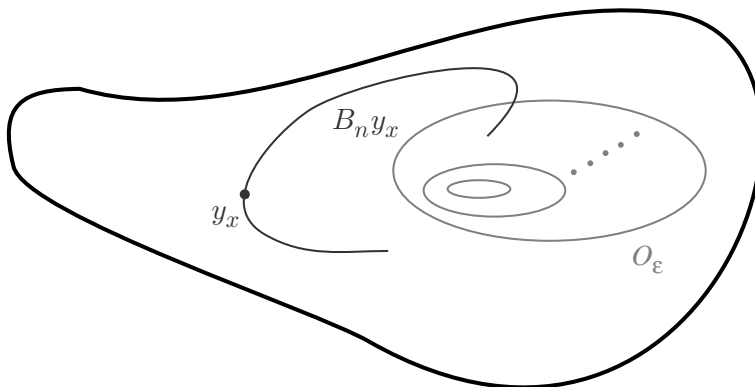


FIGURE 1. Shrinking target property

TABLE 1. Dynamical correspondence

Diophantine approximation	Dynamics
$X(\mathbb{Z}[1/p]) \subset X(\mathbb{R})$	$G(\mathbb{Q}_p) \curvearrowright Y$
$x \in X(\mathbb{R})$	$y_x \in Y$
$\ x - r\ \leq \epsilon$	$O_\epsilon \subset Y$
$\left\{ \begin{array}{l} \ x - r\ \leq \epsilon, \\ D(r) \leq p^n \end{array} \right. \text{ has a } \mathbb{Z}[1/p]\text{-solution}$	$B_n^{-1} y_x \cap O_\epsilon \neq \emptyset$

7.2. Diophantine approximation and the duality principle. To address the problem of Diophantine approximation in our setting, we consider the p -adic balls

$$(7.4) \quad B_n = \{g \in G(\mathbb{Q}_p) : \|g\|_p \leq p^n\}$$

and construct a suitable shrinking family of neighborhoods O_ϵ in Y . An important property of the family O_ϵ is that it models the family of neighborhoods $\|x - r\| \leq \epsilon$ in $X(\mathbb{R})$ and, in particular, $|O_\epsilon|$ decays as $\text{const} \cdot \epsilon^{\dim(X)}$ when $\epsilon \rightarrow 0^+$. One can show that every point $x \in X(\mathbb{R})$ is naturally associated to a point $y_x \in Y$, such that the system of inequalities

$$\|x - r\| \leq \epsilon, \quad D(r) \leq p^n,$$

has a solution $r \in X(\mathbb{Z}[1/p])$ provided that

$$(7.5) \quad B_n^{-1} y_x \cap O_\epsilon \neq \emptyset.$$

This leads to a useful connection (summarized in Table 1) between the problem of Diophantine approximation on X and the distribution of orbits of Y . In the classical setting, an analogous correspondence was observed by S. G. Dani in [D85], and was called the Dani correspondence in [KM98].

7.3. Spectral estimates. We prove the shrinking target property by analyzing the behavior of suitable averaging operators on the space Y . Let β_n denote the uniform probability measure supported on the set $B_n \subset G(\mathbb{Q}_p)$. Then we have the sequence of averaging operators

$$\pi_Y(\beta_n) : L^2(Y) \rightarrow L^2(Y)$$

defined by

$$\pi_Y(\beta_n)\phi(y) = \int_{G(\mathbb{Q}_p)} \phi(g^{-1}y) d\beta_n(g), \quad \phi \in L^2(Y).$$

Ideally, one hopes that, as in the classical von Neumann ergodic theorem, the averages $\pi_Y(\beta_n)\phi$ converge to the space average $\int_Y \phi$ as $n \rightarrow \infty$, but this is not always true due to presence of non-trivial unitary characters of $G(\mathbb{Q}_p)$ in the space $L^2(Y)$. We denote by $L_{00}^2(Y)$ the subspace of $L^2(Y)$ orthogonal to all characters of $G(\mathbb{R}) \times G(\mathbb{Q}_p)$. Then one can show that $\pi_Y(\beta_n)\phi \rightarrow 0$ in L^2 -norm for every $\phi \in L_{00}^2(Y)$. The rate of convergence here will play a crucial role. This rate is deduced from the spectral gap property, which was discussed in Sections 4 and 6, and is determined by a suitable integrability exponent. We fix a good maximal compact subgroup U_p of $G(\mathbb{Q}_p)$. For functions $\phi, \psi \in L^2(Y)$, we denote by $c_{\phi, \psi}(g) = \int_Y \phi(gy) \overline{\psi(y)} dy$ the corresponding matrix coefficient. Similarly to (3.8) we define

the *spherical integrability exponent*

$$q_p(\mathbf{G}) = \inf \{q > 0 : c_{\phi, \psi} \in L^q(\mathbf{G}(\mathbb{Q}_p)) \text{ for all } U_p\text{-inv. } \phi, \psi \in L^2_{00}(Y)\}.$$

It follows from the result of L. Clozel [C03] that the integrability exponent $q_p(\mathbf{G})$ is finite, and we establish the following quantitative mean ergodic theorem. For every sequence B_n of U_p -bi-invariant compact subsets of $\mathbf{G}(\mathbb{Q}_p)$, every $\phi \in L^2_{00}(Y)$ and $\delta > 0$,

$$(7.6) \quad \|\pi_Y(\beta_n)(\phi)\|_2 \leq \text{const}(\delta) \cdot |B_n|^{-\frac{1}{q_p(\mathbf{G})} + \delta} \|\phi\|_2.$$

This mean ergodic theorem is the main tool for establishing the required shrinking target property. We note that in order to derive optimal (or close to optimal) exponents for Diophantine approximation in (7.1), it is crucial that the error term is controlled by the L^2 -norm rather than a Sobolev norm, which compromises the rate. The actual value of the exponent with which $|B_n|$ appears in (7.6), namely the rate of convergence in the mean ergodic theorem, is the key parameter which ultimately controls the quality of Diophantine approximation, as elaborated further below. Thus this approach provides a connection between the spectral bounds in the automorphic representation (or generalized Ramanujan conjectures) and the quality of intrinsic Diophantine approximation on homogeneous varieties.

Now coming back to the original Diophantine approximation problem, we apply estimate (7.6) to the averages supported on the sets B_n as in (7.4) (or, more precisely, on the U_p -bi-invariant sets $U_p B_n U_p$). We must also estimate the behavior of these averages on the orthogonal complement of $L^2_{00}(Y)$ in $L^2(Y)$, but since the orthogonal complement is finite dimensional, this does not present a particular challenge. Assuming that

$$|O_\epsilon| \geq \text{const} \cdot |B_n|^{-\frac{2}{q_p(\mathbf{G})} + \delta}$$

with fixed $\delta > 0$, we use (7.6) to produce an upper estimate on the measure of the set $y \in Y$ for which $B_n^{-1}y \cap O_\epsilon = \emptyset$. Then using a Borel–Cantelli type argument, we deduce a shrinking target property that holds for almost every $y \in Y$. This implies a Diophantine approximation result via the dynamical correspondence (Table 1). If we impose the stronger assumption that

$$|O_\epsilon| \geq \text{const} \cdot |B_n|^{-\frac{1}{q_p(\mathbf{G})} + \delta}$$

with some fixed $\delta > 0$, then the above argument can be modified to show that $B_n^{-1}y \cap O_\epsilon = \emptyset$ for all $y \in Y$ provided that ϵ is sufficiently small. This produces a Diophantine approximation result which is valid for all points in $\overline{X(\mathbb{Z}[1/p])}$.

This completes the outline of the following result, which is part of a more general theorem stated in [GGN1, Theorem 1.6].

Theorem 7.1. *Let X be an algebraic variety defined over \mathbb{Q} equipped with a transitive action of connected almost simple algebraic group \mathbf{G} defined over \mathbb{Q} . We assume that $X(\mathbb{Z}[1/p])$ is not discrete in $X(\mathbb{R})$.** Then

*One can show that under these assumptions $\overline{X(\mathbb{Z}[1/p])}$ is open in $X(\mathbb{R})$.

- for almost every $x \in \overline{X(\mathbb{Z}[1/p])} \subset X(\mathbb{R})$, every $\delta > 0$, and $\epsilon \in (0, \epsilon_0(x, \delta))$, the system of inequalities

$$(7.7) \quad \|x - r\| \leq \epsilon \quad \text{and} \quad D(r) \leq \left(\epsilon^{-\frac{\dim(X)}{a_p(G)} - \delta} \right)^{q_p(G)/2}$$

has a solution $r \in X(\mathbb{Z}[1/p])$;

- for every $x \in \overline{X(\mathbb{Z}[1/p])} \subset X(\mathbb{R})$, every $\delta > 0$, and $\epsilon \in (0, \epsilon_0(\delta))$, the system of inequalities

$$\|x - r\| \leq \epsilon \quad \text{and} \quad D(r) \leq \left(\epsilon^{-\frac{\dim(X)}{a_p(G)} - \delta} \right)^{q_p(G)}$$

has a solution $r \in X(\mathbb{Z}[1/p])$.

We remark that in the paper [GGN2] (joint with A. Ghosh), we also prove analogues of Khinchin's and Jarnik's theorems for Diophantine approximation on homogeneous varieties.

The quality of Diophantine approximation in Theorem 7.1 is controlled by the integrability exponent $q_p(G)$. When $q_p(G) = 2$, then estimate (7.7) gives the best possible result up to arbitrary small $\delta > 0$. Coming back to the example (7.2), we note that in this case G is the orthogonal group and $q_p(G) = 2$, which follows from the work of P. Deligne [D71] on the classical Ramanujan conjecture (see [C] for a comprehensive treatment) combined with the Jacquet–Langlands correspondence [JL70]. A detailed survey can be found in [R].

In conclusion we note the intriguing fact that conversely, Diophantine approximation considerations can provide lower bounds towards the generalized Ramanujan conjectures; namely, one can deduce estimates on the integrability exponents in the automorphic representation. For instance, analyzing the quality of Diophantine approximation in \mathbb{R}^d by the rational points in $\mathbb{Z}[1/p]^d$ directly, one can deduce that

$$q_p(\mathrm{SL}_d) \geq 2(d-1).$$

We refer to [GGN1, Corollary 1.8] for a more general result in this direction.

8. ERGODIC THEOREMS FOR LATTICE SUBGROUPS

8.1. Quantitative ergodic theorems and fast equidistribution. We turn to discuss ergodic theorems for countable groups such as $\Gamma = \mathrm{SL}_n(\mathbb{Z})$ and more generally for arithmetic subgroups in semisimple algebraic groups. When G is a connected semisimple group, we have approached this problem via spectral methods, and detailed information about the unitary representation theory of G was essential. However, for discrete lattice subgroups this approach is not feasible, since the irreducible unitary representations of a lattice cannot be classified (in a Borel manner), and a general unitary representation of a lattice cannot be represented uniquely as a direct integral of irreducible ones. Thus, non-amenable discrete groups appear as the hardest case in ergodic theory because they possess neither an asymptotically invariant family of sets nor a usable spectral theory, and so other approaches must be developed.

