

Scaling entropy and automorphisms with purely point spectrum*

A. M. Vershik

To the memory of my friend Misha Birman

Abstract

We study the dynamics of metrics generated by measure-preserving transformations. We consider sequences of average metrics and ϵ -entropies of the measure with respect to these metrics. The main result, which gives a criterion for checking that the spectrum of a transformation is purely point, is that the *scaling sequence for the ϵ -entropies with respect to the averages of an admissible metric is bounded if and only if the automorphism has a purely point spectrum*. This paper is one of a series of papers by the author devoted to the asymptotic theory of sequences of metric measure spaces and its applications to ergodic theory.

Contents

1 Introduction

Among the many mathematical and nonmathematical problems we have been discussing with Misha Birman for many years after our acquaintance began in the early 1960s, the most intriguing one was the parallel between scattering theory and ergodic theory. Recently, I have returned to the (yet nonexistent) “ergodic scattering theory” and some forgotten questions related to it. However, this paper deals with another subject, which also correlates with M. Sh. Birman’s research.

*Partially supported by the grants RFBR-08-01-00379-a and RFBR-09-01-12175-ofi-m.

It is well known that the problem of deciding whether or not the spectrum of a given (say, differential) operator is discrete is quite difficult. Several remarkable early papers by M. Sh. Birman (in the first place, [?]) dealt with exactly this problem. In ergodic theory and the theory of dynamical systems, this problem (whether or not the system of eigenfunctions is complete) is also very difficult. In what follows, we use the terminology common in ergodic theory that slightly differs from the one adopted in operator theory: we use the term “discrete spectrum” as a synonym for “pure point spectrum.” This is justified by the fact that a discrete spectrum in the sense of operator theory almost never appears in the theory of dynamical systems, since almost always one deals with unitary operators.

As an example demonstrating the difficulty of this problem, we can mention the theory of substitutions, or stationary adic transformations [?], on which an extensive literature exists. The most intriguing problems concern nonstationary adic transformations with subexponential growth of the number of vertices in the corresponding Bratteli–Vershik diagram (see [?]). The simplest and most popular example of such a transformation is the Pascal automorphism defined in [?]; in this case, the measure space (i. e., the phase space of the dynamical system) is the space of infinite paths in the Pascal graph endowed with a Bernoulli measure, and the transformation sends a path to its successor in the natural lexicographic order. In spite of many efforts, we still do not know whether the corresponding unitary operator has a discrete (i. e., pure point) or mixed spectrum, or, which seems most likely, its spectrum is purely continuous. Attempts to directly construct its eigenfunctions failed; another approach, based on the characterization of systems with discrete spectrum in terms of Kushnirenko’s sequential entropy, [?] has not been carried out.

In the paper [?] (for more details, see [?, ?, ?]), we suggested a new notion, the so-called *scaling entropy*, which generalizes the notion of Kolmogorov’s entropy. The main point is that we suggest to *average the shifts of the metric* with respect to a given transformation and then compute the ε -entropy of the average metric. The class of increasing sequences of positive integers that normalize the growth of this ε -entropy over all admissible metrics does not depend on the choice of an admissible metric, so that the asymptotics of the growth of these sequences is a new metric invariant of automorphisms. Admissible metrics are measurable metrics satisfying some conditions that do or do not depend on the automorphism (see Section 2). It is important that *we consider not merely the ε -entropy of a metric, but the ε -entropy of*

a metric in a measure space. Admissible metrics play the same role as measurable generating partitions in the classical theory of Kolmogorov’s entropy according to Sinai’s definition. At first sight, the difference between using partitions and metrics looks rather technical: a partition determines a semimetric of a special form (a so-called “cut semimetric,” see [?]). However, our approach has two important differences from the classical theory. First, we use the ε -entropy of the iterated metric on a measure space rather than the entropy of a partition; this is a generalization of Kolmogorov’s entropy, which allows one to distinguish automorphisms with zero entropy. The second, more important, difference is that we use the average metric (rather than the supremum of metrics, which corresponds to the supremum of partitions), which has no interpretation in terms of partitions and which contains more information about the automorphism than the supremum of metrics.

In this paper, we give a necessary and sufficient condition for an automorphism to have a discrete spectrum in terms of the scaling sequence. The condition is that this sequence is bounded. This result generalizes a theorem of S. Ferenczi [?, ?], who considered the measure-theoretic complexity of symbolic systems by analogy with the ordinary complexity in symbolic dynamics. Our approach substantially differs from that of [?]: we consider an arbitrary admissible metric rather than the Hamming metric only, and, which is most important, introduce the average metric and show that it is admissible in many cases and, in particular, for the Hamming metric.

Thus criterion, i.e., the boundedness of the growth of the scaling entropy, should first be applied to adic transformations, e.g., to the Pascal automorphism (see [?]). Although it is not yet applied to checking that the spectrum of the Pascal automorphism is not discrete, the corresponding combinatorics is already developed and described in the recent paper [?], where a lower bound is obtained on the scaling sequence of the sup metric for the same Pascal automorphism. Supposedly, one can extend this bound ($\ln n$) to the average metric using the same techniques.

A wider context is presented in our papers [?, ?], where we suggest a plan for the study of the *dynamics of metrics in a measure space* as a source of new invariants of automorphisms. It is important that the notion of scaling entropy provides an answer to the question of whether or not the spectrum of a given transformation is discrete.

Sections 2–4 are of general nature and are intended not only for the purposes of this paper, which is devoted mostly to automorphisms with discrete spectrum. Here we introduce our main objects: admissible metrics, aver-

ages, scaling sequences, and the scaling entropy of an automorphism. In Section 5, we study the dynamics of metrics on a group and find conditions under which the average metric is admissible. The main result is given in Section 6, where we present a criterion for checking whether the spectrum of an automorphism is completely or partially discrete. In the last section, we sketch possible applications, links to the ordinary construction of entropy, and general remarks about the dynamics of metrics.

2 Admissible metrics

We consider various metrics and semimetrics on a measure space (X, \mathfrak{A}, μ) . In what follows, it is assumed to be a standard space with continuous measure μ and σ -algebra \mathfrak{A} of mod 0 classes of measurable sets, i.e., a Lebesgue space with continuous measure in the sense of Rokhlin (see [?]). The space $X \times X$ is endowed with the σ -algebra $\mathfrak{A} \times \mathfrak{A}$ and the measure $\mu \times \mu$.

We define a class of (semi)metrics on a measure space, which plays an important role in what follows.

Definition 1. *A measurable function $\rho : X \times X \rightarrow \mathbb{R}_+$ is called an admissible (semi)metric if*

1) ρ is a (semi)metric in the ordinary sense on a subset $X' \subset X$ of full measure ($\mu X' = 1$), i.e., $\rho(x, y) \geq 0$, $\rho(x, y) = \rho(y, x)$, $\rho(x, y) + \rho(y, z) \geq \rho(x, z)$ for all triples $(x, y, z) \in X' \times X' \times X'$, $(\mu \times \mu)\{(x, x), x \in X'\} = 0$, and

$$\int_X \int_X \rho(x, y) d\mu(x) d\mu(y) < \infty.$$

In order to formulate the next condition, observe that if ρ satisfies Condition 1), then the partition ψ_ρ of the space X into the classes of points $C_x = \{y : \rho(x, y) = 0\}$, $x \in X$, is measurable. Hence we have a well-defined quotient space $X_\rho \equiv X/\psi_\rho$ endowed with the quotient metric denoted by the same letter ρ and the quotient measure $\mu_{\psi_\rho} \equiv \mu_\rho$. For the quotient space (X_ρ, ρ, μ_ρ) , Condition 1) is still satisfied.

2) *The (completion of the) metric space (X, ρ) (if ρ is a metric) or (X_ρ, ρ) (if ρ on X is a semimetric) is a Polish (= metric, separable, complete) space with a Borel probability measure μ (respectively, μ_ρ).*

Following the measure-theoretic tradition, we must identify (semi)metrics (and hence the corresponding spaces) if they coincide almost everywhere as

measurable functions on the space $(X \times X, \mu \times \mu)$. Of course, a (semi)metric that coincides almost everywhere with an admissible (semi)metric is admissible.

Condition 2) means that the σ -algebra of Borel sets in the metric space (X, ρ) (or (X_ρ, ρ) in the case of a semimetric) is dense in the σ -algebra \mathfrak{A} of all measurable sets and, therefore, the measure μ (respectively, μ_ρ) is a Borel probability measure. It is obvious from the definition that a semimetric ρ is admissible if and only if the corresponding metric in the quotient space X_ρ is admissible.

An equivalent definition of an admissible metric is as follows: almost every pair of points can be separated by balls of positive measure, or, in other words, the σ -subalgebra generated by the open balls mod0 separates points of the space.

One can also formulate the admissibility condition in terms of the notion of a pure function from our paper [?]:

Lemma 1. *A metric ρ is admissible if and only if it satisfies Condition 1) and, regarded as a function of two variables, is pure in the sense of [?]; the latter means that for $\mu \times \mu$ -almost all pairs of points (x, y) , we have*

$$\mu\{z : \rho(x, z) \neq \rho(y, z)\} > 0.^1$$

If ρ is a semimetric, then this condition must hold for the metric ρ on the quotient space (X_ρ, ρ, μ_ρ) .

Indeed, the purity condition implies that the σ -algebra of sets generated by the balls separates points and hence is dense in the σ -algebra (X, \mathfrak{A}, μ) ; this also implies the separability. The converse immediately follows from the definition of an admissible metric.

It is well known (see [?]) that if (X, ρ) is a Polish space, then every nondegenerate Borel probability measure μ on X turns (X, ρ) into a Lebesgue space. In other words, the metric ρ on a Polish space (X, ρ) endowed with a Borel probability measure μ is an admissible metric on the space (X, μ) .

As in other our papers, in the definition of an admissible metric we reverse the tradition and *consider various metrics and semimetrics on a fixed measure space rather than various measures on a given metric space*. Recall that triples (X, ρ, μ) , consisting of a metric space endowed with a measure,

¹In other words, almost every point is uniquely determined by the collection of distances from this point to all points of some set of full measure (which may depend on the point).

in M. Gromov's book [?] were called *mm*-spaces, and in the paper [?], metric triples or Gromov triples.

It is useful to regard admissible metrics as *densities of some finite measures equivalent to the measure $\mu \times \mu$ on $X \times X$* :

$$dM_{\mu,\rho} = \rho(x, y)d\mu(x)d\mu(y).$$

If we set $\int_{X \times X} \rho(x, y)d\mu(x)d\mu(y) = 1$, which can be done by normalizing the metric, then the new measure is also a probability measure. Observe the following important property of admissible metrics implied by this interpretation.

Theorem 1. *For almost every point $x \in X$, there is a uniquely defined mod 0 conditional measure μ^x on X , which is given by the formula $d\mu^x(A) = \int_A \rho(x, y)d\mu(y)$ for every measurable set $A \subset X$. The family of measures $\{\mu^x; x \in X\}$ satisfies the condition*

$$\mu(A) = \int_X \mu^x(A)d\mu(x).$$

The metric ρ , regarded as a metric on the measure space (X, μ^x) , is admissible.

Proof. Consider the new measure $M_{\mu,\rho}$ on the space $X \times X$ and the measurable partition into the classes of points $C^x = \{(x, *)\} \subset X \times X$, and use the classical theorem on the existence of conditional measures (see, e.g., [?]), which implies the desired formula and the uniqueness mod 0 of the family of conditional measures. Now consider the space (X, ρ, μ^x) for a fixed x ; the metric ρ on this space is admissible since the admissibility of a metric is obviously preserved under replacing a measure with an equivalent one. \square

The conditional measure $\mu^x(A)$ can be interpreted as the “average distance,” or the conditional expectation of the distance from the set A to the point x .

In these terms, the condition that the metric, regarded as a function of two variables, is pure means that the conditional expectations do not coincide mod 0 for almost all pairs of points.

Now we can give a convenient criterion of admissibility.

Proposition 1. *A measurable function $\rho(\cdot, \cdot)$ satisfying Condition 1) of Definition 1 is admissible if and only if the following nondegeneracy condition holds:*

for μ -almost all points x , the measure of any ball of positive radius centered at x is positive: $\mu\{y : \rho(x, y) < \epsilon\} > 0$ for every $\epsilon > 0$, or, equivalently,

$$\mu^x([0, \epsilon]) > 0$$

for almost all x and any positive ϵ .

Proof. The fact that an admissible metric satisfies the condition in question was observed above. Now assume that this condition is satisfied. We must prove that the space (X, ρ) (if ρ is a metric) or the quotient space X_ρ (if ρ is a semimetric) is separable. It suffices to consider the case of a metric.

The condition stated in the proposition implies that for every ϵ there exists $\delta = \delta(\epsilon)$ with $\lim_{\epsilon \rightarrow 0} \delta(\epsilon) = 0$ such that some set of measure $> 1 - \delta$ contains a finite ϵ -net. But this means exactly that the space is separable and the measure is concentrated on a σ -compact set. \square

Several examples. 1. An important example is the following class of semimetrics, which was in fact intensively used in entropy theory. Every partition ξ of a space (X, μ) into finitely or countably many measurable sets gives rise to a semimetric:

$$\rho_\xi(x, y) = \delta_{(\xi(x), \xi(y))},$$

where $\xi(z)$ is the element of ξ that contains z . In this case, X_ρ is a finite or countable metric space. Such (finite) semimetrics are called *cuts*, and their linear combinations are called *cut semimetrics* (in the terminology of [?]). It is easy to see that cut (semi)metrics are admissible.

2. The very important metric defined by the formula $\rho(x, y) = \text{const}$ for $x \neq y$ determines a discrete uncountable space. We will call it the constant metric. *The constant metric on a space with continuous measure is not admissible.* In this case, the σ -algebra generated by the open sets is trivial and does not separate points of the space, i.e., a point is not determined by the collection of distances to the other points.