Although our approach can be applied to lattices in general groups, for the sake of keeping our exposition simple, we will present it for lattices Γ in connected simple Lie groups G . For more general formulations, we refer to [GN10]. We take a family of increasing compact domains B_t in G , set $\Gamma_t = \Gamma \cap B_t$, and denote by

λ_t the uniform probability measures supported on the sets Γ_t . Given an arbitrary measure-preserving ergodic action of Γ on a (standard) probability space (X, μ) , we consider the corresponding averaging operators

$$\pi_X(\lambda_t)f(x) = \frac{1}{|\Gamma_t|} \sum_{\gamma \in \Gamma_t} f(\gamma^{-1}x), \quad f \in L^p(X).$$

In order for the sets Γ_t to behave in a regular manner, we assume that the domains B_t are admissible in the sense of Definition 4.2. This mild regularity assumption turns out to be sufficient to establish general quantitative ergodic theorems:

Theorem 8.1. *Assume that the action of Γ on (X, μ) has a spectral gap. Then the following hold.*

- Mean ergodic theorem. *For every $f \in L^p(X)$ with $1 < p < \infty$,*

$$\left\| \pi_X(\lambda_t)f - \int_X f d\mu \right\|_p \leq C_p |\Gamma_t|^{-\theta_p} \|f\|_p$$

with uniform $C_p, \theta_p > 0$, independent of f .

- Pointwise ergodic theorem. *For every $f \in L^p(X)$ with $p > 1$ and for μ -almost every x ,*

$$\left| \pi_X(\lambda_t)f - \int_X f d\mu \right| \leq C_p(f, x) |\Gamma_t|^{-\theta'_p}$$

with uniform $\theta'_p > 0$, independent of f and x .

We emphasize that this result holds for all actions of the discrete group Γ which have a spectral gap, and furthermore, if Γ has property (T) of Kazhdan, the parameters C_p, θ_p , and θ'_p are independent of the action. The actions have nothing to do with G , in general, and in fact the only connection of the statements in Theorem 8.1 to the fact that Γ is a lattice in G is in the definition of the sets Γ_t . We remark that $C_p(f, \cdot)$ satisfies an L^s -norm bound, for a suitable s depending on p . Finally, when the action of Γ does not have a spectral gap, we also establish mean and pointwise ergodic theorem (but without a rate of convergence, of course). We refer to [GN10] for details.

Recalling the discussion in the Introduction, note that Theorem 8.1 exhibits a remarkable contrast to the ergodic theory of amenable groups, where no rate of convergence can be established. Let us give an example to illustrate this point.

- Consider the action $\Gamma = \mathrm{SL}_n(\mathbb{Z})$ on the torus $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$, and set

$$\Gamma_t = \{\gamma \in \Gamma : \log \|\gamma\| \leq t\}.$$

Then for every $f \in L^p(\mathbb{T}^n)$ with $p > 1$ and almost every $x \in \mathbb{T}^n$,

$$\left| \frac{1}{|\Gamma_t|} \sum_{\gamma \in \Gamma_t} f(\gamma^{-1}x) - \int_{\mathbb{T}^n} f d\mu \right| \leq C_p(f, x) |\Gamma_t|^{-\theta'_p}$$

with uniform $\theta'_p > 0$.

As is well known, the orbit structure in the above example is very complicated. The torus contains the dense set of rational points which consists of finite Γ -orbits, but the orbit of every irrational point is dense. In particular, the size of the estimator $C_p(f, x)$ depends very sensitively on Diophantine properties of x , and it would be

interesting to make this dependence more explicit. In the setting of random walks on the torus, a similar problem was recently investigated in [BFLM11].

In the case of isometric actions of lattices, we derive the following quantitative equidistribution result, which again has no classical amenable analogue:

Corollary 8.2. *Let Γ (as in Theorem 8.1) act isometrically on a compact Riemannian manifold X equipped with a smooth measure μ of full support. If the action has spectral gap, then for every Hölder function $f \in C^a(X)$ and for every point x in the manifold,*

$$\frac{1}{|\Gamma_t|} \sum_{\gamma \in \Gamma_t} f(\gamma^{-1}x) = \int_X f d\mu + O(|\Gamma_t|^{-\theta_a} \|f\|_{C^a})$$

with uniform $\theta_a > 0$.

8.2. The method of proof of the ergodic theorems. Our approach to proving the ergodic theorems for the averages $\pi_X(\lambda_t)$ is based on the idea of inducing the action of Γ to obtain an action of G on a larger space $Y = (G \times X)/\Gamma$. The latter space has the structure of a fiber bundle over the space G/Γ with fibers isomorphic to X . We then reduce the ergodic theorems for the averaging operators $\pi_X(\lambda_t)$, supported on $\Gamma \cap B_t$ and acting on $L^p(X)$, to the ergodic theorems for the averaging operators $\pi_Y(\beta_t)$, supported on the domains B_t and acting on $L^p(Y)$. This involves developing a series of approximation arguments which constitute generalizations of the argument used in the solution of the lattice point counting problem in Section 5. The essence of the matter is to approximate

$$\pi_X(\lambda_t)\phi(x) = \frac{1}{|\Gamma \cap B_t|} \sum_{\gamma \in \Gamma \cap B_t} \phi(\gamma^{-1}x), \quad \phi \in L^p(X),$$

by

$$\pi_Y(\beta_{t \pm c\varepsilon})f_\varepsilon(y) = \frac{1}{|B_{t \pm c\varepsilon}|} \int_{B_{t \pm c\varepsilon}} f_\varepsilon(g^{-1}y) dg$$

with suitably chosen $f_\varepsilon \in L^p(Y)$. The link between the two expressions is given by setting $y = (h, x)\Gamma$ and

$$f_\varepsilon((h, x)\Gamma) = \sum_{\gamma \in \Gamma} \chi_\varepsilon(h\gamma)\phi(\gamma^{-1}x),$$

where χ_ε is the normalized characteristic function of an identity neighborhood O_ε .

The ergodic averages $\pi_Y(\beta_{t \pm c\varepsilon})f_\varepsilon$ can then be rewritten in full as

$$\sum_{\gamma \in \Gamma} \left(\frac{1}{|B_{t \pm c\varepsilon}|} \int_{B_{t \pm c\varepsilon}} \chi_\varepsilon(g^{-1}h\gamma) dg \right) \phi(\gamma^{-1}x).$$

When $h \in O_\varepsilon$, we would like the expression in parentheses to be equal to 1 when $\gamma \in \Gamma \cap B_{t-c\varepsilon}$ and equal to 0 when $\gamma \notin \Gamma \cap B_{t+c\varepsilon}$, in order to be able to compare it to $\pi_X(\lambda_t)\phi$. These favorable lower and upper estimates depend only on the regularity properties of the sets B_t , namely, on the admissibility property. Indeed, the lower bound arises since if $\chi_\varepsilon(g^{-1}h\gamma) \neq 0$ and $g \in B_{t-c\varepsilon}$, then $\gamma \in B_t$. The upper bound holds since for $\gamma \in \Gamma \cap B_t$, the support of function $g \mapsto \chi_\varepsilon(g^{-1}h\gamma)$ is contained in $B_{t+c\varepsilon}$.

A fundamental role in the completion of the proof of the ergodic theorems is played by an *invariance principle* which holds in our context. Indeed, for any

given $f \in L^p(Y)$, the pointwise ergodic theorem for the averages $\pi_Y(\beta_t)$ holds for a set of points which contains a G -invariant set of full measure. Since the space $Y = (G \times X)/\Gamma$ is a G -equivariant bundle over the G -transitive space G/Γ , this implies that for every point $h\Gamma$, the set of points $x \in X$ in the fiber above it where the pointwise ergodic theorem holds is conull in X . This allows us to deduce that the set of points where $\pi_X(\lambda_t)\phi(x)$ converges is also a set of full measure in X .

9. DISTRIBUTION OF ORBITS ON ALGEBRAIC VARIETIES

9.1. Distribution of orbits in the de-Sitter space. Let us now turn to discuss the distribution of dense orbits for actions of arithmetic groups on algebraic varieties. These actions do not preserve a finite measure (besides the measures supported on periodic points), so that this problem falls in the realm of infinite ergodic theory.

Consider an action of a discrete group Γ on a space X , and anticipating our arguments below, we consider here a right action. The distribution of orbits of Γ is encoded by the properties of the averaging operators

$$(9.1) \quad \mathcal{A}_t\phi(x) = \frac{1}{|\Gamma_t|} \sum_{\gamma \in \Gamma_t} \phi(x\gamma)$$

defined for an increasing sequence of finite subsets of Γ that exhaust Γ and for a family of test functions ϕ on X . For example, if $\Gamma = \langle \gamma \rangle$ is a cyclic group and the action preserves a (possibly infinite) measure μ , then as in Section 1 we obtain that

$$(9.2) \quad \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} \phi(x\gamma^n) \rightarrow \mathcal{P}\phi(x) \quad \text{in } L^2(X),$$

where \mathcal{P} denotes the orthogonal projection from $L^2(X)$ to the subspace of Γ -invariant functions in $L^2(X)$. However, if the measure is infinite and the action is ergodic, then $\mathcal{P} = 0$, and (9.2) contains no information about the distribution of orbits. Initially, one might expect that if we pick a correct normalization factor such that $\frac{a(N)}{N} \rightarrow 0$, then the averages $\frac{1}{a(N)} \sum_{n=0}^{N-1} \phi(x\gamma^n)$ may exhibit a non-trivial limit. But in fact, J. Aaronson [Aa, §2.4] showed that if the action of the cyclic group $\Gamma = \langle \gamma \rangle$ is ergodic and conservative, then for any normalization $a(N)$, either for every non-negative $\phi \in L^1(X)$,

$$\liminf_{N \rightarrow \infty} \frac{1}{a(N)} \sum_{n=0}^{N-1} \phi(x\gamma^n) = 0 \quad \text{almost everywhere,}$$

or there exists a subsequence N_k such that such that for every non-negative $\phi \in L^1(X)$, $\phi \neq 0$,

$$\lim_{k \rightarrow \infty} \frac{1}{a(N_k)} \sum_{n=0}^{N_k-1} \phi(x\gamma^n) = \infty \quad \text{almost everywhere.}$$

However, in contrast to averages on the orbits of \mathbb{Z} -actions on infinite measure spaces, it turns out that averages on orbits of our lattice subgroups behave in a more regular fashion. Let us demonstrate this point by analyzing the distribution of the orbits of lattices in the groups of isometries of the d -dimensional de-Sitter space dS_d (see Figure 2). This space can be realized as a hypersurface

$$x_1^2 + \cdots + x_d^2 - x_{d+1}^2 = 1$$

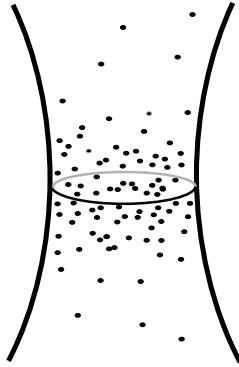


FIGURE 2. The de-Sitter space

in \mathbb{R}^{d+1} equipped with the Minkowski metric $ds^2 = dx_1^2 + \cdots + dx_d^2 - dx_{d+1}^2$, and the problem of distribution of lattice orbits on it was raised by V. Arnol'd [Ar, 1996-15, 2002-16]. After stating the result, we will explain a general strategy for how to analyze the distribution of orbits on infinite volume homogeneous spaces. Our exposition is based on [GW07, GN]. We also refer to [L99, N02, Ma02, LP03, G03, G04, GM05, Gu10, N10, MW11] for related results.