3. The condition defining an admissible metric can be strengthened by requiring, in condition 1), that

$$\int_X \int_X \rho(x, y)^p d\mu(x) d\mu(y) < \infty$$

for $p > 1$; in this case, we say that the metric is *p-admissible*. Let us say that ∞ -admissible metrics (semimetrics) are *bounded*; this class of admissible metrics will be most useful in what follows.

4. In combinatorial examples, it often suffices to consider metrics with which the space is compact or precompact (or, in the case of a semimetric, quasi-compact). For example, adic transformations [?, ?] act in the space of infinite paths of an \mathbb{N} -graded graph, which is a totally disconnected compact space.

The set of admissible metrics on a Lebesgue space (X, μ) with continuous measure is a convex cone $\mathcal{R}(X, \mu) = \mathcal{R}$ with respect to the operation of taking a linear combination of metrics with nonnegative coefficients. This cone is a canonical object (by the uniqueness of a Lebesgue space up to isomorphism) and plays a role similar to the role of the simplex of Borel probability measures in topological dynamics. In many cases, it suffices to consider admissible (semi)metrics that produce compact spaces (after completion), but we do not exclude the case of a noncompact space. One may consider different topologies on the cone \mathcal{R} ; the most natural of them is the weak topology, in which an ε -neighborhood of a metric ρ is the collection of metrics

$$\left\{ \theta : (\mu \times \mu) \{ (x, y) : |\rho(x, y) - \theta(x, y)| < \varepsilon \} > 1 - \varepsilon \right\}.$$

The property of being an admissible metric is invariant under measure-preserving transformations: if a metric ρ is admissible and a transformation T of the space (X, μ) preserves the measure μ , then the image ρ^T of ρ , defined by the formula $\rho^T(\cdot, \cdot) = \rho(T\cdot, T\cdot)$, is also admissible. Thus the group of (classes of) measure-preserving transformations acts on the cone \mathcal{R} in a natural way.

3 Average and maximal metrics

Let T be a measure-preserving transformation (in what follows, it will be an automorphism). When considering automorphisms or groups of automorphisms in spaces with admissible semimetrics, it is natural to assume that there exist an invariant set of full measure on which the (semi)metric is admissible in the sense of our definition.

In addition to the notion of admissible (semi)metrics, we define the class of *T-admissible metrics*. The *T*-admissibility condition must be invariant, in

the sense that if \mathcal{M} is the class of T -admissible metrics, then

$$V\mathcal{M} \equiv \{\rho : \rho(x, y) = \rho_1(Vx, Vy), \rho_1 \in \mathcal{M}\}$$

is the class of VTV^{-1} -admissible metrics in the same sense. Among many possible versions, we choose the class of admissible (semi)metrics for which T is a *Lipschitz transformation* almost everywhere: there exists a positive constant C such that for $(\mu \times \mu)$ -almost all pairs (x, y) , the condition

$$\rho(Tx, Ty) \leq C\rho(x, y)$$

holds; let us say that metrics from this class are Lipschitz T -admissible (semi)metrics.

The choice of an appropriate class depends on the problem under consideration and the properties of the automorphism. For example, in the case of adic automorphisms, it is most convenient to consider the class of Lipschitz metrics. This class can also be defined for countable groups of automorphisms.

For an arbitrary admissible semimetric ρ , we have defined the partition ψ_ρ of the space (X, μ) . In the same way, given an arbitrary automorphism T of the space (X, μ) , we consider the T -invariant partition $\psi_\rho^T = \bigvee_{k=0}^{\infty} T^k \psi_\rho$. We say that the semimetric ρ is *generating* for T if ψ_ρ^T is the partition into separate points mod0 (which we denote by ε). If ρ is the metric generated by a finite partition, then this is the ordinary definition of a generator (see [?]).

Let us define the average metric and the sup-metric for a given automorphism.

Definition 2. *Let T be an automorphism of a space (X, μ) , and let ρ be an admissible metric.*

The average metric ρ_n^T is defined by the formula

$$\hat{\rho}_n^T(x, y) = \frac{1}{n} \sum_{k=0}^{n-1} \rho(T^k x, T^k y).$$

The sup-metric is defined by the formula

$$\bar{\rho}_n^T(x, y) = \sup_{0 \leq k < n} \rho(T^k x, T^k y).$$

The following important result is a direct corollary of the pointwise ergodic theorem.

Theorem 2. *For any automorphism T and any admissible (semi)metric ρ , the limit of the sequence of average (semi)metrics ρ_n^T , which we denote by $\hat{\rho}$, exists almost everywhere in the space $(X \times X, \mu \times \mu)$:*

$$\hat{\rho}^T(x, y) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \rho(T^k x, T^k y);$$

$\hat{\rho}$ is a metric if and only if ρ is a metric or a generating semimetric.

The existence of the limit follows from the fact that the integral $\int_X \int_X \rho(x, y) d\mu(x) d\mu(y)$ is finite and the ergodic theorem.

Definition 3. *The metric*

$$\hat{\rho}^T(x, y) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \rho(T^k x, T^k y)$$

(the limit exists μ -a.e.) is called the average, or the l^1 -average of ρ with respect to the automorphism T .

The metric

$$\bar{\rho}^T(x, y) = \sup_{k \geq 0} \rho(T^k x, T^k y)$$

is called the limiting sup-metric of ρ with respect to T .

It is clear that $\hat{\rho}^T$ and $\bar{\rho}^T$ satisfy all conditions of the definition of a (semi)metric; in what follows, we will consider only admissible metrics or generating semimetrics ρ , so that $\hat{\rho}^T$, and $\bar{\rho}^T$ for an ergodic automorphism T , are metrics. The superscript T in the notation for the average and sup-metrics will be omitted if the automorphism is clear from the context. It is also obvious that $\hat{\rho} \leq \bar{\rho}$. However, the metrics $\hat{\rho}, \bar{\rho}$ may not be admissible even if the (semi)metric ρ is admissible.

In what follows, we will mainly consider the average metric $\hat{\rho}$. In most interesting cases, namely if T is weakly mixing, i.e., its spectrum in the orthogonal complement to the space of constants is continuous, this metric is constant and hence not admissible; the sup-metric may be constant even for automorphisms with discrete spectrum. However, in all cases we will

be interested not in the limiting metrics themselves, but in the *asymptotic behavior of the averages* ρ_n^T as $n \rightarrow \infty$. As we will see, automorphisms with discrete spectrum stand apart, since for them the average metric is often admissible.