In the setting of Arnol'd's question regarding the de-Sitter space, the following holds. Let Γ be a discrete subgroup of $\text{Isom}(dS_d)$ with finite covolume. We identify $\text{Isom}(dS_d)$ with the orthogonal group $\text{SO}(d, 1)$ and denote by $\|\cdot\|$ the standard Euclidean norm on the space of $(d+1)$ -dimensional matrices.

- When $d = 2$, for every $\phi \in L^1(dS^d)$ with compact support,

$$(9.3) \quad \lim_{t \rightarrow \infty} \frac{1}{t} \sum_{\gamma \in \Gamma: \log \|\gamma\| \leq t} \phi(x\gamma) = \int_{dS^d} \phi(y) d\nu(y) \quad \text{almost everywhere,}$$

where ν denotes a (non-zero) invariant measure on dS_d (whose normalization depends on Γ) [GN, Thm. 1.1].

- When $d \geq 3$, for every continuous function ϕ with compact support and every $x \in dS^d$ with dense Γ -orbit,

$$(9.4) \quad \lim_{t \rightarrow \infty} \frac{1}{e^{(d-2)t}} \sum_{\gamma \in \Gamma: \log \|\gamma\| \leq t} \phi(x\gamma) = \int_{dS^d} \phi(y) d\nu_x(y),$$

where the limit measure is $d\nu_x(y) = \frac{d\nu(y)}{(1+\|x\|^2)^{(d-2)/2}(1+\|y\|^2)^{(d-2)/2}}$ [GW07].

Moreover, if ϕ is of compact support and in L^p , $p > 1$, then for almost every $x \in dS^d$,

$$(9.5) \quad \frac{1}{e^{(d-2)t}} \sum_{\gamma \in \Gamma: \log \|\gamma\| \leq t} \phi(x\gamma) = \int_{dS^d} \phi(y) d\nu_x(y) + O_{\phi, x}(e^{-\delta_p t})$$

with $\delta_p > 0$ independent of x and ϕ [GN, Ex. 11.1].

The foregoing results have several remarkable features. Note that in the case of dimension 2, while the cardinality of the set $\{\gamma \in \Gamma : \log \|\gamma\| \leq t\}$ grows exponentially as $\text{const} \cdot e^t$, it turns out that only for a polynomial number of points γ do the orbit points γx in (9.3) come back to a compact set. Nonetheless, the

small fraction of points returning by time t becomes almost surely equidistributed in the compact set, with respect to the invariant measure, as $t \rightarrow \infty$.

When $d \geq 3$, while the cardinality of the set $\{\gamma \in \Gamma : \log \|\gamma\| \leq t\}$ grows as $\text{const} \cdot e^{(d-1)t}$, only an exponentially small fraction of them satisfy that γx returns to a compact set. The set of returning points does become almost surely equidistributed in the compact set, but this time the limiting measures ν_x are not invariant under Γ and, furthermore, they depend non-trivially on x . The measures ν_x , $x \in X$, on the pseudo-Riemannian space X should be considered as analogues of the Patterson–Sullivan measures in this context.

9.2. The duality principle on homogeneous spaces. The techniques developed in [GN] to analyze action on infinite-volume homogeneous spaces apply in the following general setting. Let $G \subset \text{SL}_d(\mathbb{R})$ be connected and of finite index in an algebraic group, and let Γ be a discrete subgroup of G with finite covolume. Let X be an algebraic homogeneous space of G equipped with a G -invariant smooth measure on which Γ acts ergodically. We fix a proper positive homogeneous polynomial $p : \text{M}_d(\mathbb{R}) \rightarrow [0, \infty)$ and consider the sets

$$\Gamma_t = \{\gamma \in \Gamma : \log p(\gamma) \leq t\}.$$

Our goal is to describe the asymptotic distribution of orbits of Γ or, in other words, the asymptotic behaviour of the sums $\sum_{\gamma \in \Gamma_t} \phi(x\gamma)$ as $t \rightarrow \infty$ for a sufficiently rich collection of functions ϕ on X with compact support. In order to find the right normalization for the sum $\sum_{\gamma \in \Gamma_t} \phi(x\gamma)$, one needs to compute what proportion of points in the orbit returns to compact subsets $\Omega \subset X$. Given a compact subset Ω of X with non-empty interior and $x \in X$, we set

$$a = \limsup_{t \rightarrow \infty} \frac{\log |x\Gamma_t \cap \Omega|}{t}.$$

One can check that this quantity is independent of the choices of Ω and almost every x . We will distinguish two cases according to whether $a = 0$, i.e., the return rate is subexponential, in which case it is in fact polynomial, and $a > 0$, i.e., the return rate is exponential.

Theorem 9.1 (Polynomial return rate). *Assume that $a = 0$. Then there exists $b \in \mathbb{N}_{>0}$ such that the averaging operator*

$$\mathcal{A}_t \phi(x) = \frac{1}{t^b} \sum_{\gamma \in \Gamma_t} \phi(x\gamma)$$

satisfies the following.

- Strong maximal inequality. *For every $p > 1$, a compact $D \subset X$, and $\phi \in L^p(D)$,*

$$\left\| \sup_{t \geq t_0} |\mathcal{A}_t \phi| \right\|_{L^p(D)} \leq \text{const}(p, D) \cdot \|\phi\|_{L^p(D)}.$$

- Pointwise ergodic theorem. *For every $\phi \in L^1(X)$ with compact support,*

$$\lim_{t \rightarrow \infty} \mathcal{A}_t \phi(x) = \int_X \phi d\nu \quad \text{almost everywhere,}$$

where ν is a (non-)zero G -invariant measure on X .

When $a > 0$, we prove analogous results for functions lying in a suitable Sobolev space $L^p_l(X)$. In this case it is also possible to establish a rate of convergence.

Theorem 9.2 (Exponential return rate). *Assume that $a > 0$. Then there exist $b \in \mathbb{N}_{\geq 0}$ and $l \in \mathbb{N}_{\geq 0}$ such that the averaging operators*

$$\mathcal{A}_t \phi(x) = \frac{1}{e^{at} t^b} \sum_{\gamma \in \Gamma_t} \phi(x\gamma)$$

satisfy the following.

- Strong maximal inequality. For every $p > 1$, a compact $D \subset X$, and a non-negative $\phi \in L^p_l(D)$,

$$\left\| \sup_{t \geq t_0} |\mathcal{A}_t \phi| \right\|_{L^p(D)} \leq \text{const}(p, D) \cdot \|\phi\|_{L^p_l(D)}.$$

- Pointwise ergodic theorem. For every $p > 1$ and a non-negative bounded $\phi \in L^p_l(X)$ with compact support,

$$(9.6) \quad \lim_{t \rightarrow \infty} \mathcal{A}_t \phi(x) = \int_X \phi d\nu_x \quad \text{almost everywhere,}$$

where ν_x , $x \in X$, is a family of absolutely continuous measures on X with positive densities.

- Quantitative pointwise ergodic theorem. For every $p > 1$ and a non-negative continuous subanalytic $\phi \in L^p_l(X)$ with compact support,

$$(9.7) \quad \mathcal{A}_t \phi(x) = \int_X \phi d\nu_x + \sum_{i=1}^b c_i(x, \phi) t^{-i} + O_{\phi, x}(e^{-\delta t}) \quad \text{almost everywhere}$$

with $\delta = \delta(x, \phi) > 0$. (Note that when $b = 0$, the sum should be interpreted as zero.)

We remark that the formulation of Theorem 9.2 is optimal in the generality in which it is stated; see [GN, §1.6.1]. However, if we additionally assume that the stabilizer of a point x in X is semisimple, then the stated results can be considerably improved. First let us recall that it has been shown in [GW07] that for continuous functions ϕ with compact support, the asymptotic formula (9.6) holds for all $x \in X$ whose Γ -orbit in X is dense. Second, it is possible to formulate a suitable quantitative ergodic theorem for all functions of bounded support which are in the Lebesgue spaces L^p , as follows (see [GN, Theorem 1.5]).

Theorem 9.3. *Assume that the following conditions are satisfied:*

- the stability group $H = \text{Stab}_G(x) \subset G$, $x \in X$, is semisimple and every simple component has a spectral gap in $L^2_0(G/\Gamma)$,
- the sets Γ_t are defined by a norm $p(x) = \|x\|$ on $\text{Mat}_d(\mathbb{R})$,
- the volumes of H_t satisfy $\text{vol}(H_t) \sim c \cdot e^{at}$ as $t \rightarrow \infty$, with $a, c > 0$.

Then the normalized sampling operators

$$\mathcal{A}_t \phi(x) = \frac{1}{e^{at}} \sum_{\gamma \in \Gamma_t} \phi(x\gamma)$$

satisfy, for every compact domain $D \subset X$ and $t \geq t_0$, the following.

- Quantitative mean ergodic theorem in Lebesgue spaces. For every $1 < p < \infty$ and $\phi \in L^p(D)$,

$$\left\| \mathcal{A}_t \phi(x) - \int_X \phi d\nu_x \right\|_{L^p(D)} \ll_{p,D} \|\phi\|_{L^p(D)} e^{-\delta_p t},$$

where $\delta_p > 0$ is independent of ϕ and D .

- Quantitative maximal ergodic theorem in Lebesgue spaces. For every $1 < p < q \leq \infty$ and $\phi \in L^p(D)$,

$$\left\| \sup_{t \geq t_0} e^{\delta_{p,q} t} \left| \mathcal{A}_t \phi(x) - \int_D \phi d\nu_x \right| \right\|_{L^p(D)} \ll_{p,q,D} \|\phi\|_{L^q(D)},$$

where $\delta_{p,q} > 0$ is independent of ϕ and D .

- Quantitative pointwise ergodic theorem in Lebesgue spaces. For every $1 < p < \infty$, $\phi \in L^p(D)$, and almost every $x \in X$,

$$\left| \mathcal{A}_t \phi(x) - \int_X \phi d\nu_x \right| \ll_{x,\phi,D} e^{-\delta_p t},$$

where $\delta_p > 0$ is independent of x , ϕ and D .

Again the formulation of Theorem 9.3 is optimal, up to a precise determination of the rate of convergence.

Finally, note that the solution to Arnol'd's problem on de-Sitter space stated in Section 9.1 follows from Theorem 9.1 in the case $d = 2$ and from Theorem 9.3 in the case $d \geq 3$.

9.3. Equidistribution in algebraic number fields. It seems unlikely that in the generality of Theorem 9.3, the error estimate stated in its third part can be proved for an explicit set of $x \in X$ of (say) positive measure, but there is an important case when this can be achieved. Assuming that the action of Γ on X preserves a Riemannian metric, one can show (see [GN12b]) that all Γ -orbits in X become equidistributed with a uniform rate. The observation that, in the presence of a mean ergodic theorem, all orbits of isometric actions on compact spaces are equidistributed goes back to the work of Y. Guivarc'h [G69], and in [GN12b] we have developed a quantitative version of this argument which also applies to the infinite-measure setting. For instance, fix a non-square integer $d > 0$, and let us consider the action of $\Gamma = \mathrm{SL}_2(\mathbb{Z}[\sqrt{d}])$ on the upper half-plane \mathbb{H}^2 by fractional linear transformations. We set

$$\Gamma_t = \{\gamma \in \Gamma : \log(\|\gamma\|^2 + \|\bar{\gamma}\|^2) \leq t\},$$

where $\|\cdot\|$ is the standard Euclidean norm and $\bar{\gamma}$ denotes the Galois involution of the field $\mathbb{Q}(\sqrt{d})$. Since Γ embeds diagonally in $G = \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$ as an irreducible lattice and \mathbb{H}^2 is a homogeneous space of the group G , Theorem 9.2 applies in this case and gives information about the asymptotic distribution of orbits Γx for almost every $x \in X$. Using this fact, the method of [GN12b, Theorem 1.5] leads to the following sharp equidistribution result: there exists $\delta > 0$ such that given a compact subset $D \subset \mathbb{H}^2$, for every $x \in D$ and every $\phi \in C^1(\mathbb{H}^2)$ with support contained in D ,

$$\frac{1}{e^t} \sum_{\gamma \in \Gamma_t} \phi(x\gamma) = \int_{\mathbb{H}^2} \phi(z) d\nu(z) + O_D(e^{-\delta t}),$$

where ν is a (non-zero) G -invariant measure on \mathbb{H}^2 .

Of course, it follows immediately that a quantitative ratio equidistribution result holds. Namely, if ψ satisfies the same properties as ϕ and is non-negative with $\int_{\mathbb{H}^2} \psi(z) d\nu(z) > 0$, then for every $x \in D$,

$$\frac{\sum_{\gamma \in \Gamma_t} \phi(x\gamma)}{\sum_{\gamma \in \Gamma_t} \psi(x\gamma)} = \frac{\int_{\mathbb{H}^2} \phi(z) d\nu(z)}{\int_{\mathbb{H}^2} \psi(z) d\nu(z)} + O_D(e^{-\delta't}).$$

Similar considerations apply also to the action of Γ on G by right translations, and this establishes a quantitative form of the fast ratio equidistribution result in algebraic number fields alluded to in the introduction.

9.4. Ideas of proof of quantitative duality results. Let us turn to indicate some ideas that lie behind the proof of Theorems 9.1, 9.2, and 9.3. The argument can be divided into two main steps.

(I) Compare the asymptotic behavior of the discrete averages

$$(9.8) \quad \sum_{\gamma \in \Gamma: \log p(\gamma) \leq t} \phi(x\gamma)$$

with the asymptotic behavior of the continuous averages

$$(9.9) \quad \int_{g \in G: \log p(g) \leq t} \phi(xg) dm_G(g),$$

where m_G is an invariant measure on G .

(II) Establish the asymptotics for the continuous averages (9.9).

To achieve step (I), think of G as a fiber bundle over X , and show that the fibers of this bundle become equidistributed on the space G/Γ . More explicitly, we identify X with the homogeneous space $H \backslash G$ where H is a subgroup of G . For $u, v \in G$, we set

$$H_t[u, v] = \{h \in H : \log p(uhv) \leq t\}.$$

Then we show that the sum (9.8) can be approximated by a finite linear combination of integrals

$$\int_{H_{t_i}[u_i, v_i]} f_i(h^{-1}y_i) dm_H(h)$$

with suitably chosen $t_i \approx t$, $u_i, v_i \in G$, $f_i : G/\Gamma \rightarrow \mathbb{R}$, and $y_i \in G/\Gamma$, where m_H is a left-invariant measure on H . The asymptotic behavior of these integrals can be analyzed using either representation-theoretic techniques as in [GN] or the theory of unipotent flows as in [GW07]. In both cases we conclude that as $t \rightarrow \infty$,

$$\frac{1}{|H_t[u, v]|} \int_{H_t[u, v]} f(h^{-1}y) dm_H(h) \rightarrow \int_{G/\Gamma} f dm_{G/\Gamma},$$

where $m_{G/\Gamma}$ denotes the normalized invariant measure on G/Γ . Finally, this allows us to conclude that the above linear combination is approximately equal to the integral (9.9), which completes step (I).

In step (II), we establish an asymptotic formula for integral

$$v(t) = \int_{g \in G: \log p(g) \leq t} \phi(xg) dm(g),$$

where ϕ is a continuous subanalytic function on X with compact support. For this, in the case of Theorem 9.2 we consider its transform

$$f(s) = \int_0^\infty t^{-s} v(\log t) dt,$$

which converges when $\operatorname{Re}(s)$ is sufficiently large. Using a suitable version of resolution of singularities as in [P94], we show that $f(s)$ has a meromorphic continuation beyond the first pole. Then the asymptotic formula for $v(t)$ follows via a Tauberian argument (see also [CT10] for a similar argument).

The completion of the argument proving Theorem 9.3 is different and requires several arguments. A key role is played by the limiting density $d\nu_x(y)$, which is given by a positive kernel on $H \backslash G$. It has been shown in [GW07] (see also [GN]) to be given by the explicit formula

$$d\nu_x(y) = \left(\lim_{t \rightarrow \infty} \frac{m_H(H_t[\mathfrak{s}(x), \mathfrak{s}(y)])}{m_H(H_t)} \right) dm_{H \backslash G}(y),$$

where $\mathfrak{s} : H \backslash G \rightarrow G$ is a section of the natural projection. The stability and regularity properties of the sets H_t and of their volume growth, as well as the speed of convergence to the limiting density, must all be established in order to obtain quantitative convergence to the ergodic limit of the continuous averages (9.9). This latter estimate is then combined with a quantitative estimate on the difference between the continuous averages and the discrete averages (9.8). Such an estimate is obtained by applying a stable form of the quantitative ergodic theorem for H -actions with a spectral gap, which applies to the whole family of averages $\int_{H_{t_i}[u_i, v_i]} f_i(h^{-1}y_i) dm_H(h)$ acting on $L^2(G/\Gamma)$, introduced above. The stable ergodic theorem utilizes the fact that H is semisimple, so that the spectral transfer principle holds.

10. RECENT DEVELOPMENTS: BEST POSSIBLE DIOPHANTINE EXPONENTS

The problem of establishing exponents of intrinsic Diophantine approximation discussed Section 7 is an instance of a much more general problem, namely establishing a rate of distribution for dense lattice orbits on homogeneous spaces of the ambient group. In recent joint work with A. Ghosh [GGN3], [GGN4] a general approach to establishing quantitative density of orbits was developed. This approach is based on the duality principle discussed in Section 9, combined with the quantitative mean ergodic theorem, discussed in Section 4. It forms another instance of the dynamical correspondence mentioned in Section 7 based on the duality principle and, more explicitly, states that the quantitative density of lattice orbits $x\Gamma$ in the homogeneous space $xG = X$ is equivalent to the quantitative density of the corresponding orbits of H_x in the space $Y = G/\Gamma$, where H_x denotes the stabilizer of x in G . A remarkable feature of this approach is that it often yields best possible Diophantine exponents, a problem that has been inaccessible thus far for non-Abelian groups G , as far as the authors are aware.

To give a concrete example of this approach, let Γ denote a discrete subgroup of the algebraic group $G \subset \operatorname{SL}_d(\mathbb{R})$, and fix a norm on \mathbb{R}^d and $\operatorname{Mat}_d(\mathbb{R})$. Effective density of Γ -orbits amounts to estimating Diophantine exponents, which we define as follows.

Definition 10.1. Assume that for $x, x_0 \in \overline{x\Gamma}$, there exist constants $c = c(x, x_0)$ and $\epsilon_0 = \epsilon_0(x, x_0)$ such that for all $\epsilon < \epsilon_0$, the system of inequalities

$$\|x\gamma - x_0\|_\infty \leq \epsilon \quad \text{and} \quad \|\gamma\| \leq c\epsilon^{-\kappa}$$

has a solution $\gamma \in \Gamma$. Define the Diophantine approximation exponent $\kappa_\Gamma(x, x_0)$ as the infimum of $\kappa > 0$ such that the foregoing inequalities have a solution.

The Diophantine exponent admits a very general lower bound based on geometric arguments. Let us define the growth exponent of Γ -orbits in the variety X by

$$a_\Gamma(x) = \sup_{\text{compact } \Omega \subset X(\mathbb{R})} \limsup_{R \rightarrow \infty} \frac{\log N_R(\Omega, x)}{\log R},$$

where $N_R(\Omega, x)$ denotes the number of elements $\gamma \in \Gamma$ such that $\|\gamma\| \leq R$ and $x\gamma$ belongs to Ω . It is possible to show that for almost every $x_0 \in X$,

$$\kappa_\Gamma(x, x_0) \geq \frac{\dim(X)}{a_\Gamma(x)}.$$

The quantity $a_\Gamma(x)$ could be difficult to estimate in general, but if the variety X is homogeneous, it can be estimated in terms of volume growth of a suitable subgroup. X can then be identified with the homogeneous space $H \backslash G$, where H is a closed subgroup of G . We set

$$a(X) = \limsup_{t \rightarrow \infty} \frac{\log m_H(H_t)}{t},$$

where $H_t = \{h \in H; \log \|h\| \leq t\}$ and m_H is a right-invariant Haar measure on H . Then one can show using discreteness of Γ that, for every $x \in X$,

$$a_\Gamma(x) \leq a(X).$$

In particular, for almost every $x_0 \in X$,

$$(10.1) \quad \kappa_\Gamma(x, x_0) \geq \frac{\dim(X)}{a(X)}.$$

Thus the fundamental question that arises is to determine when this lower bound is in fact sharp (for almost all x), and in general to give an upper bound of $\kappa_\Gamma(x, x_0)$. This problem is closely related to the shrinking target problem for the orbit Hy in $Y = G/\Gamma$ with respect to the family of neighborhoods $\mathcal{O}_\epsilon(y_0)$, where $y_0 \in G/\Gamma$ is associated with $x_0 \in H \backslash G$.