It is clear that $\hat{\rho}$ is nothing else but the projection of the function ρ , regarded as an element of the space $L^1(X \times X, \mu \times \mu)$, to the subspace of $(T \times T)$ -invariant functions, i.e., the expectation of the metric ρ with respect to the subspace of invariant functions on $X \times X$. This space consists of constants if and only if T has no nontrivial eigenfunctions (in other words, T is weakly mixing). In this case, $\hat{\rho}$ is almost everywhere a constant, which is equal to the average ρ -distance between the points of the space X . At the same time, if T is not weakly mixing, then the spectrum of T contains a discrete component and $\hat{\rho}$ may be a nonconstant T -invariant (semi)metric. In this case, one may obtain bounds on $\hat{\rho}$ using Fourier analysis.

Here is an example of computing the average metric generated by the cut semimetric in the case of a rotation of the circle.

Example. Let $X = \mathbb{T}^1 = \mathbb{R}/\mathbb{Z}$, and let $\lambda \in X$ be an irrational number. Consider the semimetric $\rho(x, y) = |\chi_A(x) - \chi_A(y)|$, where χ_A is the indicator of a measurable set $A \subset X$; the metric $\rho(x, y)$ is T -admissible in the sense of our definition for the shift T_λ by any irrational number λ . The corresponding average metric is shift-invariant and looks as follows: $\hat{\rho}(x, y) = m[(A + x)\Delta\mu(A + y)] = m[A\Delta(A + x - y)]$; it is obviously admissible.

4 Entropy and scaling entropy

4.1 The ε -entropy of a measure in a metric space

Recall that the ε -entropy of a compact metric space (X, ρ) is the function $\varepsilon \mapsto H_\rho(\varepsilon)$ whose value is equal to the minimum number of points in an ε -net of (X, ρ) .

Definition 4. *The ε -entropy of a measure space (X, μ) with an admissible metric ρ is the function*

$$\varepsilon \mapsto H_{\rho, \varepsilon}(\mu) = \inf\{H(\nu) : k_\rho(\nu, \mu) < \varepsilon\},$$

where ν ranges over the set of all discrete measures with finite entropy and $k_\rho(\cdot, \cdot)$ is the Kantorovich metric on the space of Borel probability measures

on (X, ρ) . The entropy of a discrete measure $\nu = \sum_i c_i \delta_{x_i}$ is defined in the usual way: $H(\nu) = -\sum_i c_i \ln c_i$.

For our purposes, it is more convenient to use in the above definition another characteristic instead of $H(\mu)$, namely,

$$H'_{\rho, \varepsilon}(\mu) = \min \left\{ \ln k : \exists X', \mu(X') > 1 - \varepsilon, \exists \{x_i\}_1^k : X' \subset \bigcup_{i=1}^k V_\varepsilon(x_i) \right\},$$

where $V_\varepsilon(x)$ is the ε -ball centered at x ; thus $\{x_1, x_2, \dots, x_k\}$ is an ε -net in X' . The finiteness of H' follows from the fact that a Borel probability measure in a Polish space is concentrated, up to ε , on a compact set; H' is more convenient for computations than H .

In this paper, we use the following simple inequality.

Lemma 2. *For every compact metric space (X, ρ) and every nondegenerate Borel measure μ , the following inequality holds:*

$$H_{\rho, (d+1)\varepsilon}(\mu) \leq H'_{\rho, \varepsilon}(\mu),$$

where d is the diameter of the space.

Proof. Assume that the diameter of the compact space does not exceed 1. Assume that the measure of a set X' is greater than $1 - \varepsilon$ and $X' \subset \bigcup_{i=1}^k V_\varepsilon(x_i)$. Thus the points x_1, \dots, x_k form an ε -net in X' . Consider the discrete measure ν supported by the points x_1, \dots, x_k with charges $\nu(x_i)$ equal to the measures $\mu(V(x_i))$ of the corresponding balls (if two balls have a nonempty intersection, then we distribute the measure of the intersection proportionally between their centers). Choose an arbitrary point x_∞ and set its charge equal to $1 - \mu(X')$. Now consider the Monge–Kantorovich transportation problem with input measure μ and output measure ν . It is easy to see that we have in fact determined an admissible plan Ψ for this problem: the transportation from a point $x \in X'$ goes to the points x_i for which $x \in V_\varepsilon(x_i)$, and the remaining part of the measure μ on the set $X \setminus X'$ goes to the point x_∞ . It is easy to compute the cost of this plan; this gives a bound on the Kantorovich distance between the measures ν and μ :

$$k_\rho(\nu, \mu) \leq \varepsilon(1 - \varepsilon) + \varepsilon < 2\varepsilon.$$

On the other hand, we have $H(\nu) \leq \ln k = H'_{\rho, \varepsilon}(\mu)$. □

4.2 Scaling sequence and scaling entropy

Let us define the notion of a scaling sequence for the entropy of an automorphism. If an automorphism T is fixed, we omit the superscript in the notation for the average entropy and write simply $\hat{\rho}_n$.

Definition 5. *Let T be an automorphism of a Lebesgue space (X, μ) with a T -invariant measure μ . By definition, the class of scaling sequences for the automorphism T and a given (semi)metric ρ on X is the class, denoted by $\mathcal{H}_{\rho, \varepsilon}(T)$, of increasing sequences of positive numbers $\{c_n, n \in \mathbb{N}\}$ such that*

$$\mathcal{H}_{\rho, \varepsilon}(T) = \left\{ \{c_n\} : 0 < \liminf_{n \rightarrow \infty} \frac{H_{\hat{\rho}_n, \varepsilon}(\mu)}{c_n} \leq \limsup_{n \rightarrow \infty} \frac{H_{\hat{\rho}_n, \varepsilon}(\mu)}{c_n} < \infty \right\}.$$

In many cases, the class of scaling sequences for a given metric ρ does not depend on sufficiently small ε . In this case, it is obvious that all sequences $\{c_n\}$ from $\mathcal{H}_{\rho}(T)$ are equivalent.

Definition 6. *Assume that for a given ergodic automorphism T of a space (X, μ) there exists a (semi)metric ρ_0 such that the class of scaling sequences for ρ_0 is the maximal one (i.e., for any other (semi)metric, sequences $\{c_n\}$ from the corresponding class grow not faster than for ρ). In symbols, we write this fact as*

$$\mathcal{H}_{\rho_0}(T) = \sup_{\rho} \mathcal{H}_{\rho}(T).$$

Then we say that $\mathcal{H}_{\rho_0}(T)$ is the class of scaling sequences for the automorphism T and the metric ρ_0 is T -maximal.

It seems that such a metric exists for every automorphism. If we have chosen some T -maximal scaling sequence and the corresponding limit of entropies does exist, then it is called the *scaling entropy*.

Conjecture 1. *For every ergodic automorphism T , a generic T -admissible Lipschitz metric is T -maximal. In particular, for a K -automorphism (i.e., an automorphism with completely positive entropy), the scaling sequence is equivalent to the sequence $c_n = h(T)n$, where $h(T)$ is the entropy of T , for every Lipschitz metric.*

A preparatory result in this direction was obtained in [?].

In this paper, we will prove that for an automorphism with purely discrete spectrum and a T -admissible metric, the class $\mathcal{H}_{\rho}(T)$ of scaling sequences is the class of bounded sequences.