The shrinking target problem can be approached using a quantitative mean ergodic theorem for the action of H on Y . Assume that the discrete group Γ has finite covolume in G and that the family of averaging operators $\pi_Y(\beta_t) : L^2(Y) \rightarrow L^2(Y)$ satisfies the quantitative mean ergodic theorem of Section 4, namely that there exist $C, \theta > 0$ such that for all sufficiently large t ,

$$(10.2) \quad \left\| \pi_Y(\beta_t)(\phi) - \int_Y \phi dm_Y \right\| \leq C m_H(H_t)^{-\theta} \|\phi\|_2,$$

where m_Y denotes the normalized Haar measure on Y . Let $\theta_\Gamma(X)$ denote the supremum over θ 's for which the estimate (7.6) holds. Then we have

Theorem 10.2 ([GGN3]). *For every $x_0 \in X$ and almost every $x \in X$,*

$$\kappa_\Gamma(x, x_0) \leq (2\theta_\Gamma(X))^{-1} \frac{\dim(X)}{a(X)}.$$

Moreover, the constants $c(x, x_0)$ and $\epsilon_0(x, x_0)$ appearing in Definition 10.1 are uniform over x, x_0 in compact sets.

Looking at the case $\theta_\Gamma(X) = 1/2$, we have the following sample conclusion.

Corollary 10.3 ([GGN3]). *If G and the stability group H are semisimple and non-compact and the representation of H on $L_0^2(\Gamma \backslash G)$ is a tempered representation of H , then the Diophantine exponent of Γ -action on $X = G/H$ is best possible and is given by*

$$\kappa_\Gamma(x, x_0) = \frac{\dim(X)}{a(X)}$$

for every x_0 and almost every x in X .

Finally, we note that the phenomenon of *subgroup temperedness*, whereby the restriction of the automorphic representation of a semisimple group G on $L_0^2(G/\Gamma)$ to a closed semisimple subgroup $H \subset G$ is a tempered representation of H , is a surprisingly common phenomenon which manifests itself in many interesting examples. For more on this phenomenon and for many examples of estimates and computations of the Diophantine exponent κ , we refer to [GGN3] and [GGN4].

ACKNOWLEDGMENTS

Part of this paper was written in Spring 2011 when the first author was visiting École Polytechnique Fédérale de Lausanne during the program “Group Actions in Number Theory”, and he would like to thank Emmanuel Kowalski, Philippe Michel, and the Centre Interfacultaire Bernoulli for their hospitality.

Both authors would like to express their gratitude to the Ergodic Theory Group at the Fédération Denis Poisson for the opportunity to explain the present work in the lecture series “*Théorie ergodique des actions de groupes*” held at the University of Tours in April 2011. In particular we would like to thank the organizers, Claire Anantharaman-Delaroche, Jean-Philippe Anker, and Emmanuel Lesigne.

We would like to thank Angel David Martinez for carefully reading a preliminary version of the paper and for supplying us with list of corrections.

The first author would also like to thank Manfred Einsiedler and the Institute for Mathematical Research at ETH, Zürich for their hospitality during his visit in Autumn 2012 when the work on this survey was completed.

ABOUT THE AUTHORS

Alexander Gorodnik is professor of mathematics at University of Bristol. He received his Ph.D. from the Ohio State University in 2003 and has held positions at University of Michigan and California Institute of Technology.

Amos Nevo is professor of mathematics at Technion, Israel Institute of Technology.

REFERENCES

- [Aa] Jon Aaronson, *An introduction to infinite ergodic theory*, Mathematical Surveys and Monographs, vol. 50, American Mathematical Society, Providence, RI, 1997. MR1450400 (99d:28025)
- [AAB] Claire Anantharaman, Jean-Philippe Anker, Martine Babillot, Aline Bonami, Bruno Demange, Sandrine Grellier, François Havard, Philippe Jaming, Emmanuel Lesigne, Patrick Maheux, Jean-Pierre Otal, Barbara Schapira, and Jean-Pierre Schreiber, *Théorèmes ergodiques pour les actions de groupes* (French), Monographies de L'Enseignement Mathématique [Monographs of L'Enseignement Mathématique], vol. 41, L'Enseignement Mathématique, Geneva, 2010. With a foreword in English by Amos Nevo. MR2643350 (2011h:47010)
- [Ar] Vladimir I. Arnold, *Arnold's problems*, Springer-Verlag, Berlin; PHASIS, Moscow, 2004. Translated and revised edition of the 2000 Russian original; With a preface by V. Philippov, A. Yakivchik and M. Peters. MR2078115 (2005c:58001)
- [AK63] V. I. Arnol'd and A. L. Krylov, *Uniform distribution of points on a sphere and certain ergodic properties of solutions of linear ordinary differential equations in a complex domain* (Russian), Dokl. Akad. Nauk SSSR **148** (1963), 9–12. MR0150374 (27 #375)
- [B02] Martine Babillot, *Points entiers et groupes discrets: de l'analyse aux systèmes dynamiques* (French, with English and French summaries), Rigidité, groupe fondamental et dynamique, Panor. Synthèses, vol. 13, Soc. Math. France, Paris, 2002, pp. 1–119. With an appendix by Emmanuel Breuillard. MR1993148 (2004i:37057)
- [B82] Hans-Jochen Bartels, *Nichteuklidische Gitterpunktprobleme und Gleichverteilung in linearen algebraischen Gruppen* (German), Comment. Math. Helv. **57** (1982), no. 1, 158–172, DOI 10.1007/BF02565853. MR672852 (84c:22013)
- [BHV] B. Bekka, P. de la Harpe and A. Valette, *Kazhdan's property (T)*. New Mathematical Monographs, 11. Cambridge University Press, Cambridge, 2008.
- [BO12] Yves Benoist and Hee Oh, *Effective equidistribution of S -integral points on symmetric varieties* (English, with English and French summaries), Ann. Inst. Fourier (Grenoble) **62** (2012), no. 5, 1889–1942, DOI 10.5802/aif.2738. MR3025156
- [B31] G. D. Birkhoff, *Proof of the ergodic theorem*, Proc. Nat. Acad. Sci. USA **17** (1931), 656–660.
- [BB13] Valentin Blomer and Farrell Brumley, *The role of the Ramanujan conjecture in analytic number theory*, Bull. Amer. Math. Soc. (N.S.) **50** (2013), no. 2, 267–320, DOI 10.1090/S0273-0979-2013-01404-6. MR3020828
- [BW80] Armand Borel and Nolan R. Wallach, *Continuous cohomology, discrete subgroups, and representations of reductive groups*, Annals of Mathematics Studies, vol. 94, Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1980. MR554917 (83c:22018)
- [BF11] Jean Bourgain and Elena Fuchs, *A proof of the positive density conjecture for integer Apollonian circle packings*, J. Amer. Math. Soc. **24** (2011), no. 4, 945–967, DOI 10.1090/S0894-0347-2011-00707-8. MR2813334 (2012d:11072)
- [Bo83] J. Bourgain, *Averages in the plane over convex curves and maximal operators*, J. Analyse Math. **47** (1986), 69–85, DOI 10.1007/BF02792533. MR874045 (88f:42036)
- [Bo89] Jean Bourgain, *Pointwise ergodic theorems for arithmetic sets*, Inst. Hautes Études Sci. Publ. Math. **69** (1989), 5–45. With an appendix by the author, Harry Furstenberg, Yitzhak Katznelson and Donald S. Ornstein. MR1019960 (90k:28030)
- [BFLM11] Jean Bourgain, Alex Furman, Elon Lindenstrauss, and Shahar Mozes, *Stationary measures and equidistribution for orbits of nonabelian semigroups on the torus*, J. Amer. Math. Soc. **24** (2011), no. 1, 231–280, DOI 10.1090/S0894-0347-2010-00674-1. MR2726604 (2011k:37008)
- [BGS06] Jean Bourgain, Alex Gamburd, and Peter Sarnak, *Sieving and expanders* (English, with English and French summaries), C. R. Math. Acad. Sci. Paris **343** (2006), no. 3, 155–159, DOI 10.1016/j.crma.2006.05.023. MR2246331 (2007b:11139)
- [BGS10] Jean Bourgain, Alex Gamburd, and Peter Sarnak, *Affine linear sieve, expanders, and sum-product*, Invent. Math. **179** (2010), no. 3, 559–644, DOI 10.1007/s00222-009-0225-3. MR2587341 (2011d:11018)

- [BK10] Jean Bourgain and Alex Kontorovich, *On representations of integers in thin subgroups of $SL_2(\mathbb{Z})$* , *Geom. Funct. Anal.* **20** (2010), no. 5, 1144–1174, DOI 10.1007/s00039-010-0093-4. MR2746949 (2012i:11008)
- [BMW99] R. W. Bruggeman, R. J. Miatello, and N. R. Wallach, *Resolvent and lattice points on symmetric spaces of strictly negative curvature*, *Math. Ann.* **315** (1999), no. 4, 617–639, DOI 10.1007/s002080050331. MR1731464 (2001f:11170)
- [BGM11] Roelof Wichert Bruggeman, Fritz Grunewald, and Roberto Jorge Miatello, *New lattice point asymptotics for products of upper half-planes*, *Int. Math. Res. Not. IMRN* **7** (2011), 1510–1559, DOI 10.1093/imrn/rnq120. MR2806513 (2012e:11095)
- [BS91] M. Burger and P. Sarnak, *Ramanujan duals. II*, *Invent. Math.* **106** (1991), no. 1, 1–11, DOI 10.1007/BF01243900. MR1123369 (92m:22005)
- [C53] A. P. Calderon, *A general ergodic theorem*, *Ann. of Math. (2)* **58** (1953), 182–191. MR0055415 (14,1071a)
- [CT10] Antoine Chambert-Loir and Yuri Tschinkel, *Igusa integrals and volume asymptotics in analytic and adelic geometry*, *Confluentes Math.* **2** (2010), no. 3, 351–429, DOI 10.1142/S1793744210000223. MR2740045 (2012d:11143)
- [C03] Laurent Clozel, *Démonstration de la conjecture τ* (French), *Invent. Math.* **151** (2003), no. 2, 297–328, DOI 10.1007/s00222-002-0253-8. MR1953260 (2004f:11049)
- [CW76] Ronald R. Coifman and Guido Weiss, *Transference methods in analysis*, American Mathematical Society, Providence, R.I., 1976. Conference Board of the Mathematical Sciences Regional Conference Series in Mathematics, No. 31. MR0481928 (58 #2019)
- [C] B. Conrad, *Modular forms and the Ramanujan conjecture*, Cambridge University Press, 2011.
- [C78] Michael Cowling, *The Kunze-Stein phenomenon*, *Ann. Math. (2)* **107** (1978), no. 2, 209–234. MR0507240 (58 #22398)
- [C79] Michael Cowling, *Sur les coefficients des représentations unitaires des groupes de Lie simples* (French), *Analyse harmonique sur les groupes de Lie (Sém., Nancy-Strasbourg 1976–1978)*, II, *Lecture Notes in Math.*, vol. 739, Springer, Berlin, 1979, pp. 132–178. MR560837 (81e:22019)
- [CN01] M. Cowling and A. Nevo, *Uniform estimates for spherical functions on complex semisimple Lie groups*, *Geom. Funct. Anal.* **11** (2001), no. 5, 900–932, DOI 10.1007/s00039-001-8220-x. MR1873133 (2002k:43005)
- [D85] S. G. Dani, *Divergent trajectories of flows on homogeneous spaces and Diophantine approximation*, *J. Reine Angew. Math.* **359** (1985), 55–89; erratum: *J. Reine Angew. Math.* **359** (1985), 214.
- [D86] S. G. Dani, *Bounded orbits of flows on homogeneous spaces*, *Comment. Math. Helv.* **61** (1986), no. 4, 636–660, DOI 10.1007/BF02621936. MR870710 (88i:22011)
- [D89] S. G. Dani, *On badly approximable numbers, Schmidt games and bounded orbits of flows*, *Number theory and dynamical systems (York, 1987)*, *London Math. Soc. Lecture Note Ser.*, vol. 134, Cambridge Univ. Press, Cambridge, 1989, pp. 69–86, DOI 10.1017/CBO9780511661983.006. MR1043706 (91d:58200)
- [D71] Pierre Deligne, *Formes modulaires et représentations l -adiques* (French), *Séminaire Bourbaki. Vol. 1968/69: Exposés 347–363*, *Lecture Notes in Math.*, vol. 175, Springer, Berlin, 1971, pp. Exp. No. 355, 139–172. MR3077124
- [D42] Jean Delsarte, *Sur le gitter fuchsien* (French), *C. R. Acad. Sci. Paris* **214** (1942), 147–179. MR0007769 (4,191a)
- [DH] Harold G. Diamond and H. Halberstam, *A higher-dimensional sieve method*, *Cambridge Tracts in Mathematics*, vol. 177, Cambridge University Press, Cambridge, 2008. With an appendix (“Procedures for computing sieve functions”) by William F. Galway. MR2458547 (2009h:11151)
- [Du03] W. Duke, *Rational points on the sphere*, *Ramanujan J.* **7** (2003), no. 1-3, 235–239, DOI 10.1023/A:1026203430418. Rankin memorial issues. MR2035804 (2005h:11094)
- [DRS93] W. Duke, Z. Rudnick, and P. Sarnak, *Density of integer points on affine homogeneous varieties*, *Duke Math. J.* **71** (1993), no. 1, 143–179, DOI 10.1215/S0012-7094-93-07107-4. MR1230289 (94k:11072)
- [EG74] Frederick P. Greenleaf and William R. Emerson, *Group structure and the pointwise ergodic theorem for connected amenable groups*, *Advances in Math.* **14** (1974), 153–172. MR0384997 (52 #5867)