5 Invariant metrics on groups and averages of admissible metrics

5.1 Invariant metrics and discrete spectrum

Let us recall some known facts about ergodic automorphisms with discrete spectrum. It obviously follows from the character theory of commutative groups that the spectrum of a translation on a compact Abelian group is discrete. By the classical von Neumann theorem, the converse is also true: an ergodic automorphism with discrete spectrum is metrically isomorphic to the translation T on a compact Abelian group G endowed with the Haar measure m by an element whose powers form a dense subgroup:

$$x \mapsto Tx = x + g, \quad \text{Cl}\{ng, n \in \mathbb{Z}\} = G$$

(we use the additive notation). Note that on a compact group G there are many metrics that are invariant under the whole group of translations and determine the standard group topology. We will need the following assertion (which is, possibly, partially known).

Proposition 2. *The spectrum of an ergodic automorphism T of a measure space (X, μ) for which there exists a T -invariant admissible semimetric ρ contains a discrete component. Moreover, if ρ is a metric, then the spectrum of T is discrete and, consequently, T is isomorphic to a translation on a compact Abelian group.*

Proof. Since an admissible (semi)metric lies in the space $L^1(X \times X, \mu \times \mu)$, the tensor square of the operator U_T , which corresponds to the automorphism $T \times T$, has nonconstant eigenfunctions. This can happen only if the spectrum of the unitary operator U_T contains a discrete component, and the first claim is proved. In other words, T is an extension of some quotient automorphism with discrete spectrum, which may coincide with T itself. This means that T is a skew product over a base with discrete spectrum. Denote by $H \subset L^2(X, \mu)$ the subspace spanned by all eigenfunctions of U_T . All invariant functions of the operator $U_T \otimes U_T$ belong to the tensor square $H \otimes H$; hence these functions, regarded as functions of two variables, do not change when the argument ranges over a fiber of the skew product. But if these fibers are not single point sets, it follows that the metric does not distinguish points in fibers and hence is a semimetric. Thus if ρ is a metric, then each fiber

necessarily consists of a single point, the spectrum of T is purely discrete, and, by the ergodicity, T is isomorphic to a translation on a group. \square

Let us supplement this proof with an important refinement. Assume that the automorphism T has an orbit that is everywhere dense with respect to the metric ρ , i.e., T is topologically transitive (though we do not assume that it is a priori continuous). It is clear from above that this condition follows in fact from the existence of an invariant metric. Since the metric is admissible, we may assume without loss of generality that (X, ρ) is a Polish space. Consider the dense orbit $O = \{T^n x, n \in \mathbb{Z}\}$ of some point x . The restriction of ρ to O is a translation-invariant metric on the group \mathbb{Z} , and all translations are isometries. Hence the completion of O is an Abelian group to which we can extend the translations and their limits. Therefore X is a Polish monothetic group.² Obviously, the measure μ is invariant under the action of the closures of powers of T , i.e., it is an invariant probability measure on the whole group. Hence, by Weil's theorem, X is a compact commutative group and T is the translation by an element whose powers are everywhere dense.

5.2 Admissible invariant metrics

As we have already observed, for weakly mixing automorphisms, the average of every metric is constant, since there are no other invariant metrics. However, for automorphisms whose spectra contain a discrete component or are purely discrete, there are many invariant (semi)metrics. Hence, in order to study such automorphisms, we should investigate the question when the average metric for an automorphism T with a discrete spectrum is admissible.

Given a translation-invariant metric ρ on a compact commutative group with the Haar measure, consider the function $\phi_\rho(r) = \rho(x, x+r) = \rho(0, r)$. In the example from Section 3, it looked as $\phi(r) \equiv \hat{\rho}(x, x+r) = m[A\Delta(A+r)]$.

One can easily write down necessary and sufficient algebraic conditions on a measurable function ϕ that guarantee that it can be written as ϕ_ρ for a measurable invariant metrics:

$$\phi(\mathbf{0}) = 0, \phi(x) \geq 0, \phi(-x) = \phi(x), \phi(x) + \phi(y) \geq \phi(x + y).$$

²A topological group that contains a dense infinite cyclic subgroup is called monothetic. Note that there are many non-locally compact monothetic groups on which there is an invariant metric, but there is no invariant measure. A recent example is the Urysohn universal space regarded as a commutative group, see [?].

We will not need these conditions; the only important fact is that the admissibility condition from Lemma 1 can easily be reformulated in terms of this function.

Theorem 3. *An invariant measurable (semi)metric on a commutative compact group (satisfying Condition 1 from Definition 1) is admissible if and only if any of the following conditions holds.*

1. *The corresponding function ϕ_ρ is measurable, and*

$$\mu\{z : \phi_\rho(z) \neq \phi_\rho(g + z)\} > 0$$

for almost all g .

- 2.

$$\text{ess inf}_{g \in V \setminus \{0\}} \phi_\rho(g) = 0,$$

where V is an arbitrary neighborhood of the zero of the group in the standard topology.

Proof. Condition 1 is exactly equivalent to the condition of Lemma 1. It is useful to give a direct proof that this condition is necessary. Assume that it is not satisfied, i.e., for all elements g from some set of positive Haar measure, $\phi_\rho(z) = \phi_\rho(g + z)$ for almost all z . Therefore, the measurable function ϕ is constant on cosets of the subgroup generated by g . However, every set of positive Haar measure in a nondiscrete Abelian compact group on which there is an ergodic translation contains an element g that generates a dense cyclic subgroup. Then, since the function ϕ_ρ is measurable, it follows that it is constant almost everywhere. But if the function ϕ_ρ is constant, then the metric ρ is also constant and hence not admissible.

Let us prove that Condition 2 of the theorem is necessary. Consider the function ϕ_ρ for an invariant admissible metric ρ . Assume that $\theta = \text{ess inf}$ from Condition 2 is positive; then, using the invariance of ρ , we can construct a continuum of points lying at a fixed positive ρ -distance greater than θ , which contradicts the admissibility. For the same reason, the set of values of ϕ_ρ is dense in some neighborhood of the zero. The fact that Condition 2 is sufficient follows from Proposition 1. \square

The last assertion implies the following corollary.

Corollary 1. *For every admissible invariant metric there exists a sequence of group elements that converges to zero both in the standard topology and with respect to the (semi)metric.*

5.3 Admissibility of the average metric

Now we can explicitly write down the condition that guarantees the admissibility of the average, i.e., invariant, metric in terms of the original metric. Consider an arbitrary measurable (non necessarily admissible) metric ρ on a compact commutative group G and an ergodic translation T by an element g . Let us write down an expression for the average metric:

$$\begin{aligned}\hat{\rho}^T(x, y) &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^n \rho(x + kg, y + kg) = \int_G \rho(g + x, g + y) dm(g) \\ &= \int_G \rho(z, z + y - x) dm(z)\end{aligned}$$

(we have used the fact that the measure m is invariant). Hence the function $\phi(r)$, regarded as a function of r , is measurable and has the form

$$\phi(r) = \int_G \rho(z, z + r) dm(z) = \hat{\rho}(x, y); \quad y - x = r.$$

Obviously, $\hat{\rho}$ is a measurable function on the group $G \times G$. Now we can check the admissibility condition for the average metric $\hat{\rho}$.