- [EM93] Alex Eskin and Curt McMullen, *Mixing, counting, and equidistribution in Lie groups*, Duke Math. J. **71** (1993), no. 1, 181–209, DOI 10.1215/S0012-7094-93-07108-6. MR1230290 (95b:22025)
- [EMS96] Alex Eskin, Shahar Mozes, and Nimish Shah, *Unipotent flows and counting lattice points on homogeneous varieties*, Ann. of Math. (2) **143** (1996), no. 2, 253–299, DOI 10.2307/2118644. MR1381987 (97d:22012)
- [F06] John B. Friedlander, *Producing prime numbers via sieve methods*, Analytic number theory, Lecture Notes in Math., vol. 1891, Springer, Berlin, 2006, pp. 1–49, DOI 10.1007/978-3-540-36364-4_1. MR2277657 (2007k:11153)
- [FI] John Friedlander and Henryk Iwaniec, *Opera de cribro*, American Mathematical Society Colloquium Publications, vol. 57, American Mathematical Society, Providence, RI, 2010. MR2647984 (2011d:11227)
- [GJ78] Stephen Gelbart and Hervé Jacquet, *A relation between automorphic representations of $GL(2)$ and $GL(3)$* , Ann. Sci. École Norm. Sup. (4) **11** (1978), no. 4, 471–542. MR533066 (81e:10025)
- [GGN1] Anish Ghosh, Alexander Gorodnik, and Amos Nevo, *Diophantine approximation and automorphic spectrum*, Int. Math. Res. Not. IMRN **21** (2013), 5002–5058. MR3123673
- [GGN2] A. Ghosh, A. Gorodnik and A. Nevo, *Metric Diophantine approximation on homogeneous varieties*. Math arXiv:1205.4426. To appear in Compositio Mathematica.
- [GGN3] A. Ghosh, A. Gorodnik and A. Nevo, *Best possible rate of distribution of dense lattice orbits on homogeneous varieties*. In preparation.
- [GGN4] A. Ghosh, A. Gorodnik and A. Nevo, *Diophantine approximation exponents on homogeneous varieties*. To appear in Contemporary Math, 2014.
- [G83] Anton Good, *Local analysis of Selberg’s trace formula*, Lecture Notes in Mathematics, vol. 1040, Springer-Verlag, Berlin, 1983. MR727476 (85k:11026)
- [G03] Alexander Gorodnik, *Lattice action on the boundary of $SL(n, \mathbb{R})$* , Ergodic Theory Dynam. Systems **23** (2003), no. 6, 1817–1837, DOI 10.1017/S0143385703000154. MR2032490 (2005c:22015)
- [G04] Alexander Gorodnik, *Uniform distribution of orbits of lattices on spaces of frames*, Duke Math. J. **122** (2004), no. 3, 549–589, DOI 10.1215/S0012-7094-04-12234-1. MR2057018 (2005f:37009)
- [GM05] Alexander Gorodnik and Francois Maucourant, *Proximity and equidistribution on the Furstenberg boundary*, Geom. Dedicata **113** (2005), 197–213, DOI 10.1007/s10711-005-5539-8. MR2171305 (2006e:37015)
- [GN10] Alexander Gorodnik and Amos Nevo, *The ergodic theory of lattice subgroups*, Annals of Mathematics Studies, vol. 172, Princeton University Press, Princeton, NJ, 2010. MR2573139 (2011c:22006)
- [GN12a] Alexander Gorodnik and Amos Nevo, *Counting lattice points*, J. Reine Angew. Math. **663** (2012), 127–176, DOI 10.1515/CRELLE.2011.096. MR2889708
- [GN12b] Alexander Gorodnik and Amos Nevo, *On Arnold’s and Kazhdan’s equidistribution problems*, Ergodic Theory Dynam. Systems **32** (2012), no. 6, 1972–1990, DOI 10.1017/S0143385711000721. MR2995880
- [GN12c] A. Gorodnik and A. Nevo, *Lifting, restricting and sifting integral points on affine homogeneous varieties*, Compositio Math. 2012, doi:10.1112/S0010437X12000516.
- [GN] Alexander Gorodnik and Amos Nevo, *Ergodic theory and the duality principle on homogeneous spaces*, Geom. Funct. Anal. **24** (2014), no. 1, 159–244, DOI 10.1007/s00039-014-0257-8. MR3177381
- [GW07] Alex Gorodnik and Barak Weiss, *Distribution of lattice orbits on homogeneous varieties*, Geom. Funct. Anal. **17** (2007), no. 1, 58–115, DOI 10.1007/s00039-006-0583-6. MR2306653 (2008i:37012)
- [G] George Greaves, *Sieves in number theory*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 43, Springer-Verlag, Berlin, 2001. MR1836967 (2002i:11092)
- [Gr10] Andrew Granville, *Different approaches to the distribution of primes*, Milan J. Math. **78** (2010), no. 1, 65–84, DOI 10.1007/s00032-010-0122-7. MR2684773 (2011g:11178)
- [Gu10] Antonin Guilloux, *Polynomial dynamic and lattice orbits in S -arithmetic homogeneous spaces*, Confluentes Math. **2** (2010), no. 1, 1–35, DOI 10.1142/S1793744210000120. MR2649235 (2011i:37003)

- [G69] Yves Guivarc'h, *Généralisation d'un théorème de von Neumann* (French), C. R. Acad. Sci. Paris Sér. A-B **268** (1969), A1020–A1023. MR0251191 (40 #4422)
- [HR] H. Halberstam and H.-E. Richert, *Sieve methods*, Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], London-New York, 1974. London Mathematical Society Monographs, No. 4. MR0424730 (54 #12689)
- [H] Glyn Harman, *Prime-detecting sieves*, London Mathematical Society Monographs Series, vol. 33, Princeton University Press, Princeton, NJ, 2007. MR2331072 (2008k:11097)
- [HP01] Sa'ar Hersensky and Frédéric Paulin, *Hausdorff dimension of Diophantine geodesics in negatively curved manifolds*, J. Reine Angew. Math. **539** (2001), 29–43, DOI 10.1515/crll.2001.077. MR1863852 (2002i:53053)
- [HP02a] Sa'ar Hersensky and Frédéric Paulin, *Diophantine approximation in negatively curved manifolds and in the Heisenberg group*, Rigidity in dynamics and geometry (Cambridge, 2000), Springer, Berlin, 2002, pp. 203–226. MR1919402 (2003i:11112)
- [HP02b] Sa'ar Hersensky and Frédéric Paulin, *Diophantine approximation for negatively curved manifolds*, Math. Z. **241** (2002), no. 1, 181–226, DOI 10.1007/s002090200412. MR1930990 (2003g:53051)
- [HM79] Roger E. Howe and Calvin C. Moore, *Asymptotic properties of unitary representations*, J. Funct. Anal. **32** (1979), no. 1, 72–96, DOI 10.1016/0022-1236(79)90078-8. MR533220 (80g:22017)
- [HT] Roger Howe and Eng-Chye Tan, *Nonabelian harmonic analysis*, Universitext, Springer-Verlag, New York, 1992. Applications of $SL(2, \mathbf{R})$. MR1151617 (93f:22009)
- [H56] Heinz Huber, *Über eine neue Klasse automorpher Funktionen und ein Gitterpunktproblem in der hyperbolischen Ebene. I* (German), Comment. Math. Helv. **30** (1956), 20–62 (1955). MR0074536 (17,603b)
- [I78] Henryk Iwaniec, *Almost-primes represented by quadratic polynomials*, Invent. Math. **47** (1978), no. 2, 171–188. MR0485740 (58 #5553)
- [I96] Henryk Iwaniec, *The lowest eigenvalue for congruence groups*, Topics in geometry, Progr. Nonlinear Differential Equations Appl., vol. 20, Birkhäuser Boston, Boston, MA, 1996, pp. 203–212. MR1390315 (97e:11058)
- [IK] Henryk Iwaniec and Emmanuel Kowalski, *Analytic number theory*, American Mathematical Society Colloquium Publications, vol. 53, American Mathematical Society, Providence, RI, 2004. MR2061214 (2005h:11005)
- [JL70] H. Jacquet and R. P. Langlands, *Automorphic forms on $GL(2)$* , Lecture Notes in Mathematics, Vol. 114, Springer-Verlag, Berlin-New York, 1970. MR0401654 (53 #5481)
- [Jo93] Roger L. Jones, *Ergodic averages on spheres*, J. Anal. Math. **61** (1993), 29–45, DOI 10.1007/BF02788837. MR1253437 (95g:28031)
- [JR79] Andrés del Junco and Joseph Rosenblatt, *Counterexamples in ergodic theory and number theory*, Math. Ann. **245** (1979), no. 3, 185–197, DOI 10.1007/BF01673506. MR553340 (81d:10042)
- [KW82] Yitzhak Katznelson and Benjamin Weiss, *A simple proof of some ergodic theorems*, Israel J. Math. **42** (1982), no. 4, 291–296, DOI 10.1007/BF02761409. MR682312 (84i:28020)
- [K65] D. A. Každan, *Uniform distribution on a plane* (Russian), Trudy Moskov. Mat. Obšč. **14** (1965), 299–305. MR0193187 (33 #1408)
- [K67] D. A. Každan, *On the connection of the dual space of a group with the structure of its closed subgroups* (Russian), Funkcional. Anal. i Priložen. **1** (1967), 71–74. MR0209390 (35 #288)
- [KS03] H. Kim and P. Sarnak, *Refined estimates towards the Ramanujan and Selberg conjectures*, appendix in H. Kim, *Functoriality for the exterior square of $GL(4)$ and the symmetric fourth of $GL(2)$* . J. Amer. Math. Soc. **16** (2003), no. 1, 139–183.
- [K98a] Dmitry Y. Kleinbock, *Flows on homogeneous spaces and Diophantine properties of matrices*, Duke Math. J. **95** (1998), no. 1, 107–124, DOI 10.1215/S0012-7094-98-09503-5. MR1646538 (99k:11107)
- [K98b] Dmitry Y. Kleinbock, *Bounded orbits conjecture and Diophantine approximation*, Lie groups and ergodic theory (Mumbai, 1996), Tata Inst. Fund. Res. Stud. Math., vol. 14, Tata Inst. Fund. Res., Bombay, 1998, pp. 119–130. MR1699361 (2000k:11089)