Definition 7. *We say that an admissible metric ρ on a compact commutative group G with Haar measure m is semicontinuous at zero in mean if*

$$\liminf_{r \rightarrow 0} \int_G \rho(x, x + r) dm(x) = 0;$$

we say that it is semicontinuous at zero in measure if the following condition holds (in which “meas” means convergence in measure):

$$\liminf_{r \rightarrow 0} (\text{meas}) \rho(x, x + r) = 0.$$

Note that the second condition follows from the first one, and that both conditions are stated in purely group terms, i.e., do not depend on the particular translation T .

Thus we have the following admissibility criterion for the average metric.

Proposition 3. *The average metric $\hat{\rho}^T$ for an ergodic translation T on a compact commutative group G is admissible if and only if the original (semi)metric ρ is admissible and semicontinuous at zero in mean.*

Proof. The “if” part is proved above; the “only if” part follows from the equality $\phi_{\hat{\rho}}(r) = \int_G \rho(x, x+r) dm(x)$ and the previous proposition. \square

Now we are ready to formulate and prove the following important fact.

Theorem 4. *For every bounded admissible metric on a compact commutative group G , the average metric is admissible.*

Proof. Assume that the metric ρ is not semicontinuous at zero and there exists a positive number $c > 0$ such that

$$\liminf_{r \rightarrow 0} \int_G \rho(x, x+r) dm(x) > c.$$

Assume that the metric is normalized so that the diameter of the space X is equal to 1. Then it follows from our assumption that for sufficiently small (i.e., belonging to a small neighborhood of the zero in the group G) r there exists a (depending on r) subset in X of measure α , which does not depend on r and is greater than $\frac{c}{2}$, on which $\rho(x, x+r) > \frac{c}{2}$. Indeed, $1 \cdot \alpha + (1 - \alpha)\frac{c}{2} > \int_G \rho(x, x+r) dm(x) > c$. But since the group is compact, there is a set of positive measure for all points x of which the inequality $\rho(x, x+r) > \frac{c}{2}$ holds for all sufficiently small r from some set of positive measure. This in turn contradicts the admissibility of the metric ρ in the formulation of Proposition 1: arbitrarily small values $\phi(r)$ for small r cannot interlace with values greater than $\frac{c}{2}$ by the triangle inequality. Thus we have proved that ρ is semicontinuous. \square

Question. Does there exist an unbounded admissible metric ρ on the circle $S^1 = \mathbb{R}/\mathbb{Z}$ for which the average metric $\hat{\rho}$ is not admissible? Does there exist an unbounded p -admissible metric, with $1 < p < \infty$, on a compact Abelian group for which the average metric is not admissible?³

³While the paper was in press, F. Petrov and P. Zatitskiy proved that the average of any (in particular, unbounded) admissible metric on the circle is admissible. Thus the additional assumptions that the average metric is admissible in Proposition 2 and Theorem 5 are superfluous.

6 Criterion for the discreteness and continuity of the spectrum in terms of the scaling entropy

Now we formulate our main result.

Theorem 5. *For an ergodic automorphism with discrete spectrum realized as a translation on a compact commutative group with an arbitrary bounded admissible metric, or, more generally, with a metric for which the average metric is admissible, the scaling sequence is bounded.*

Proof. Since the set $B_\varepsilon = \phi^{-1}([0, \varepsilon])$ is, by definition, the ball of radius ε centered at $\mathbf{0} \in G$ in the metric $\hat{\rho}$, which is admissible by assumption, it follows that $mB_\varepsilon > 0$, since the metric is nondegenerate (see the definition of an admissible metric). But the sum $B_\varepsilon + B_\varepsilon$, like the sum $A + A$ for every set A of positive Haar measure in a locally compact group, contains a neighborhood V of the zero in the standard topology (see, e.g., [?]). It follows from the triangle inequality that $B_\varepsilon + B_\varepsilon \subset B_{2\varepsilon}$, so that $V \subset B_{2\varepsilon}$. Since $\varepsilon > 0$ is arbitrary, we see that *the topology on G determined by the average metric $\hat{\rho}$ coincides with the standard topology*, i.e., in the topology determined by $\hat{\rho}$, the group G is compact and contains a finite ε -net for every ε .

The pointwise a.e. convergence

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \rho(x + kg, y + kg) = \hat{\rho}_T(x, y),$$

which follows from the pointwise ergodic theorem, implies that the number of points in an ε -net for (G, ρ_n) tends to the number of points in an ε -net for $(G, \hat{\rho})$. This means that the sequence $H_{\rho_n, \varepsilon}(X)$ converges to $H_{\hat{\rho}, \varepsilon}(G)$. From the inequalities of Lemma 3 we see that the sequence $H_{\rho_n, \varepsilon}(\mu)$ is bounded as $n \rightarrow \infty$, and thus the scaling sequence for the automorphism T , which acts on the metric triple (G, ρ, m) , is bounded. \square

Combining this theorem with the previous one, we obtain the following result.

Theorem 6. *An ergodic automorphism T has a discrete spectrum if and only if the scaling sequence for T is bounded for some, and hence for every, bounded admissible metric.*

Proof. Above we have proved that if an automorphism has a discrete spectrum and the average metric is admissible, then the scaling sequence is bounded. But the average metric is always admissible provided that the original metric is bounded and admissible.

Assume that the scaling sequence is bounded for an automorphism T and an admissible metric ρ . Recall that the average metric is indeed a metric (and not a semimetric). Consequently, the space $(X, \hat{\rho})$ is precompact, and hence the metric $\hat{\rho}$ is admissible. Since it is T -invariant, it follows from Theorem 1 that T has a purely discrete spectrum. \square

Combining the last theorems with the previous results yields a criterion for the discreteness and continuity of the spectrum in terms of the automorphism T and an arbitrary admissible metric.

Theorem 7. *Let T be an ergodic automorphism, and let ρ be a bounded admissible (semi)metric. If the corresponding scaling sequence is not bounded, then the spectrum of T contains a continuous component. If the scaling sequence is not bounded for every admissible (semi)metric, then the spectrum of T is purely continuous.*

In the next section, we will show how one could apply this criterion.

7 Comparison with the traditional approach, the Pascal automorphism, and concluding remarks

7.1 Supremum metrics

The entropy theory of dynamical systems, developed mainly by Kolmogorov, Sinai, and Rokhlin, essentially uses the tools of the theory of measurable partitions. In Sinai's definition, the entropy appears as an asymptotic invariant of the dynamics of finite partitions under the automorphism:

$$\lim_n \frac{H(\bigvee_{k=0}^{n-1} T^k \xi)}{n} = h(T, \xi).$$

As a result of this theory, the study of the class of automorphisms with completely positive entropy was differentiated into a separate field, whose

methods do not apply to automorphisms with zero entropy. For example, one cannot obtain a new invariant for such automorphisms following the same scheme. This can be seen from the following simple fact.

Proposition 4. *For every transformation T and every increasing sequence of positive numbers $\{c_n, n \in \mathbb{N}\}$ satisfying the condition $\lim_n \frac{c_n}{n} = 0$, there exists a generating partition ξ such that*

$$\lim_n \frac{H(\bigvee_{k=1}^n T^k \xi)}{c_n} = \infty.$$

This means that the maximum growth of the entropy $H(\bigvee_{k=1}^n T^k \xi)$ either is linear (for automorphisms with positive entropy), or, in the case it is sublinear, it is arbitrarily close to linear for every automorphism. Thus we obtain no new information.

The metric corresponding to the supremum (product) $\xi_n = \bigvee_{k=1}^n T^k \xi$ of partitions is the supremum of the shifted metrics: $\bar{\rho}_n^T(x, y) = \sup_{0 \leq k < n} \rho(T^k x, T^k y)$. Hence, following our plan, we can use the ε -entropy of the metric $\bar{\rho}_n^T(x, y)$ instead of the entropy of the partition ξ_n itself. Then, using the definitions from Section 4, for a given metric ρ we can introduce an analog of the function $\mathcal{H}_{\rho, \varepsilon}(T)$ with the metric $\hat{\rho}_n$ replaced by $\bar{\rho}_n$:

$$\bar{\mathcal{H}}_{\rho, \varepsilon}(T) = \left\{ \{c_n\} : 0 < \liminf_{n \rightarrow \infty} \frac{H_{\bar{\rho}_n, \varepsilon}(\mu)}{c_n} \leq \limsup_{n \rightarrow \infty} \frac{H_{\bar{\rho}_n, \varepsilon}(\mu)}{c_n} < \infty \right\}.$$

In this way we define the class of *sup-scaling sequences* \bar{c}_n for a given metric ρ . This also allows us to extend the classical entropy theory following the above scheme. Though it is somewhat easier to deal with the sup-metric than with the average metric, the former is much less useful than the latter. The metric $\bar{\rho}$ more often happens to be constant for an automorphism with discrete spectrum, while, as we have seen, $\hat{\rho}$ is always admissible if the original metric is bounded. Let us illustrate the important difference between the operations of taking the average and supremum metrics by the following example.

Example. Let T be an irrational rotation of the unit circle, and let ρ be the semimetric corresponding to a generating two-block partition (i.e., a partition into two sets of positive measure); the semimetric ρ is T -admissible, hence, as we have seen, $\hat{\rho}$ is an invariant admissible metric. At the same time, $\bar{\rho}$ is the constant metric. This means that the scaling sequence c_n is bounded, but \bar{c}_n

(the scaling sequence for the sup-metric) is not; namely, we have $\bar{c}_n \sim \ln n$. Thus the difference manifests itself even in the case of a discrete spectrum.

Does it make sense to use intermediate averages, e.g., the l^p -averages

$$\lim \left[\frac{1}{n} \sum_{k=0}^{n-1} \rho(T^k x, T^k y)^p \right]^{\frac{1}{p}} = \hat{\rho}^p(x, y)$$

for $p \in (1, \infty)$, instead of l^1 ? Apparently, they do not lead to any new effects: these metrics behave in the same way as the l^1 -average metric. For instance, in the above example, the p -average of the metric determined by a two-block partition of the unit circle is $\hat{\rho}^p(x, y) = \{m[A\Delta(A+x-y)]\}^{\frac{1}{p}}$. For $p = \infty$ (i.e., the sup-metric), the picture is completely different, as in other interpolation theories. Thus in entropy theory, the use of average metrics substantially supplements the classical considerations.

7.2 Application of the discreteness criterion

As noted above, the problem of determining whether or not the spectrum of an automorphism is discrete, is not at all simple. Theorems 5–7 provide convenient nonspectral criteria for checking that the spectrum is not purely discrete; for this, one should bound the entropy from below for one admissible metric satisfying the conditions of Section 2 by a sequence that grows arbitrarily slow with n .

One of the intriguing examples of automorphisms for which the discreteness of the spectrum has not been neither proved nor disproved since the 1980s is the Pascal automorphism. It was introduced by the author in 1980 (see [?, ?]) as an example of an *adic transformation*, and is defined as a natural transformation in the space of paths in the Pascal graph regarded as a Bratteli–Vershik diagram with lexicographic ordering of paths. One can give a short combinatorial description of this transformation by encoding these paths with sequences of zeros and ones and identifying the space of paths with the compact space $X = \{0; 1\}^\infty = \mathbf{Z}_2$. Then the Pascal automorphism is defined by the formula

$$T(\{1^i 0^j 1 * * \}) = \{0^{j-1} 1^{i+1} 0 * * \};$$

here $i \geq 0$, $j > 0$, and the domain of T and T^{-1} is the whole X except for the countable set of sequences having finitely many zeros or ones. The most

natural metric on X is the 2-adic metric $\rho(\{x_k\}, \{y_k\}) = 2^{-n}$, where n is the first digit with $x_k \neq y_k$. This metric is admissible, and the Pascal automorphism satisfies the Lipschitz condition almost everywhere. The orbits of this automorphism coincide with the orbits of the action of the infinite symmetric group. The Bernoulli measures are T -invariant. The spectrum of the Pascal transformation was studied in the papers [?, ?, ?, ?], where some interesting properties were established (e.g., it was proved that T is loosely Bernoulli, the complexity of T was computed, etc.), but the question about the type of the spectrum remains open. In [?, ?] it was conjectured that the study of the behavior of scaling sequences may turn to be useful. The corresponding plan was carried out in [?], but in that paper a logarithmic lower bound was obtained on the scaling sequence for the sup-metric, and not for the average metric; this is not sufficient for the conclusion that the spectrum is not discrete. Nevertheless, one may hope that the combinatorics developed in [?] will help to prove that the scaling sequence is unbounded also for the ε -entropy of the average metric, which, by our theorem, would imply that the spectrum is not discrete. There are many adic transformations similar to the Pascal automorphism for which the same question is also of great interest. For example, if we replace the Pascal graph with its multidimensional analog or the Young graph, we will obtain automorphisms that supposedly have continuous spectra. As observed above, in order to prove that there are no nontrivial eigenfunctions, one should obtain a growing lower bound on the scaling sequence not for one, but for all (or for some representative set of) bounded admissible (semi)metrics.