- [K99] Dmitry Kleinbock, *Badly approximable systems of affine forms*, J. Number Theory **79** (1999), no. 1, 83–102, DOI 10.1006/jnth.1999.2419. MR1724255 (2001b:11064)
- [K01] Dmitry Kleinbock, *Some applications of homogeneous dynamics to number theory*, Smooth ergodic theory and its applications (Seattle, WA, 1999), Proc. Sympos. Pure Math., vol. 69, Amer. Math. Soc., Providence, RI, 2001, pp. 639–660, DOI 10.1090/pspum/069/1858548. MR1858548 (2002g:37009)
- [K10] Dmitry Kleinbock, *Quantitative nondivergence and its Diophantine applications*, Homogeneous flows, moduli spaces and arithmetic, Clay Math. Proc., vol. 10, Amer. Math. Soc., Providence, RI, 2010, pp. 131–153. MR2648694 (2011m:11144)
- [KM98] D. Y. Kleinbock and G. A. Margulis, *Flows on homogeneous spaces and Diophantine approximation on manifolds*, Ann. of Math. (2) **148** (1998), no. 1, 339–360, DOI 10.2307/120997. MR1652916 (99j:11083)
- [KM99] D. Y. Kleinbock and G. A. Margulis, *Logarithm laws for flows on homogeneous spaces*, Invent. Math. **138** (1999), no. 3, 451–494, DOI 10.1007/s002220050350. MR1719827 (2001i:37046)
- [KM] D. Kleinbock and K. Merrill, *Rational approximation on spheres*. arXiv:1301.0989.
- [K09] Alex V. Kontorovich, *The hyperbolic lattice point count in infinite volume with applications to sieves*, Duke Math. J. **149** (2009), no. 1, 1–36, DOI 10.1215/00127094-2009-035. MR2541126 (2011f:11125)
- [K] Alex Kontorovich, *From Apollonius to Zaremba: local-global phenomena in thin orbits*, Bull. Amer. Math. Soc. (N.S.) **50** (2013), no. 2, 187–228, DOI 10.1090/S0273-0979-2013-01402-2. MR3020826
- [KO11] Alex Kontorovich and Hee Oh, *Apollonian circle packings and closed horospheres on hyperbolic 3-manifolds*, J. Amer. Math. Soc. **24** (2011), no. 3, 603–648, DOI 10.1090/S0894-0347-2011-00691-7. With an appendix by Oh and Nimish Shah. MR2784325
- [K1] E. Kowalski, *The large sieve and its applications*, Cambridge Tracts in Mathematics, vol. 175, Cambridge University Press, Cambridge, 2008. Arithmetic geometry, random walks and discrete groups. MR2426239 (2009f:11123)
- [K2] E. Kowalski, *Sieve in expansion*, Séminaire Bourbaki, Exposé 1028 (November 2010)
- [K3] E. Kowalski, *Sieves in discrete groups*, especially sparse. Math. arXiv, 1207.7051, January 2013.
- [KO12] Alex Kontorovich and Hee Oh, *Almost prime Pythagorean triples in thin orbits*, J. Reine Angew. Math. **667** (2012), 89–131. MR2929673
- [KS60] R. A. Kunze and E. M. Stein, *Uniformly bounded representations and harmonic analysis of the 2×2 real unimodular group*, Amer. J. Math. **82** (1960), 1–62. MR0163988 (29 #1287)
- [K69] Bertram Kostant, *On the existence and irreducibility of certain series of representations*, Bull. Amer. Math. Soc. **75** (1969), 627–642. MR0245725 (39 #7031)
- [La95] Michael T. Lacey, *Ergodic averages on circles*, J. Anal. Math. **67** (1995), 199–206, DOI 10.1007/BF02787789. MR1383493 (97f:28045)
- [L89] Steven P. Lalley, *Renewal theorems in symbolic dynamics, with applications to geodesic flows, non-Euclidean tessellations and their fractal limits*, Acta Math. **163** (1989), no. 1-2, 1–55, DOI 10.1007/BF02392732. MR1007619 (91c:58112)
- [L65] Serge Lang, *Report on Diophantine approximations*, Bull. Soc. Math. France **93** (1965), 177–192. MR0193064 (33 #1286)
- [LP82] Peter D. Lax and Ralph S. Phillips, *The asymptotic distribution of lattice points in Euclidean and non-Euclidean spaces*, J. Funct. Anal. **46** (1982), no. 3, 280–350, DOI 10.1016/0022-1236(82)90050-7. MR661875 (83j:10057)
- [L01] Elon Lindenstrauss, *Pointwise theorems for amenable groups*, Invent. Math. **146** (2001), no. 2, 259–295, DOI 10.1007/s002220100162. MR1865397 (2002h:37005)
- [L99] François Ledrappier, *Distribution des orbites des réseaux sur le plan réel* (French, with English and French summaries), C. R. Acad. Sci. Paris Sér. I Math. **329** (1999), no. 1, 61–64, DOI 10.1016/S0764-4442(99)80462-5. MR1703338 (2000c:22009)
- [LP03] F. Ledrappier and M. Pollicott, *Ergodic properties of linear actions of (2×2) -matrices*, Duke Math. J. **116** (2003), no. 2, 353–388, DOI 10.1215/S0012-7094-03-11626-9. MR1953296 (2003j:37041)

- [L95] Jian-Shu Li, *The minimal decay of matrix coefficients for classical groups*, Harmonic analysis in China, Math. Appl., vol. 327, Kluwer Acad. Publ., Dordrecht, 1995, pp. 146–169. MR1355801 (98d:22009)
- [LZ96] Jian-Shu Li and Chen-Bo Zhu, *On the decay of matrix coefficients for exceptional groups*, Math. Ann. **305** (1996), no. 2, 249–270, DOI 10.1007/BF01444220. MR1391214 (97f:22029)
- [L44a] U. V. Linnik, *On the least prime in an arithmetic progression. I. The basic theorem* (English, with Russian summary), Rec. Math. [Mat. Sbornik] N.S. **15(57)** (1944), 139–178. MR0012111 (6,260b)
- [L44b] U. V. Linnik, *On the least prime in an arithmetic progression. II. The Deuring-Heilbronn phenomenon* (English, with Russian summary), Rec. Math. [Mat. Sbornik] N.S. **15(57)** (1944), 347–368. MR0012112 (6,260c)
- [LS10] Jianya Liu and Peter Sarnak, *Integral points on quadrics in three variables whose coordinates have few prime factors*, Israel J. Math. **178** (2010), 393–426, DOI 10.1007/s11856-010-0069-y. MR2733075 (2011k:11052)
- [Lu12] Alexander Lubotzky, *Expander graphs in pure and applied mathematics*, Bull. Amer. Math. Soc. (N.S.) **49** (2012), no. 1, 113–162, DOI 10.1090/S0273-0979-2011-01359-3. MR2869010 (2012m:05003)
- [LPS86] A. Lubotzky, R. Phillips, and P. Sarnak, *Hecke operators and distributing points on the sphere. I*, Comm. Pure Appl. Math. **39** (1986), no. S, suppl., S149–S186, DOI 10.1002/cpa.3160390710. Frontiers of the mathematical sciences: 1985 (New York, 1985). MR861487 (88m:11025a)
- [LPS87] A. Lubotzky, R. Phillips, and P. Sarnak, *Hecke operators and distributing points on S^2 . II*, Comm. Pure Appl. Math. **40** (1987), no. 4, 401–420, DOI 10.1002/cpa.3160400402. MR890171 (88m:11025b)
- [LRS95] W. Luo, Z. Rudnick, and P. Sarnak, *On Selberg’s eigenvalue conjecture*, Geom. Funct. Anal. **5** (1995), no. 2, 387–401, DOI 10.1007/BF01895672. MR1334872 (96h:11045)
- [Mag02] Akos Magyar, *Diophantine equations and ergodic theorems*, Amer. J. Math. **124** (2002), no. 5, 921–953. MR1925339 (2003f:37015)
- [M02] Gregory Margulis, *Diophantine approximation, lattices and flows on homogeneous spaces*, A panorama of number theory or the view from Baker’s garden (Zürich, 1999), Cambridge Univ. Press, Cambridge, 2002, pp. 280–310, DOI 10.1017/CBO9780511542961.019. MR1975458 (2004h:11031)
- [M04] Grigoriy A. Margulis, *On some aspects of the theory of Anosov systems*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2004. With a survey by Richard Sharp: Periodic orbits of hyperbolic flows; Translated from the Russian by Valentina Vladimirovna Szulikowska. MR2035655 (2004m:37049)
- [MNS00] G. A. Margulis, A. Nevo, and E. M. Stein, *Analogs of Wiener’s ergodic theorems for semisimple Lie groups. II*, Duke Math. J. **103** (2000), no. 2, 233–259, DOI 10.1215/S0012-7094-00-10323-7. MR1760627 (2001h:22008)
- [Ma02] F. Maucourant, *Approximation diophantienne, dynamique des chambres de Weyl et répartition d’orbites de réseaux*, PhD Thesis, Université de Lille, 2002.
- [Ma07] François Maucourant, *Homogeneous asymptotic limits of Haar measures of semisimple linear groups and their lattices*, Duke Math. J. **136** (2007), no. 2, 357–399, DOI 10.1215/S0012-7094-07-13626-3. MR2286635 (2007j:22022)
- [MW11] François Maucourant and Barak Weiss, *Lattice actions on the plane revisited*, Geom. Dedicata **157** (2012), 1–21, DOI 10.1007/s10711-011-9596-x. MR2893477
- [MW92] R. Miatello and N. R. Wallach, *The resolvent of the Laplacian on locally symmetric spaces*, J. Differential Geom. **36** (1992), no. 3, 663–698. MR1189500 (93i:58160)
- [Mol10] R. A. Mollin, *An overview of sieve methods*, Int. J. Contemp. Math. Sci. **5** (2010), no. 1-4, 67–80. MR2667834
- [Mo08] Yoichi Motohashi, *An overview of sieve methods and their history [translation of Sūgaku **57** (2005), no. 2, 138–163; MR2142054]*, Sugaku Expositions **21** (2008), no. 1, 1–32. Sugaku Expositions. MR2406271
- [N94] Amos Nevo, *Harmonic analysis and pointwise ergodic theorems for noncommuting transformations*, J. Amer. Math. Soc. **7** (1994), no. 4, 875–902, DOI 10.2307/2152735. MR1266737 (95h:22006)