7.3 The dynamics of metrics

Recall that the general approach that consists in studying the asymptotic behavior of metrics is not exhausted by considering the asymptotics of the ε -entropy of the average or supremum metric, i.e., does not reduce to studying the growth of scaling sequences; this is only its simplest version. In fact, we consider the original measure space (X, \mathfrak{A}, μ) with an action of an automorphism T (or a group of automorphisms G), fix an appropriate metric ρ , and study the sequence of metric triples

$$(X, \rho_n^T, \mu), \quad \text{where } \rho_n^T(x, y) = \frac{1}{n} \sum_{k=0}^{n-1} \rho(T^k x, T^k y).$$

The conjecture is that, for a fixed measure and a fixed automorphism (or group of automorphisms), the asymptotic properties of this sequence of metric triples do not depend (or weakly depend) on the choice of an individual admissible metric from a wide class. These properties include not only the scaling entropy, but also more complicated characteristics of the sequence, say the mutual properties of several consecutive metric triples. Since the classification of metric triples up to measure-preserving isometry is known (see [?, ?]), one may hope to apply it to this problem. In this field there are many traditional and nontraditional questions. For example, what is the distribution of the fluctuations of the sequence of average metrics, regarded as functions of two variables on $(X \times X, \mu \times \mu)$, as they converge to the constant metric (for weakly mixing transformations, e.g., K -automorphisms)? What can be said about the asymptotic properties of neighboring pairs of metric triples (with indices n and $n + 1$)? Etc.

In conclusion, it is worth mentioning that the concept of scaling entropy appeared in connection with the classification of filtrations in [?] and was used in [?]. In terms of the present paper, the scaling entropy for filtrations, i.e., decreasing sequences of measurable partitions or σ -algebras, is the scaling entropy for an action of a locally finite group such as $\sum \mathbb{Z}/2$ instead of an action of \mathbb{Z} considered here. The definitions we have given for an action of \mathbb{Z} essentially coincide with those given in [?] for locally finite groups.

References

- [1] M. S. Birman, *The spectrum of singular boundary problems*, Mat. Sb., N. Ser. **55(97)**, No. 2, (1961), 125–174. English Translation in: *Eleven Papers on Analysis*, Amer. Math. Soc. Transl. **53**, 23–80, AMS, Providence, RI, 1966.
- [2] A. M. Vershik, *Uniform algebraic approximation of shift and multiplication operators*. Dokl. Akad. Nauk SSSR **259**, No. 3, 526–529 (1981). English translation: Sov. Math. Dokl. **24**, 97–100 (1981).
- [3] A. M. Vershik, *A theorem on periodic Markov approximation in ergodic theory*. Zapiski Nauchn. Semin. LOMI **115**, 72–82 (1982). English translation: J. Sov. Math. **28**, 667–674 (1985).

- [4] A. M. Vershik, *Dynamics of metrics in measure spaces and their asymptotic invariants*. Markov Proc. Related Fields, **16**, No. 1, 169–185 (2010).
- [5] M. Gromov, *Metric Structures for Riemannian and Non-Riemannian Spaces*. Birkhauser, 1999.
- [6] A. M. Vershik, *Classification of measurable functions of several variables and invariantly distributed random matrices*, Funkts. Anal. Prilozh. **36**, No. 2, 12–28 (2002). English translation: Funct. Anal. Appl. **36**, No. 2, 93–105 (2002).
- [7] A. M. Vershik, *Random metric spaces and universality*, Uspekhi Mat. Nauk **59**, No. 2(356), 65–104 (2004). English translation: Russian Math. Surveys **59**, No. 2, 259–295 (2004).
- [8] A. M. Vershik, *Information, entropy, dynamics*, in: *Mathematics of the 20th Century: A View from Petersburg* [in Russian], MCCME, 2010, pp. 47–76.
- [9] P. Cameron and A. M. Vershik, *Some isometry groups of Urysohn spaces*. Ann. Pure Appl. Logic **143**, No. 1–3, 70–78 (2006).
- [10] A. G. Kushnirenko, *Metric invariants of entropy type*, Uspekhi Mat. Nauk **22**, No. 5, 57–65 (1967). English translation: Russian Math. Surv. **22**, No. 5, 53–61 (1967).
- [11] A. M. Vershik, *Orbit theory, locally finite permutations, and Morse arithmetics*, Contemp. Math. **532**, 115–136 (2010).
- [12] J. Feldman, *r -entropy, equipartition and Ornstein’s isomorphism theorem in R^n* . Israel J. Math. **36**, 321–345 (1980).
- [13] M. Ratner, *Some invariants of Kakutani equivalence*. Israel J. Math. **38**, 232–240 (1981).
- [14] A. M. Vershik, *Dynamic theory of growth in groups: Entropy, boundaries, examples* Uspekhi Mat. Nauk **55**, No. 4(334), 59–128 (2000). English translation: Russian Math. Surv. **55**, No. 4, 667–733 (2000).
- [15] S. Ferenczi, *Measure theoretic complexity of ergodic systems*, Israel J. Math. **100**, 189–207 (1997).

- [16] S. Ferenczi and K. Park, *Entropy dimensions and a class of constructive examples*, Discr. Cont. Dyn. Syst. **17**, 133–141 (2007).
- [17] M. Keith, a private communication.
- [18] M. M. Deza and M. Laurent, *Geometry of Cuts and Metrics*, Springer, 1997.
- [19] A. A. Lodkin, I. E. Manaev, and A. R. Minabutdinov, *Asymptotic behavior of the scaling sequence of the Pascal adic transformation*, Zapiski Nauchn. Semin. POMI **378**, 58–72 (2010). English translation to appear in J. Math. Sci.
- [20] B. Glasner, B. Tsirelson, and B. Weiss, *The automorphism group of the Gaussian measure cannot act pointwise*, Israel J. Math. **148**, 305–329 (2005).
- [21] S. Bailey, M. Keane, K. Petersen, and I. Salama, *Ergodicity of the adic transformation on the Euler graph*, Math. Proc. Camb. Phil. Soc. **141**, 231–238 (2006).
- [22] K. Peterson and A. Varchenko, *The Euler adic dynamical system and path counts in the Euler graph*, to appear in Tokyo J. Math., 2010.
- [23] K. Peterson and X. Mela, *Dynamical properties of the Pascal adic transformation*, Ergodic Theory Dynam. Systems **25**, 227–256 (2005).
- [24] V. A. Rokhlin, *On the fundamental ideas of measure theory*, Mat. Sb., N. Ser. **25(67)**, No. 2, 107–150 (1949).
- [25] V. A. Rokhlin, *Lectures on the entropy theory of measure-preserving transformations*, Uspekhi Mat. Nauk **22**, No. 5(137), 3–56 (1967). English translation: Russian Math. Surveys **22**, No. 5, 1–52 (1967).
- [26] E. Janvresse and T. de la Rue, *The Pascal adic transformation is loosely Bernoulli*, Ann. Inst. H. Poincaré Prob. Stat. **40**, No. 2, 133–139 (2004).
- [27] A. M. Vershik and A. D. Gorbulsky, *Scaled entropy of filtrations of sigma-fields*, Teor. Veroyatnost. i Primenen **52**, No. 3, 446–467 (2007). English translation: Probab. Theory Appl. **52**, No. 3, 493–508 (2008).

- [28] A. Weil, *L'integration dans les groupes topologiques et ses applications*.
Paris, 1940.