- [N94b] Amos Nevo, *Pointwise ergodic theorems for radial averages on simple Lie groups. I*, Duke Math. J. **76** (1994), no. 1, 113–140, DOI 10.1215/S0012-7094-94-07605-9. MR1301188 (96c:28027)
- [N97] Amos Nevo, *Pointwise ergodic theorems for radial averages on simple Lie groups. II*, Duke Math. J. **86** (1997), no. 2, 239–259, DOI 10.1215/S0012-7094-97-08607-5. MR1430433 (98m:28041)
- [N98] Amos Nevo, *Spectral transfer and pointwise ergodic theorems for semi-simple Kazhdan groups*, Math. Res. Lett. **5** (1998), no. 3, 305–325, DOI 10.4310/MRL.1998.v5.n3.a5. MR1637840 (99e:28030)
- [N06] Amos Nevo, *Pointwise ergodic theorems for actions of groups*, Handbook of dynamical systems. Vol. 1B, Elsevier B. V., Amsterdam, 2006, pp. 871–982, DOI 10.1016/S1874-575X(06)80038-X. MR2186253 (2006k:37006)
- [NS10] Amos Nevo and Peter Sarnak, *Prime and almost prime integral points on principal homogeneous spaces*, Acta Math. **205** (2010), no. 2, 361–402, DOI 10.1007/s11511-010-0057-4. MR2746350 (2011m:22040)
- [NS94] Amos Nevo and Elias M. Stein, *A generalization of Birkhoff’s pointwise ergodic theorem*, Acta Math. **173** (1994), no. 1, 135–154, DOI 10.1007/BF02392571. MR1294672 (95m:28025)
- [NS97] Amos Nevo and Elias M. Stein, *Analogs of Wiener’s ergodic theorems for semisimple groups. I*, Ann. of Math. (2) **145** (1997), no. 3, 565–595, DOI 10.2307/2951845. MR1454704 (98m:22007)
- [NT97] Amos Nevo and Sundaram Thangavelu, *Pointwise ergodic theorems for radial averages on the Heisenberg group*, Adv. Math. **127** (1997), no. 2, 307–334, DOI 10.1006/aima.1997.1641. MR1448717 (98f:22005)
- [N03] Markus Neuhauser, *Kazhdan constants and matrix coefficients of $\mathrm{Sp}(n, \mathbf{R})$* , J. Lie Theory **13** (2003), no. 1, 133–154. MR1958578 (2003m:22015)
- [N02] Arnaldo Nogueira, *Orbit distribution on \mathbb{R}^2 under the natural action of $\mathrm{SL}(2, \mathbb{Z})$* , Indag. Math. (N.S.) **13** (2002), no. 1, 103–124, DOI 10.1016/S0019-3577(02)90009-1. MR2014978 (2005h:22034)
- [N10] Arnaldo Nogueira, *Lattice orbit distribution on \mathbb{R}^2* , Ergodic Theory Dynam. Systems **30** (2010), no. 4, 1201–1214, DOI 10.1017/S0143385709000558. MR2669417 (2011h:11076)
- [O98] Hee Oh, *Tempered subgroups and representations with minimal decay of matrix coefficients* (English, with English and French summaries), Bull. Soc. Math. France **126** (1998), no. 3, 355–380. Erratum, <http://gauss.math.yale.edu/~ho2/error.pdf>. MR1682805 (2000b:22015)
- [O02] Hee Oh, *Uniform pointwise bounds for matrix coefficients of unitary representations and applications to Kazhdan constants*, Duke Math. J. **113** (2002), no. 1, 133–192, DOI 10.1215/S0012-7094-02-11314-3. MR1905394 (2003d:22015)
- [O10] Hee Oh, *Orbital counting via mixing and unipotent flows*, Homogeneous flows, moduli spaces and arithmetic, Clay Math. Proc., vol. 10, Amer. Math. Soc., Providence, RI, 2010, pp. 339–375. MR2648698 (2011f:11079)
- [O11] H. Oh, *Harmonic analysis, ergodic theory and counting for thin groups*. arXiv:1208.4148.
- [O14] Hee Oh, *Apollonian circle packings: dynamics and number theory*, Jpn. J. Math. **9** (2014), no. 1, 69–97, DOI 10.1007/s11537-014-1384-6. MR3173439
- [P76] S. J. Patterson, *A lattice-point problem in hyperbolic space*, Mathematika **22** (1975), no. 1, 81–88. MR0422160 (54 #10152)
- [P94] Adam Parusiński, *Subanalytic functions*, Trans. Amer. Math. Soc. **344** (1994), no. 2, 583–595, DOI 10.2307/2154496. MR1160156 (94k:32006)
- [P95] Mark Pollicott, *A symbolic proof of a theorem of Margulis on geodesic arcs on negatively curved manifolds*, Amer. J. Math. **117** (1995), no. 2, 289–305, DOI 10.2307/2374915. MR1323676 (96k:58169)
- [R38] Frederick Riesz, *Some mean ergodic theorems*, J. London Math. Soc. **S1-13**, no. 4, 274, DOI 10.1112/jlms/s1-13.4.274. MR1574977
- [R] J. Rogawski, *Modular forms, the Ramanujan conjecture, and the Jacquet-Langlands correspondence*; appendix in A. Lubotzky, Discrete groups, expanding graphs and invariant measures. Progress in Mathematics, 125. Birkhäuser Verlag, Basel, 1994.

- [SV] A. Salehi Golsefidy and Péter P. Varjú, *Expansion in perfect groups*, *Geom. Funct. Anal.* **22** (2012), no. 6, 1832–1891, DOI 10.1007/s00039-012-0190-7. MR3000503
- [SS] Alireza Salehi Golsefidy and Peter Sarnak, *The affine sieve*, *J. Amer. Math. Soc.* **26** (2013), no. 4, 1085–1105, DOI 10.1090/S0894-0347-2013-00764-X. MR3073885
- [S95] Peter Sarnak, *Selberg’s eigenvalue conjecture*, *Notices Amer. Math. Soc.* **42** (1995), no. 11, 1272–1277. MR1355461 (97c:11059)
- [S03] Peter Sarnak, *Spectra of hyperbolic surfaces*, *Bull. Amer. Math. Soc. (N.S.)* **40** (2003), no. 4, 441–478, DOI 10.1090/S0273-0979-03-00991-1. MR1997348 (2004f:11107)
- [S05] Peter Sarnak, *Notes on the generalized Ramanujan conjectures*, *Harmonic analysis, the trace formula, and Shimura varieties*, *Clay Math. Proc.*, vol. 4, Amer. Math. Soc., Providence, RI, 2005, pp. 659–685. MR2192019 (2007a:11067)
- [S07] P. Sarnak, *Letter to Lagarias on integral Apollonian packings*, <http://www.math.princeton.edu/sarnak/>
- [S08] P. Sarnak, *Equidistribution and primes*, <http://www.math.princeton.edu/sarnak/>
- [S09] P. Sarnak, *Integral Apollonian Packings*, <http://www.math.princeton.edu/sarnak/>
- [S81] Klaus Schmidt, *Amenability, Kazhdan’s property (T), strong ergodicity and invariant means for ergodic group-actions*, *Ergodic Theory Dynamical Systems* **1** (1981), no. 2, 223–236. MR661821 (83m:43001)
- [Sch08] Eric Schmutz, *Rational points on the unit sphere*, *Cent. Eur. J. Math.* **6** (2008), no. 3, 482–487, DOI 10.2478/s11533-008-0038-4. MR2425007 (2009c:11112)
- [S56] A. Selberg, *Harmonic analysis and discontinuous groups in weakly symmetric Riemannian spaces with applications to Dirichlet series*, *J. Indian Math. Soc. (N.S.)* **20** (1956), 47–87. MR0088511 (19,531g)
- [S65] Atle Selberg, *The estimation of Fourier coefficients of modular forms*, *Proc. Sympos. Pure Math.*, Vol. VIII, Amer. Math. Soc., Providence, R.I., 1965, pp. 1–15. MR0182610 (32 #93)
- [St61] E. M. Stein, *On the maximal ergodic theorem*, *Proc. Nat. Acad. Sci. U.S.A.* **47** (1961), 1894–1897. MR0131517 (24 #A1367)
- [SW78] Elias M. Stein and Stephen Wainger, *Problems in harmonic analysis related to curvature*, *Bull. Amer. Math. Soc.* **84** (1978), no. 6, 1239–1295, DOI 10.1090/S0002-9904-1978-14554-6. MR508453 (80k:42023)
- [T67] A. A. Tempel’man, *Ergodic theorems for general dynamical systems* (Russian), *Dokl. Akad. Nauk SSSR* **176** (1967), 790–793. MR0219700 (36 #2779)
- [T92] Arkady Tempelman, *Ergodic theorems for group actions*, *Mathematics and its Applications*, vol. 78, Kluwer Academic Publishers Group, Dordrecht, 1992. Informational and thermodynamical aspects; Translated and revised from the 1986 Russian original. MR1172319 (94f:22007)
- [V] P. Varjú, *Random walks in Euclidean space*. arXiv:1205.3399.
- [vN32] J. von Neumann, *Proof of the quasi-ergodic hypothesis*, *Proc. Nat. Acad. Sci. USA* **18** (1932), 70–82.
- [W99] Michel Waldschmidt, *Density measure of rational points on abelian varieties*, *Nagoya Math. J.* **155** (1999), 27–53. MR1711387 (2000h:11077)
- [W39] Norbert Wiener, *The ergodic theorem*, *Duke Math. J.* **5** (1939), no. 1, 1–18, DOI 10.1215/S0012-7094-39-00501-6. MR1546100

SCHOOL OF MATHEMATICS, UNIVERSITY OF BRISTOL, BRISTOL, UNITED KINGDOM
E-mail address: a.gorodnik@bristol.ac.uk

DEPARTMENT OF MATHEMATICS, TECHNION, ISRAEL
E-mail address: anevo@tx.technion.ac.il